교육학석사학위논문

Learning Stress Cues for Word Segmentation by Korean Learners of English

한국인 영어 학습자들의 단어 분절 및 인지를 위한 강세 단서 학습

2016 년 8 월

서울대학교 대학원

외국어교육과 영어전공

권 나 연

Learning Stress Cues for Word Segmentation by Korean Learners of English

by

NA-YEON KWON

A thesis Submitted to
the Department of Foreign Language Education
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts in Education

At the
Graduate School of Seoul National University
August 2016

Learning Stress Cues for Word Segmentation by Korean Learners of English

한국인 영어 학습자들의 단어 분절 및 인지를 위한 강세 단서 학습

지도교수 안 현 기

이 논문을 교육학 석사 학위논문으로 제출함
2016년 6월

서울대학교 대학원
외국어교육과 영어전공
권 나 연

권나연의 석사학위논문을 인준함
2016년 8월

위 원 장 ______________________
부위원장 ______________________
위 원 ______________________

Learning Stress Cues for Word Segmentation by Korean Learners of English

APPROVED BY THESIS COMMITTEE:

SEONGEUN LEE, COMMITTEE CHAIR

YOONSUK MO

HYUNKEE AHN

Abstract

It is known that English native speakers recognize the onset of a word at the stressed syllable (Metrical Segmentation Strategy) but in a large volume of studies, Korean learners of English are not sensitive to stress, and thereby they can't the use stress cue for segmentation effectively.

The current study investigated whether Korean learners of English can learn to segment a stream of speech by recognizing the stressed syllable as the onset of a word based on statistical regularity, especially transitional probability (TP).

The experiments were conducted for 10 - 13 aged learners of English living in Seoul and divided into two parts: Experiment 1 for learning effects depending on stress patterns and the ways of learning, Experiment 2 for the stress pattern preference of Korean learners of English.

In Experiment 1, stimuli were designed with 3 sets of 18 pseudo-words composed of 3 syllables (CVCVCV) and conducted pre- and post- tests. Pre-test was proceeded on TP-only stimuli and the word-spotting test. In the post-test, learning sections were included depending on the stress patterns

(word-initial or word-final) and the ways of learning (listening or repeating aloud) and then the participants took the word-spotting test.

In Experiment 2, all the participants took the preference test in which they chose what they recognized as a word after listening to 18 chains of a repeated pseudo-word stressed at the initial syllable.

As a result of Experiment 1, in the case of being given TP-only cue, Korean learners of English could segment the speech by using TP but this was interfered with when they learned word-initial stressed words by listening. However, when they learned word-initial stressed words by repeating aloud, the accuracy increased and this effect was similar to the case that they learned word-final stressed words by listening.

Experiment 2 was investigated whether Korean learners of English showed any preference for stress pattern at word-level. The result showed that approximately 89.7% of the participants tended to recognize a word by word-final stress.

Taken together, the current study implies that producing experience can facilitate the accurate perception and segmentation in a new language, which can be explained by Motor Theory. Especially, this result suggests the possibility that segmentation at the word-initially stressed syllable could be

learned through producing experience by L2 learners. Furthermore, Korean learners might apply the L1 prosody (L-H tone pattern) to a new language prosodic pattern, which could be interpreted by PAM model or word-final lengthening effects.

A number of studies have identified that Korean learners have trouble using the stress cue for segmentation since there is not word level stress in Korean. The results of the current study imply that producing experience is a necessary process not only for speaking but also for perception of the speech. In particular, in order to learn the word level stress cue which does not exist in Korean, the experiences of the direct articulation and the memory of the feature in producing play an important role for the accurate segmentation and perception. This might be a useful way of improving listening skills in the English as a Foreign Language environment such as Korea where listening input is exceedingly limited.

Key Words: Speech perception, Word segmentation, Statistical learning, Metrical Segmentation Strategy, Motor theory

Student Number: 2014-20899

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1.

INTRODUCTION

This chapter introduces the motivation and purpose of the current study. Section 1.1 explains the purpose of the study, and Section 1.2 presents the research questions. Section 1.3 outlines the organization of the thesis.

1.1. The Purpose of the Study

It has been studied that experience-dependent factors play an important role in infants' language acquisition. In particular, recognizing words and learning them by sensing word boundaries and segmenting in fluent speech are the first strategy of acquiring a language. It has been suggested that infants can distinguish word boundaries as they implicitly learn the statistical regularities occurring in the continuous streams of speech (Saffran, Newport, & Aslin, 1996).

A great number of studies suggested that infants can acquire not only words but syntactic structures by tracking statistical regularities called Transitional Probability(TP) (Gomez & Gerken, 1999; Marcus et al., 1999).

Besides, suprasegmental features such as tone (Saffran et al., 1999) and stress (Thiessen & Saffran, 2007) can also be acquired based on TP.

Stress is one of the acoustic cues which affects segmentation in a stress-timed language such as English. According to Cutler et al. (1995), English native speakers consider a stressed syllable as the onset of a word and intuitively segment the speech ahead of it. They named the strategy as Metrical Segmentation Strategy (MSS), which is an efficient means of segmentation since a majority of English words have stress on the first syllable of a word (Cutler & Carter, 1987).

MSS is not a natural born ability of English native infants but an acquired skill by statistical information (Thiessen & Saffran, 2007). Thiessen & Saffran (2007) investigated whether English native infants have natural preference for specific stress pattern and could implicitly learn iambic stress with TP. They suggested English native infants didn't have a biased stress preference, but learned that a stressed syllable means a start of new information through a great deal of input and experience.

Meanwhile, some studies reported that Korean learners of English do not effectively use stress cues for segmentation (Kim & Cho, 2009; Kim, 2004; Kim, Broersma, & Cho, 2012) because Korean is a syllable-timed language and Koreans are not sensitive to stress in English which is stress-timed. Especially, the studies claimed that Koreans have more difficulties to perceive and use word-initial stress (Kim & Nam, 2011; Kim et al., 2012) than word-final stress. Therefore, it would be much harder for Korean learners of English to acquire and use MSS for segmentation efficiently.

The current study investigates whether it is possible that stress based strategy for segmentation, particularly word-initial stress could be acquired for Korean learners.

To this end, the research question starts with whether Korean young learners of English can segment a stream of speech by using the statistical regularity as native infants do. It is general that Korean learners of English know how to segment fluent speech by sensing words they already know. Since Korea is an EFL (English as a Foreign Language) environment where people don't speak English publically and have limited input, Korean learners lack enough opportunities of listening to learn how to segment speech implicitly. Therefore, if they could acquire efficient strategies using cues for segmentation, Korean young learners of English would be expected to recognize words more efficiently.

The other purpose of the current study is to examine whether Korean learners can learn how to use stress pattern cues and whether voice production is helpful to learn the stress based cues for segmentation. As mentioned above, since Korea is a quite input-limited environment for acquiring English, a range of assistant methods to learn the prosodic cues for segmentation could be useful, and Motor learning could support them. According to Motor Theory studies (Liberman & Mattingly, 1985; Mechsner et al., 2001; Repp, 1987), phonetic information is perceived through the phonetic module in which the relationship between articulation and acoustic patterns is constructed, hence, it is expected that the listener's experience of articulation can facilitate the ability to perceive the speech. In consequence, the current study investigates whether Korean learners of English can learn how to use the cues for segmentation and whether they can learn MSS by through speech production.

1.2. Research Questions

The current study intends to examine whether Korean learners of English could sense implicitly the statistical information of the words in the stream of continuous speech and learn how to use stress pattern for segmentation. In addition, it also investigates whether articulation could play a learning aid role in acquiring the ability to use stress pattern cues which don't exist in Korean.

To deal with those issues, experiments are conducted with following research questions:

1. Can Korean L2 learners sense statistical regularities in a string of speech for segmentation?
2. Do Korean L2 learners have preference of stress pattern for segmentation?
3. Can Korean L2 learners learn stress pattern cues for segmentation? (Especially word-initial stress pattern)
4. Can the experience of articulation in learning affect the enhancement of the perception ability?

1.3. Organization of the Thesis

The present study consists of five chapters. Chapter 1 introduces the purpose of the current study and presents the research questions. Chapter 2 presents an overview of the literature review on segmentation cues, statistical learning and the motor theory. In chapter 3, the methodology of the study is explained with the apparatus, participants, stimuli, procedures, and the data collection and analysis. Chapter 4 presents the results and discusses the research findings. Finally, Chapter 5 concludes the research with the summary of the major findings and the pedagogical implications of the present study and the suggestions for further studies.

Chapter 2.

LITERATURE REVIEW

This chapter consists of the literature overview of segmentation in language perception and involved learning theories. Section 2.1 discusses general word segmentation cues in speech perception and Section 2.2 explains statistical cues and Transitional Probability for segmentation in detail. Stress pattern cues are dealt with in Section 2.3. In Section 2.4, the relation between perception and production in language learning as well as motor theory that production can facilitate perception is discussed.

2.1. Word Segmentation Cues in Speech Perception

It is important to identify where the start or the end of the words is from a continuous speech because there are no salient boundaries between words and spoken words are rarely distinguished by pauses (Cole, Jakimik, & Cooper, 1980). Thus, not only native speakers but non-native speakers use a range of cues to sense word boundaries. For example, lexical-semantic cue is one of them. Native speakers distinguish *two lips* from *tulips* depending on the context when hearing *two lips* (Gow Jr & Gordon, 1995). Non-native speakers also use lexical-semantic cues for segmentation even though their lexical entry is limited (White, Melhorn, & Mattys, 2010). White et al. (2010) found Hungarian learners of English rely more on lexical cues than other cues like stress patterns for segmentation to the possible extent.

Acoustic-phonetic cues such as aspiration, glottal stop, or creaky voice are also reliable. For instance, people distinguish word boundaries like *keeps talking* and *keep stalking* using aspiration and *an ice man* and *a nice man* using glottal stop or creaky voice (Altenberg, 2005; Ito & Strange, 2009; Nakatani & Dukes, 1977).

Phonotactic constraints are also used as segmentation cues. McQueen (1998) found that people recognize word boundaries faster and easier when they can expect words aligned to their native language phonotactics. This effect influences on L2 segmentation, too. Weber et al.(Weber, 2000; Weber & Cutler, 2006) found that the German listeners who are also proficient in English detect embedded English words in nonsense sequences as well as the native English listeners do, which means that the English proficient Germans get the capability to recognize the words violating L2 phonotactics. Phonotactics in probability is another useful means for segmentation. Vitevitch and Luce (1998, 1999, 2005) claim that probabilistic phonotactics have been related to facilitatory effects on speech perception, especially with non-words. Just for phoneme awareness, frequent phonotactic information would be helpful cues.

Although little research was conducted, syntactic structure also can provide implying cues for segmentation. Cole, Jakimik, and Cooper (1980) studied that mispronounced syllables were detected fast when subjects were provided because listeners can expect syllables of the word in a given sentence. For example, in the following pair of sentences:

(1) The doctors said he knows drops will help the cold.

(2) The doctors said that nosedrops will help the cold.

In the continuous speech of this sentence, *he* in sentence (a) provides a cue for listeners to expect the general form of a following verb ending with –s and segment the words.

Context is also used to segment the speech and recognize words. In the same study of Cole et. al. (1980), when listening the sentence such as *"They saw the carko on the ferry"*, listeners segmented and recognize the mispronounced word *carko* as "cargo" or "car go" depending on the guiding short story.

In addition, statistical cues may be the first strategy humans use to learn a language by detecting more frequently co-occurring sounds and stress patterns also play a pivotal role as markers of language-specific word pattern. Among aforementioned segmentation cues, statistical cues and stress pattern cues will be more thoroughly in the following section.

2.2. Statistical Cues for Word Segmentation

The statistical learning ability of humans is to notice statistical regularities from the world around them to learn about the environment. There are also statistical regularities on sound sequences that comprise words, which compares with accidental sound sequences that occur across word boundaries over a corpus of speech (Hayes & Clark, 1970).

Transitional probability (TP) is one resource of the measurable statistical information. Thiessen and Saffran (2003) showed that infants can intuitively know the statistical relationship between adjacent syllables. Saffran et al. (1996) explained this notion using transitional probability and TP Y given X is computed as followed:

(a)	(b)	(c)
$\dfrac{\text{frequency of pair } XY}{\text{frequency of } X}$	$\dfrac{\text{frequency of } bay.bi}{\text{frequency of } bay}$	$\dfrac{\text{frequency of } bay\#too}{\text{frequency of } bay}$

Fig. 2.1 (a) TP calculation, (b) TP of bi given bay in a word,
(c) TP in a word external pair

Transitional probabilities are greater within a word than across words. For instance, in English word *pretty*, the syllable *pre* can be followed by a set of syllables like *ty, tend*, and *vent*. The expectable frequency of the syllable *pre* preceding *ty* is roughly 80% in the infant's language environment. In the

English phrase *pretty baby*, however, the final syllable *ty* can appear before any syllable of other English words, resulting in the extremely low probability of *ty+ba* (roughly 0.03% in speech to infants). Because of the difference in sequential probabilities of the syllables, *pretty* is more likely to be a word than *tyba* (Saffran, Aslin, & Newport, 1996).

Another interesting factor of statistical learning is prior knowledge. The studies investigating word segmentation of infants have shown that older infants may ignore TPs when they conflict with learned probabilistic cues to word boundaries that exist in the infants' native language Thiessen and Saffran (2003). Adult language learners also seem to ignore TPs when prior knowledge conflicts with new information (Finn & Hudson Kam, 2008). The researchers studied a TP use of adult learners in a continuous chain of pseudowords CCVCV (i.e. /kmodu/) violating English phonotactics, and they found that adults' knowledge of their native language system can interfere with statistical learning when their L1 knowledge collides with statistical cues to word boundaries of that of L2. It would be interpreted that when old information and the new information are incompatible; it is helpful for infants to rely on new information but for adults to depend on the old knowledge, because infants have little knowledge about any language but adults have much knowledge and already have a command of at least one language successfully. However, when clear pause is given between the words, adult learners recognize the word individually. Thus, if learning new words containing stressed syllable and chained with pause between them, learners would use both TP and stress cue for segmentation, and this learning experience of combining cues would increase word recognition.

2.3. Stress Pattern Cues for Word Segmentation

Stress is a powerful segmentation cue for the languages such as English and German. Cutler and Carter (1987) investigated a corpus of 190,000 words of spontaneous British English conversation and found that 90% of lexical words in it began with stressed syllables (in the corpus, about 60% of the lexical words were monosyllables; 28% were polysyllables with initial primary stress; 3% were polysyllables with initial secondary stress). In addition, Clopper (2002) investigated sum of frequency relating word size and stress position in words from the Hoosier Mental Lexicon (HML), an online version of Webster's Pocket Dictionary that has 20,000 words and the Brown University database. As shown in the Table 2.1, a majority of stress tends to be on the first syllables of the words.

Table 2.1 Frequency information for syllable-stress patterns in English database (Clopper, 2002)

Syll-Stress Pattern	Word Count	Sum Frequency
2syl-1pri	3624	67693
2syl-2pri	995	19881
3syl-1pri	2619	24558
3syl-2pri	1510	15278
3syl-3pri	369	1398
4syl-1pri	497	3549
4syl-2pri	1331	9014
4syl-3pri	1017	6831
4syl-4pri	37	97

* syl. for syllable, pri. for primary stress

Therefore, considering strong stress as the beginning of a new word which is known as Metrical Segmentation Strategy (MSS) (Cutler & Norris, 1988) is a reasonable strategy for the efficient word perception as English speakers do (Cutler & Butterfield, 1992).

However, stress pattern is a language specific feature. For example, Korean and French are syllable-timed language and Japanese is mora-timed one. Moreover, the stress patterns of Spanish are different from those of

English. Thus, either learning a novel segmentation cue in L2 English or applying the L1 stress pattern to L2 speech would be difficult for L2 learners. Sanders et al. (2002) studied whether L2 learners use lexical, syntactic, and stress pattern cues for segmentation and if they do, on which cues they rely the most. They asked native Japanese and native Spanish late learners of English to answer whether specific sound fell at the beginning or in the middle of the words in English sentences. In their study, they found that late-learners used the lexical cues to perform the segmentation task, as native English speakers do. Non-native speakers didn’t use syntactic information as natives do but the late learners employed stress pattern cues for segmentation, although their tendency depended on the stress pattern characteristics of their native language. They found that Japanese late learners were slightly better to use stress pattern for segmentation than Spanish late learners. They interpreted that perhaps those Japanese learners withheld their L1 segmentation cues for L2 segmentation, which led them to learn the stress pattern, the new segmentation cue. The researchers could not rule out the possibility that Japanese L1 would have the segmentation cue which can facilitate the use of stress cue in English because there is little known about segmentation cues to Japanese language. On the other hand, Spanish learners who were confused with the similar prosodic elements but different patterns from those of their native language.

Korean does not have stress in word level and it is segmented by rhythmic cues such as intonation patterns (Jun, 1993, 1998; Kim & Cho, 2009; Kim, 2004; Sohn, 2001) or final lengthening (Cho & Keating, 2001; Oh, 1998) at the Accentual Phrase (AP) level. According to Jun (1998), Korean has two prosodic units of phrase level, Accentual Phrase (AP) and Intonational Phrase (IP). English has different units which are Intermediate Phrase (ip) and Intonational Phrase (IP). Two IPs in both languages plays a similar role in speech perception of which domain is in the intonation level, which is lengthened at the phrase-final and related to semantic processing. What shows the different characteristics is Korean AP and English ip. English ip has High or Low phrase accent, while Korean AP is marked by High phrase-final boundary tone. That is, it could be expected that Korean learners would have tendency to consider lengthened or high F0 sound as a signal of the phrase-final and short or mostly low F0 sound as a signal of the phrase-initial in a sentence, and it was born out (Kim, 2004).

In a large volume of studies which dealt with the segmentation problems of Korean L2 learners of English, the reason Korean learners have trouble using the segmentation cues to English comes from those language specific attributes. For instance, Guion (2005) investigated the relationship between the age of acquisition and the language prosody acquisition of English stress patterns with early and late Korean-English bilinguals. She found that especially late bilinguals showed non-nativelike effect of lexical class and preference for stress placement on final long vowels. She interpreted the results came from the fact that Korean prosodic prominence is determined at the AP not word-level and it is realized by a tone pattern. Especially late Korean-English bilinguals were exposed to this phrase-level prosody first, which would interfere with learning word-level prosody.

In addition, Kim and Nam (2011) also found that high F0 was recognized as word-final for the Korean and segmentation was hindered when stress was on the initial position at words. Consequently, Korean learners could not use the stress pattern of English or MSS effectively for segmentation and it is estimated that the cue is difficult to be learned.

2.4. Relation between Perception and Production

"Perception leads production" is an obvious notion when discussing speech learning. That might be because perception plays a role as input or resources for production. However, the relationship between perception and production is complex one in language learning and processing.

In the phonological development of the native language, perception skills appeared to be completed in advance of production skills. Stange & Broen (1981) found that 3-year-old children who did not reach to the level of articulating /r/ and /l/ perfectly showed less stable perception than children who produced all phonemes correctly. Further, Flege (1991) mentioned that in a certain level of L2 learning, nonnative speakers might perceive an L2 sound just like native speakers, but they cannot produce the exactly the same sound with the prosody as the natives. It is not expected, however, that the opposite pattern: nonnative speakers who can produce an L2 sound authentically fail to perceive it. This implies that the ability to perceive a certain sound developed before the ability to produce it. This

Although the production ability might not develpe before the perception ability, the idea that production can facilitate perception was also investigated in many studies.

Moter Theory is the representative theory based the notion that production influences perception. Stating in a brief way, the motor theory of speech perception claims that speech perception involves access to the speech motor system (Liberman et al., 1967), which means perception has relation to production system in speech. In early motor theory, Liberman (1957) focused on a behavioral relation through the mimicry of infants when they hear and claimed that infants' mimicry leads to association between articulation and sensory consequences in one way, and the acoustic signals in the other way. In his study, he claimed that the listener's perception of /d/ in two syllables /di/ and /du/ comes from the listener's acquired articulation similarity. As it were, associating different acoustic signals for the syllables with the same response makes listeners consider they hear the same sound /d/. Thus, producing what they heard can be a shortcut for a clear perception of the similar sounds.

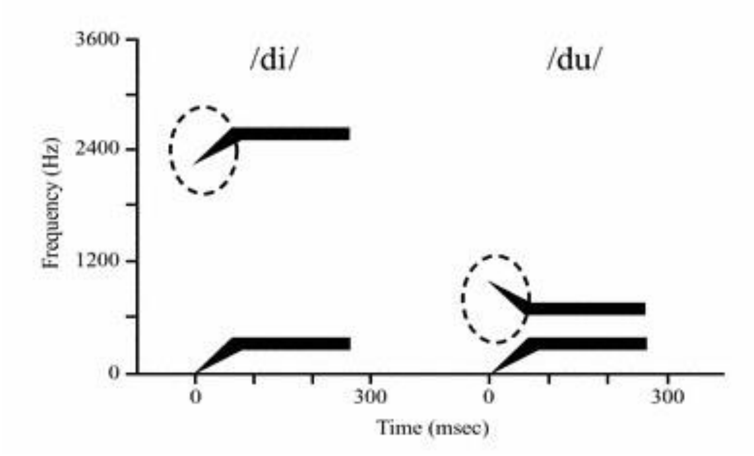

Fig. 2.2 ***The difference in formant transitions in spectrographic patterns that produce /di/ and /du/***

This version of the motor theory changed into coevolution view. Liberman et al. changed their focus from ontogenetic learning process to phylogenetic adaptations (Galantucci, Fowler, & Turvey, 2006; Liberman et al., 1967). They argued that adaptations are composed of the skills to co-articulate speech and to perceive co-articulated speech. Accordingly, the two skills are not independent and have to be coevolved, which leads to the strong link to a single mechanism. They proposed the perception procedure in which the listener analyzed the acoustic input, guessed how the speaker would articulate the sound, synthesized a virtual acoustic signal, guessed and matched the virtual signal to actual one. As a result, skilled listeners can mediate between the moments of articulation, the sensory consequence of the acoustic information through the representation of their embodied production memory. In other words, production experience is essential in order to perceive the appropriate sound or prosodic pattern in language learning.

This theory gives an idea that L2 learners can facilitate learning how to use stress pattern for segmentation by producing, not just depending on listening experience. In several recent studies, it has been reported that production can affect L2 perception (Cason, Astésano, & Schön, 2015; Kato & Tanaka, 2015; Linebaugh & Roche, 2015).

Carson et al. (2015) investigated whether audio-motor training enhances phonological processing by rhythmic priming. They had the participants perform a phoneme detection task under matching and mismatching prosodic structure with the audio-motor training learning group and the control group. The research results indicated that matching rhythmic prime facilitated phoneme detection and rhythmic audio-motor training enhanced the priming effect.

Linebaugh & Roche (2015) examined whether production can inform perception in terms of problematic L2 sounds. They conducted perceptual discrimination tests on 46 Arabic speaking learners of English and compared the results of pre-, post-, and post-post condition tests of the production group and the aural exposure group on perceiving /æ, ʌ/, /ɜ, ɔ/, /g, dʒ/. They ascertained that production can inform perception and articulation training can facilitate L2 acquisition.

Kato & Tanaka (2015) also investigated a facilitative role of sensorimotor activity in L2 speech perception of Japanese college level learners of English. The research revealed that reading aloud (RA) and listening performance have strong correlation, which suggests production accuracy/fluency played an important role in listening of both less-proficient and advanced learners.

In the previous section, it was mentioned that MSS is a hard skill for Korean learners of English because Korean speakers are not familiar with word-level stress and understand F0 rise as a phrase-final signal which can play a role as a stress signal of the onset of words in English.

The current study aims to add the experiments to investigate whether motor learning, that is producing the stress pattern which is unfamiliar with L1 can help to perceive, learn and use the stress patterns to segment the L2 speech.

Chapter 3.

METHODOLOGY

This chapter describes the methodology employed in the present study. Section 3.1 discusses the participants. Section 3.2 provides details on the instruments in terms of the text, the target words, and the tasks. The word learning assessment methods and their scoring procedures are described in Section 3.3. The procedure and the data analysis are explained in Sections 3.4 and 3.5, respectively.

3.1. Apparatus

To make the stimuli, Cenix H7 was used to record the voice of the female English native speaker and Praat (Boersma, 2002) was used to edit the sounds. Headphones (royche rhs-4100) were provided during the experiments, and SPSS (20 for Windows) was mainly employed for the statistical analysis.

3.2. Participants

A total number of 108 Korean learners of English participated in the experiment. They lived in Seoul and their age was from 10 to 13. It was considered that the participants' expatriate experience could influence on the test results due to the familiarization with the language-specific prosodic features. Thus, participants who had the experience of living abroad were removed and the test groups were divided randomly into three groups considering the gender and age equivalence. The three groups had different stress pattern cues and learning styles in the phase 2 in Experiment 1. For stress patterns, word-initial stress patterned stream was provided for Group A and Group C. Group B listened to word-final stress patterned stream. For learning styles, Group A and B learned stress patterned pseudo-words by just listening, and Group C learned by listening and mimicking the stress patterned (word-initial) artificial words. One of the main goals of this study is to investigate whether Korean learners of English can learn word-initial stress as a segmentation cue. Therefore, word-final stress stimuli were provided only as a control in the learning style of listening.

Among the participants, four members were removed due to their earlier experience of living abroad (one was in Beijing for 3 years, one was in Singapore for 2 years, and the last two were in Paris for over 5 years).

Seven participants did not understand how to take the experiment. Thus, a total of 97 learners were included in the experimental result. All the participants reported that they had no problems of hearing and speaking. The general information of the participants for groups is provided in Table 3.1.

Table 3.1 Participants Information

	Group A (n=32)	Group B (n=34)	Group C (n=31)
Gender (F/M)	17/15	16/18	15/16
Mean age	12.3	11.9	12.3

3.3. Materials

In accordance with the method used in a majority of segmentation studies using peuso-words (Endress & Mehler, 2009; Kim et al., 2012; Ordin & Nespor, 2013; Jenny R Saffran et al., 1996; Thiessen & Saffran, 2007), 18 trisyllabic(CVCVCV) pseudo-words composed of consonant b, t, n, r, f, d, m, p, v and vowel a, i, o are made. Since the participants are L2 learners of English, the consonants were selected in English and two thirds of them were common in Korean. The three sets of pseudo-words will be used at each phase in Experiment 1 are illustrated in Table 3.2.

Table 3.2 The three sets of pseudo-words by use for Experiment 1

SET 1	SET 2	SET 3
for phase 1 (TP-only cue)	for phase 2 (learning stress)	for phase 2 (TP + stress cue)
batinu, niruta, furina, firabo, tunobi, tofabu	fidumo, damipu, fudima, domupi, pavidu, povadu	bunati, tunibo, torinu, rufino, robafu, bitafo

The basic condition of the syllables is following; the duration of one syllable is 250ms and F0 is 200Hz and the detail information of the materials at each phase will be explained in the next sections.

All stimuli syllables were recorded 5 times by an Canadian woman aged 24 whose L1 is English. The syllable was produced independently with intervals and the clearest syllable was extracted among the 5 syllables. Then they were manipulated to be normalized and have the same pitch, duration.

Although higher intensity is considered related to stress (Kochanski et al., 2005; Mo, 2008), it was not considered in the present study. According to the study of Ordin & Nespor (2013), intensity can't mark prominence independently as well as differences between stressed and unstressed vowels are very small. Thus, intensity was not considered in manipulation of stress and controlled with the normalization of the mean intensity of all the syllables. A word was made by the concatenation of three syllables.

3.4. Experiment 1

The goals of Experiment 1 are to investigate whether L2 learners use TP cues for segmentation when they are exposed to continuous speech and whether they can learn stress pattern cues (i.e. word-initial or word-final stress) under two learning conditions (i.e. exposure and motor conditions).

To this end, participants are exposed to TP-only cued words without pause and conducts a word-spotting test in phase 1. After break, participants learn stress patterned words (word-initial stressed or word final stressed) one by one by exposure or repeating the given words with the same prosody. In the following, participants are exposed to the stream of stressed new words with the same TP of phase 1 and took the word-spotting test.

3.4.1. Phase 1

3.4.1.1. Materials

3.4.1.2.1. Exposure Materials

The words in SET 1: *batinu, niruta, furina, firabo, tunobi, tofabu*. TP within a word was 1.0 and between words were less than 0.25. There was no repetition of a word in a row and they were arranged quasi-randomly. All the syllables were adjusted to the same pitch: 200Hz which was F0 of the woman's voice who recorded the stimuli. Since it was around 250 ms in general to identify a specific syllable, the duration of each syllable as a baseline was controlled to 250 ms. Therefore, 1 word (3 syllables) took 750ms since each word was composed of the concatenation of the three syllables without any stress. A three syllabic word in a baseline condition is illustrated in Fig.3.1, with 200 Hz, 750 ms.

In the study of Ordin & Nespor (2013), the exposure time was 597.6 seconds (9.96 minutes) for 19-20 years old participants and in the study of Thiessen and Saffran (2007), 9-months infants were exposed to 2 minute speech chain to investigate statistical learning. Although the recognition of words was proportional to the exposure time (Endress & Mehler, 2009), in the current study, exposure time was determined between 9 minutes and 2minutes considering the age of the participants (mean age: 12) and feasible exposure time for learning effect, attention span and the total experiment time. Accordingly, one word appeared 80 times in a string of words and the total exposure time was 360 seconds (750ms X 6 words X 80 times). A sample stream of speech is */batinufurinatunobinirutatofabufirabobati.../* and 5 seconds was faded in and out in the beginning and at the end of the strings in order that the participants should guess words at the start of the string. As a result, the only cue for segmentation in this phrase was Transitional

Probability (TP).

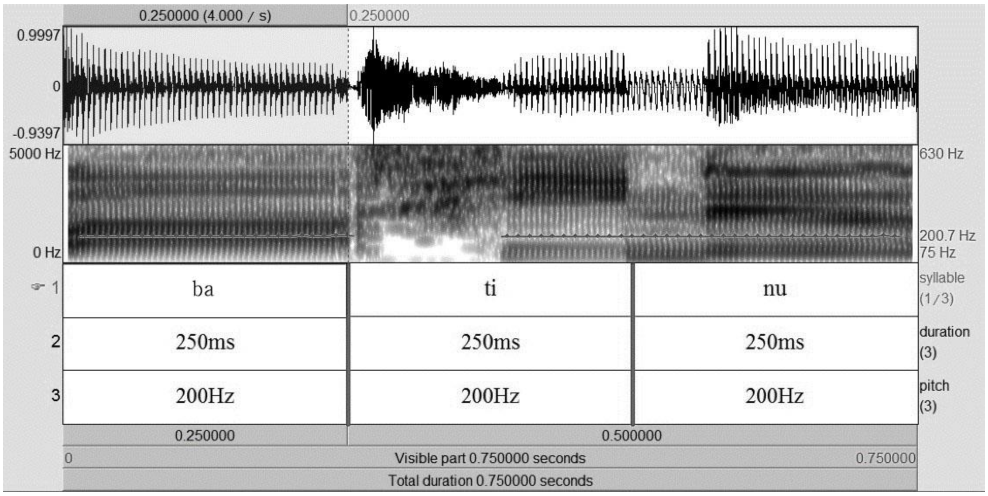

Fig. 3.1 A basic three-syllabic pseudo-word condition

3.4.1.2.2. Test Materials

The test materials were the combination of six words, three non-words and three part-words. Fig. 3.2 summarized the definition of word, part-word, and non-word from a word-stream.

Non-words mean the combinations of the appearing syllables but not concatenated. For example, BANITA consisted of the syllables appearing in the word chain, BATINU-NIRUTA, but they never had been concatenated in a row. Accordingly, TP of the non-words were 0 and the listeners did not actually hear them as words.

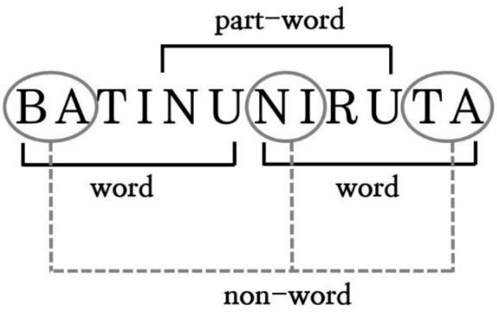

Fig. 3.2 The definition of word, part-word and non-word

Part-words mean the possible combinations among syllables of the artificial six words. For example, the word BATINU and NIRUTA could make a stream of BATINU-NIRUTA and *NUNIRU* was a C-A-B type part-word. *TINUNI* was then a B-C-A type part-word. TP within a C-A-B type part-word was 0.25 and 1; TP within a B-C-A type part-word was 1 and 0.25, which were lower than those of a word. Fig. 3.3 illustrates types of part-words and TP within a part-word.

Part-words were made considering TP and the stress pattern. Part-words were composed of neighboring syllables only whose TP was different. Thus, it is reasonable to comprising a part-word with neighboring syllables bi-directionally. Another reason was to be neutral to the stress pattern preference. When stress pattern was added to a stream, listeners might segment the chain of words depending on the stress pattern ignoring TP. There were two conditions of stress pattern: word-initial and word-final, thereby C-A-B (word-final stress as the onset) and B-C-A (word-initial stress as coda) type part-words played impartial options cross-conditionally.

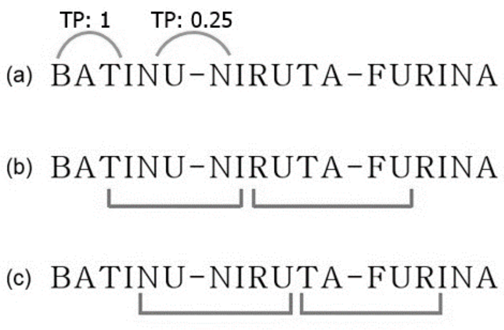

Fig. 3.3 The types of part-words for the test:
(a) a chain of words (b) B-C-A type of part-word: TINUNI, RUTAFU
(c) C-A-B type of part-word: NUNIRU, TAFURI

In this way, total 36 pairs of word and part-word, word and non-word were made for test materials. Fig. 3.4 and 3.5 gave examples two options at each test question provided to participants. There is 700 ms interval as silence between the two options. A total test materials are attached in Appendix A.

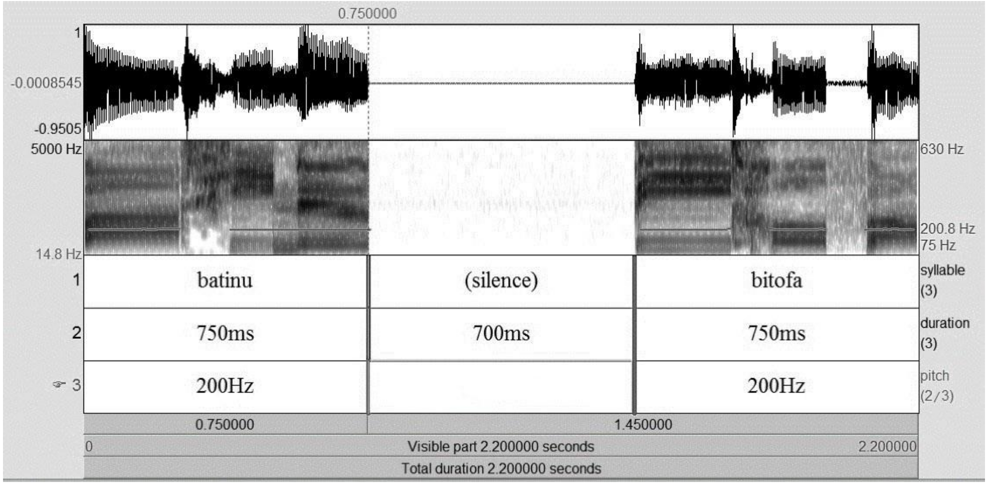

Fig. 3.4 A pair of options in the word-spotting test (word and part-word)

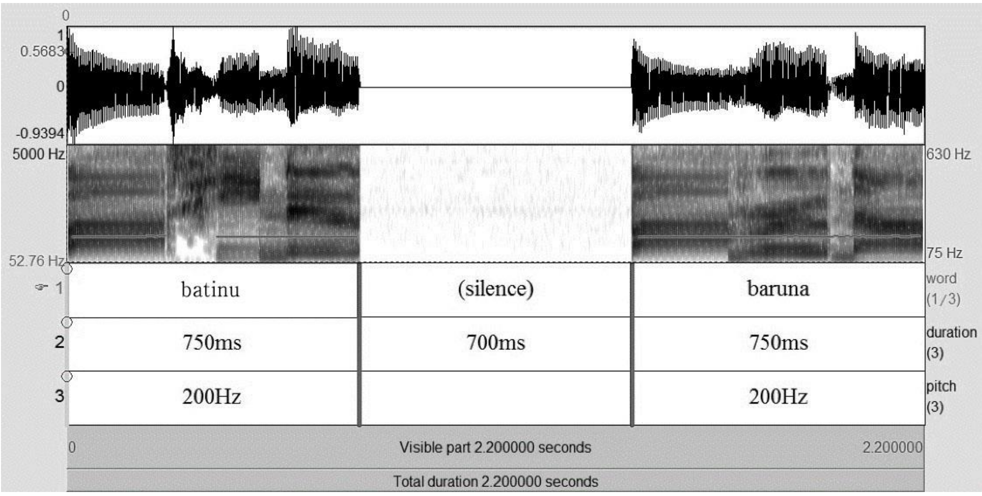

Fig. 3.5 ***A pair of options in the word-spotting test (word and non-word)***

3.4.1.2. Procedure

3.4.1.2.1. Exposure

The experiment was conducted in a quiet room with Praat software. Each participant was exposed to the 360 seconds TP-based pseudo-words stream. There was no pause. Before starting the experiment, they took practice test with an example of the stream consisting of pseudo-words which were not used in the real experiment. They wore headphones and adjusted the volume for their convenience. After listening a 6-minutes (360 seconds) stream, participants took one-minute break, and they took a word-spotting test.

3.4.1.2.2. Word Spotting Test

The participants took the 36 word-spotting tests, which were a forced-choice type. They listened to a pair of “word - 700ms silence – part (or non)-word”, and then they had to choose “1” or “2” as they thought of the “word”. If they thought all the presented examples sounded like words, they had to choose a seemingly more word-like one for them. The experimenter was always in the experiment and controlled the situation.

3.4.2. Phase 2

3.4.2.1. Materials

3.4.2.1.1. Learning Materials

The learning artificial words were the same stream as those of Phase 1. The difference was the stress pattern on the stream. The duration and pitch of a stressed syllable were manipulated. According to the study of Kim et al. (2012), duration of a baseline syllable was 252 ms and that of a stressed syllable was 446 ms (about 1.77 times). In terms of pitch, baseline was 190 Hz and stressed one was manipulated to 250 Hz based on the Korean production data (Cho & Keating, 2001).

In the current study, duration of the stressed syllable was lengthened from 250 ms to 400 ms (1.6 times) making a word length 900ms and pitch of the stressed syllable was 280 Hz comparing baseline F0, 200 Hz.

The word in SET 2 (*fidumo, damipu, fudima, domupi, pavidu, povadu*) were used as learning materials, and a sample initial stressed word is PAvidu and a final stressed one is domuPI. Each learning chain has 2 second-interval between words, thus a learning session took 294 seconds (900ms (a word) x 6 (words) x 40 times + 2000 ms (an interval) x 39, 4.9 minutes).

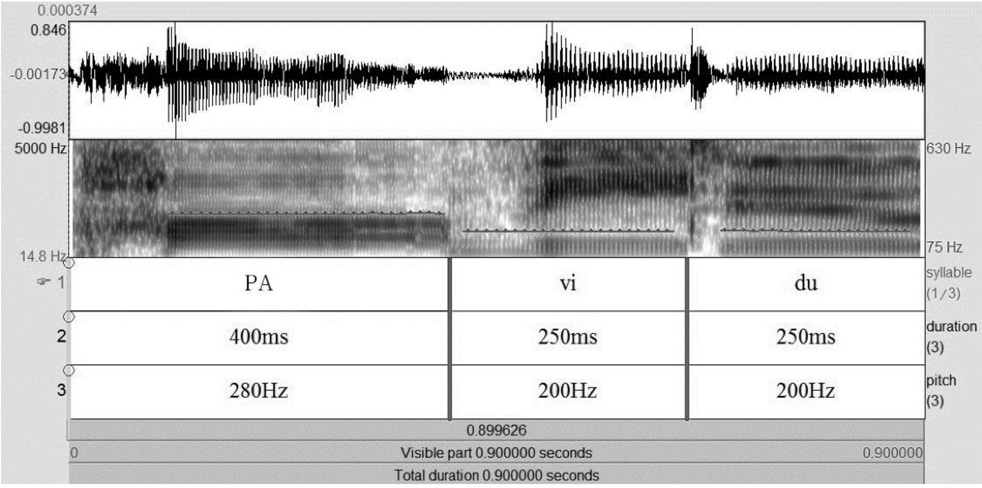

Fig. 3.6 Initial stressed word (PAvidu)

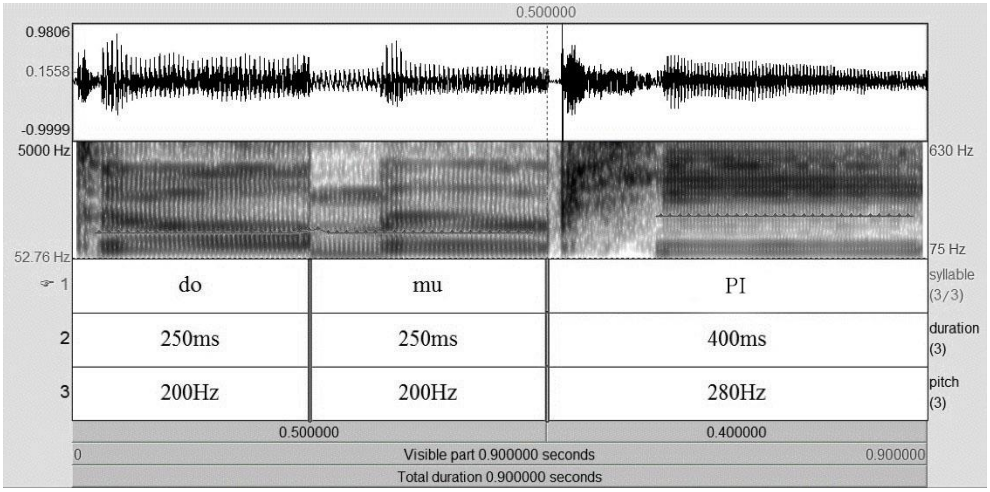

Fig. 3.7 Final stressed word (domuPI)

3.4.2.1.2. Exposure Materials

The way of making exposure materials was the same as that of learning materials in phase 1. The words in SET 3 (*bunati, tunibo, torinu, rufino, robafu*, *bitafo*) were used for a stream of pseudo-words. The condition of TP was the same as that of phase 1 - TP within a word was 1.0 and TP between words was less than 0.25. A total stressed word-chain took 432 seconds (900ms x 6 words x 80 times). A sample stream of speech is */tuniBOrufiNObunaTItoriNUbitaFOroba.../* for the word-final stress and */TUniboRUfinoBUnatiTOrinuBItafoROba…/* for the word-initial stress. 5 seconds was faded in and out in the beginning and at the end of the strings in order not to have the participants guess words at the start of the string. Accordingly, the cues for segmentation in this phrase were TP and stress patterns they learned.

3.4.2.1.3. Test Materials

The manner of making test materials was the same as that of Phase 1. For example, the word BUNATI and ROBAFU could make a stream of BUNATI-ROBAFU. *Tiroba* is a C-A-B type part-word and *Natiro* is then a B-C-A type part-word. Non-words which were presented in the word stream but not concatenated were also generated (e.g. *butiba*). In this way, 6 words, 3 part-words, and 3 non-words made a total of 36 pairs of test materials. The test materials are presented as “*natiro*”-“*tiroba*” with interval 700ms. Importantly, stress patterns were removed in the test options (Kim et al., 2012) because the purpose of the test is for participants to recognize the words not for the memory of the rhythmic patterns. As a result, the way of the test is the same as that of phase 1 and total test materials are attached in Appendix B.

3.4.2.2. Procedure

3.4.2.2.1. Learning

The participants were divided into three groups depending on the learning style and stress patterns. A and C group learned a word-initial stress patterned stream of learning materials for 294 seconds (4.9 minutes) but learning types were different. Group A were just exposed to the word chain in which 2-second pause was included between words. Thus, they learned words by listening each word with word-initial stress pattern. Groups C learned the same initial stressed words but these participants learned by vocalizing each word with the same stress pattern after listening. Groups B learned word-final stress patterned stream for 294 seconds by listening the word chain with, pauses between the words.

3.4.2.2.2. Exposure

Group A and C were exposed to a word-initial stress patterned stream of exposure materials for 432 seconds (7.2 minutes). Group B was exposed to a word-final stress patterned stream for 432 seconds. There were no pauses.

3.4.2.2.3. Word Spotting Test

After listening to the exposure stream, participants took one minute break and had the 36 word spotting test, which was forced-choice. The test procedure was exactly the same as that of phase 1.

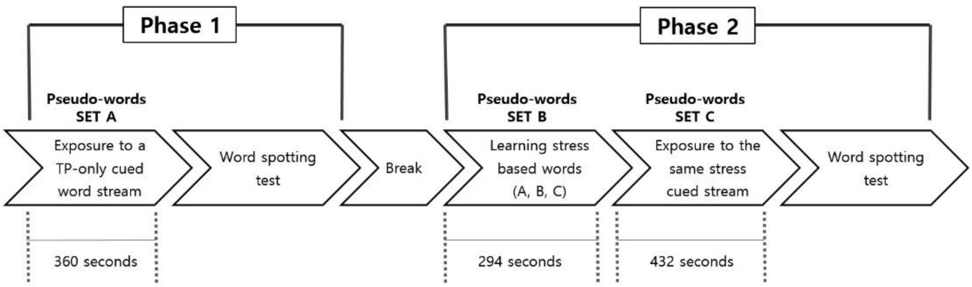

Fig. 3.8 Overall procedure of Experiment 1

3.5. Experiment 2

3.5.1. Materials

8 bisyllabic words (*diva, mipa, nubo, tira, fuda, doru, pivo, bamu*) and 8 trisyllabic words (*ropida, bapidu, tuvimo, finoda, nafito, butira, poruma, punibo*) were created for stress preference test. The first syllable of a word was manipulated as a stressed syllable, which has 400ms duration and 280Hz pitch. Other normal syllable is 250 ms and 200 Hz pitch. Each word was presented 20 times in a row.

3.5.2. Stress Preference Test

Participants listened to 16 words. One word was repeated 20 times without pause including 1.5 second fade-in and fade-out. All words were stressed initially but the words were concatenated, so the participants could not notice where the stress was located. After listening repetition of a word, participants had to take a forced-choice test either the stressed initial "word" or stressed final "part-word". For example, if participants listened to a string of /MIpaMIpaMIpaMIpaMIpaMIpaMIpa ..../, they had to choose either (1) mipa or (2) pami. If they chose /mipa/, they would prefer word-initial stress. In the same manner, if they chose /pami/, they would prefer word-final stress.

3.6. Data Collection and Analysis

In the word-spotting test at each phase of Experiment 1, the number of words were collected among what the participants recognized as words. As mentioned in the participants' section, 11 participants were removed due to their staying abroad or problems with understanding the experiment procedures.

For the statistical analysis, one sample t-test, paired t-test and Analysis of Covariance (ANCOVA) was used through the Statistical Packet for Social Science (SPSS 20 for Windows). To examine the learning effect within each group, paired t-test was conducted for each Group A, B, and C. Next, all the participants took the TP-based segmentation test in the phase 1 and it could influence on the result of phase 2 since TP has a linear correlation with segmentation. Therefore, ANCOVA was used in which the results of phase 1 were covariate in order to identify the learning effect without TP influence in phase 2.

Chapter 4.

RESULTS AND DISCUSSION

This chapter describes the results of the statistical analysis and discusses the findings. Section 4.1 reports the results of the Experiment 1. Section 4.2 discusses the results of the Experiment 2.

4.1. Experiment 1

4.1.1. Test Results

The first research question of the current study was whether Korean L2 learners sense statistical regularities in a string of speech for segmentation. To investigate the question, learners took a word spotting test after being exposed to a stream of pseudo-words only with Transitional Probability. Participants were randomly divided into three groups and their scores of the word-spotting test were collected during the TP-only exposure (phase 1) and Stress Pattern added to TP exposure (phase 2).

First of all, one sample t-test was conducted and the test value was 18 which is a half of the entire test questions, 36. The performance of TP-only test was above chance level (50% or 18 correct answers) as showed in Table 4.1.

Table 4.1 The result of one sample t-test in phase 1 (test value: 18)

Group	**N**	**t**	**p-value**
A	32	3.70	.001**
B	34	2.85	.007**
C	31	2.55	.016*

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

The overall means and the standard deviations of the test scores by the three groups in phase 1 and 2 are summarized in Table 4.2. The mean scores of the word-spotting test in phase 1 were 19.72 (Group A), 20.18 (Group B), and 19.77 (Group C) and there was no big difference, which means the participants of the three groups can use TP cue for segmentation.

Table 4.2 Descriptive Statistics of Experiment 1

Group	Stress pattern	N	Phase 1		Phase 2	
			M	SD	M	SD
A	word-initial	32	19.72	2.63	18.22	4.46
B	word-final	34	20.18	4.45	23.03	4.57
C	word-initial	31	19.77	3.88	23.13	3.93

Table 4.3 Paired t-test results of Group A

N	t	r	p-value
32	1.559	-.118	.129

Table 4.4 Paired t-test results of Group B

N	t	r	p-value
34	-3.062	.274	.004**

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

Table 4.5 Paired t-test results of Group C

N	t	r	p-value
31	-4.382	.404	.000***

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

However, the test scores of phase 2 in which TP cue and stress pattern cue were given showed difference. When word-initial stress was given (Group A), the mean test scores slightly decreased (18.22) and when word-final stress was given (Group B), the mean score pointed increased (23.03). In this case, the way of learning was just “listening”. The mean score of Group C in phase 2 was 23.13, which was also higher than that of phase 1. In this case, word-initial stress pattern was provided in the same manner of Group A but the way of learning was “repeating aloud the learning materials with the same prosodic pattern”. To verify the significance of the difference between the result of phase 1 and phase 2, paired t-test within each group was performed.

Table 4.3 to 4.5 showed the paired t-test results. The result of Group A was not significant, and there was no difference of the scores between phase 1 and phase 2, $t(32) = 1.559$, $p > .05$, the effect size $r = -.118$, Cohen’s $d = .41$ (small). The result of Group B showed significant difference between phase 1 and phase 2, $t(34) = -3.062$, $p < .05$, the effect size $r = .274$, Cohen’s $d = -.63$ (moderate). The result of Group C showed also significant difference between phase 1 and phase 2, $t(31) = -4.382$, $p < .001$, and the effect size $r = .404$, Cohen’s $d = -.86$ (large).

Fig. 4.1 illustrates the mean scores of each phase of all the three groups. In the result of Group A, although the difference was not significant, the mean score decreased in phase 2. On the other hand, the mean scores in phase 2 of Group B and Group C increased.

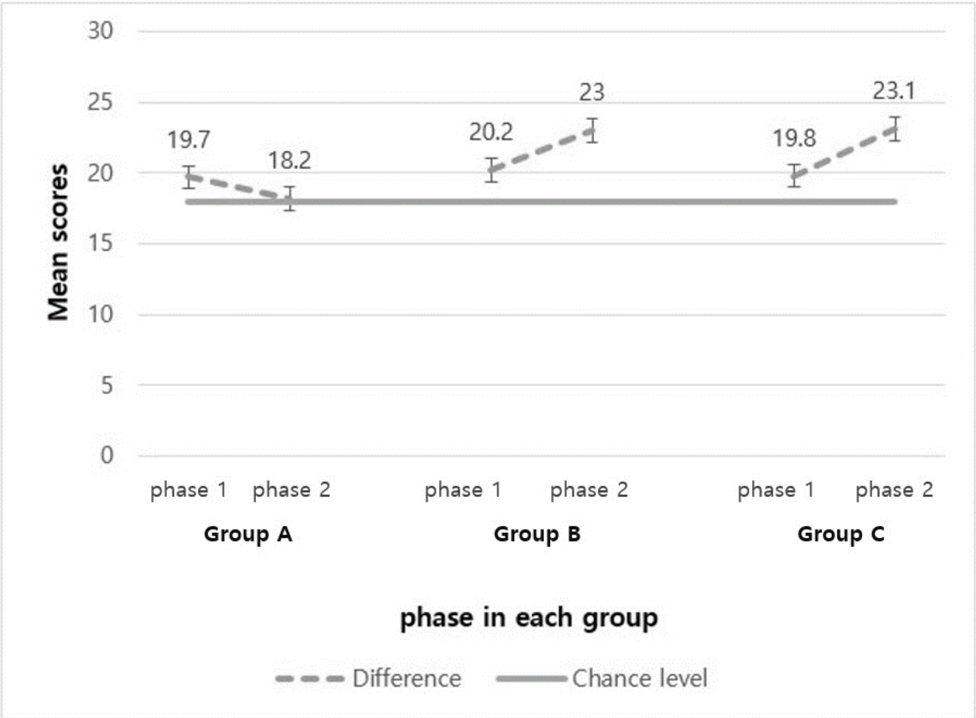

Fig. 4.1 Mean scores in the phase 1 and 2 by groups in Experiment 1

All the participants of each group took TP-only cued test in phase 1 and learned stress based segmentation strategy including TP under different learning conditions in phase 2. Accordingly, since stress pattern cues and learning types were added to TP-only cue and both cues work for segmentation, Analysis of covariance (ANCOVA) was performed with the result of TP-only cued test (phase 1) as covariate in order to analyze the difference of the learning effect among the three groups in phase 2.

Table 4.6 ANCOVA results of Experiment 1

	SS	**Df**	**MS**	**F**	**P**	**partial η^2**
TP-only	80.43	1	80.43	4.43	.038**	.045
Groups	492.64	2	246.32	13.56	.000***	.226

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

Table 4.5 revealed the ANCOVA results. TP-only cue had influence on the three groups ($F(1,95) = 4.43$, $p = .038$) though the effect size was small ($\eta_p^2 = .045$). However, the difference between the three groups was significant ($F(2,94) = 13.56$, $p = .000$) and the effect size was large ($\eta_p^2 = .23$). Post hoc testing result (Scheffe) indicated a significant difference between A and B ($p < .001$), A and C ($p < .001$), but not between B and C. Therefore, it is plausible to say that learning effects appeared in Group B and Group C and no improvement in Group A.

In each group, TP cue was a common factor and stress pattern or a way of learning were different. The result of ANCOVA indicates that the difference of the three groups without the effect of TP cue. The difference was larger in Groups than TP-only, which means stress pattern cue or a way of learning could affect more to the test scores.

The learning effect showed in Group B and Group C. The learning effect on Group B can be estimated from the correlation between the prosodic feature of the first language and that of the new language as discussed in preceding studies, and the advancement in Group C can be interpreted as the effect of the actual articulation experience by Motor Theory.

4.1.2. Discussion

4.1.2.1. Learning effect 1: TP and stress pattern

One sample t-test result revealed that TP-only cue played a role for segmentation, which means Korean learners could segment by only a statistical cue. In the exposure in phase 1, TP within words were 1.0 and TP in part-words were low, less than 0.25. Thus, participants chose words that they thought more familiar connection of syllables among non-words and part-words, which is statistically significant. The test result of phase 1 (TP-only condition) did not show difference among the three groups.

In the result of Group A, the test scores slightly decreased in phase 2, under which condition word-initial cues were given and participants learned them through listening. This means Korean learners couldn't learn how to use word-initial stress as a cue for segmentation, specifically as the onset of a word as English speakers do.

It is known that English native speakers use MSS considering word-initial stress as a signal of the onset of a new word (Cutler & Norris, 1988), and after infants (L1 learners) noticed the main prosodic pattern of English, they started to recognize stressed syllables as the onsets of words (Johnson & Jusczyk, 2001; Jusczyk, Houston, & Newsome, 1999). However, in the studies about acquiring statistical cues and stress cues for segmentation by

English L1 learners (infants), 7-month-old infants were more sensitive to TP than stress pattern, while 9-month-old infants were more sensitive to stress pattern. This implies that statistical regularity would be the first reliable cue for infants, and after enough exposure to words with prosody, they seem to learn how to use stress pattern, especially word-initial stress for segmentation applying to their first language (Thiessen & Saffran, 2007; Thiessen & Saffran, 2003). This point of view was indicated in the study of Mattys et. al (2005) in which they proposed a hierarchy of speech segmentation cues, suggesting that statistical regularity is more primary as a segmentation cue than lexical stress. Thus, using stress pattern is a language specific strategy optimized in the first language and MSS is the acquired one for native English speakers.

In the case of English L2 learners, word-initial stress seems hard to use as a cue for segmentation, which was due to the different prosodic features from their native language. In the study of Kim et. al. (2012), native Korean listeners performed worse when given initial F0 rise as a prosodic cue than when given no prosodic cues. The similar result that Korean listeners couldn't use word-initial stress as a segmentation cue came from the study of Kim & Nam (Kim & Nam, 2011) as well. Ordin & Nespor (2013) presented Italian learners of English also showed lower performance when given word initial or final stress (word-edge), and French listeners also had difficulty segmenting using word-initial stress (F0 rise) (Vroomen, Tuomainen, & de Gelder, 1998), which are presumed due to the effect of L1 prosodic structure.

Therefore, the result of the Group A can be interpreted that because of lacking word level stress pattern in the first language or conflicting stress pattern with that of L1, it is hard for L2 learners to learn how to use them as segmentation cues more primarily than statistical regularity. Indeed, since the use of word-initial stress pattern is difficult for learners whose native language doesn't have similar prosodic features, rather, it could hinder the segmentation and accordingly just brief exposure to the speech was not enough to learn.

4.1.2.2. Learning effect 2: Stress pattern and L1

The participants of Group B learned word-finally stressed new pseudo-words through listening in phase 2. In the result of Group B, it appeared the learning effect. The mean score increased from 20.2 (phase 1) to 23 (phase 2) and showed significant difference (p = .004), its effect size r (.274) and Cohen's d (-.63) was moderate. It can be interpreted that Korean learners can notice the word boundary when it's stress was in the final syllable of the word, and it is may be related to the familiarity with L1 prosody.

It was known that Koreans were sensitive to pitch (F0 rise) and duration (lengthening) at the end of a word in the stream of artificial language for segmentation (Kim et al., 2012). In another study, Kim & Cho (2009) found that Korean native speakers tended to notice word boundaries better when the prosody of the phrase level's edge was H#L, which means they consider final F0 rise as the end of a word.

Guion (2005) investigated the relationship between the age of acquisition and the language prosody acquisition of English stress patterns with early and late Korean-English bilinguals. She found that especially late bilinguals showed non-nativelike effect of lexical class and preference for stress placement on final long vowels. She interpreted the results came from the fact that Korean prosodic prominence is determined at the AP not word-level and it is realized by a tone pattern. Especially late Korean-English bilinguals were exposed to this phrase-level prosody first, which would interfere with learning word-level prosody.

Indeed, Korean has a THLH (where T= H if the AP initial segment is aspirated or tense, and T= L otherwise) tone pattern (Jun, 1993, 1998), and Koreans have segmentation strategies by the tonal information (Kim & Cho, 2009; Kim, 2004), which can influence on Foreign language segmentation.

Kim and Nam (2011) also found that high F0 was recognized as word-final for the Korean and segmentation was hindered when stress was on the initial position at words. (The result of the stress preference test in Experiment 2 in the present study supports this point of view.)

Phrasal prosody can give cues for segmentation when it's related to universal or basic stages since phonological phrase boundaries seems to act as natural boundaries (Nespor; Shukla, Nespor, & Mehler, 2007). Thus, it could be reasonable that the participants might recognize the stressed syllable as the right-edge point of a word. However, this point of view and relevant studies were conducted with Korean and some Romance languages such as French, Spanish or Italian whose stress tended to be on the right edge of a word. In this manner, L1 prosody seemingly affects L2 perception.

Several studies corroborate that the prosody of the first language affects learning foreign languages (Altmann, 2006; Lin et al., 2014; Shin & Speer, 2012; Tyler & Cutler, 2009). In particular, Tyler & Cutler (2009) investigated how the first language affected processing the foreign language in terms of prosodic cues and segmentation. They found that final lengthening might be to be used universally for segmentation but pitch movement was under the commonality in their first language. This could also be interpreted by PAM (Best, McRoberts, & Sithole, 1988).

According to the Perceptual Assimilation Model (PAM), non-native speech perception is strongly influenced by listeners' phonological knowledge of native language, and that listeners perceptually assimilate a non-native phone to one in the native phoneme systems whenever possible, based on the detection of commonalities in the articulators, constriction locations and/or constriction degrees. Korean leaners couldn't use stress in word-level as a phonological knowledge but since word-final stress is similar to L-H phrasal pattern, they might cite the tonal knowledge of the first language in order to segment the speech of a new language.

Considering this point of view, different prosodic structure can interfere with segmentation for a new language.

Ordin & Nespor (2013) clarified that lexical stress disturbs segmentation and phrasal stress facilitated segmentation. In their study, Italian learners of English were interfered with the word-edge stress for segmentation since Italian shows penultimate or antepenultimate stress pattern. However, if lexical stress was located on the edge of a word and showed F0 rise, word level stress played the same role as the phrasal stress for segmentation. Although it was marginal, it also implies this point of view with result of Group A.

Especially the prosodic characteristic of Korean is phrasal stress (AP) and L-H tonal pattern which are in line with universal prosodic pattern of final lengthening (Albin & Echols, 1996; Altmann, 2006; Bagou, Fougeron, & Frauenfelder, 2002; Cunillera, Gomila, & Rodríguez-Fornells, 2008; Echols & Newport, 1992; Endress, Scholl, & Mehler, 2005; Langus et al., 2012;

Saffran., Newport., & Aslin., 1996). This phenomenon is seen not only in right-edge stress pattern languages but in infant-directed speech (Albin & Echols, 1996).

Therefore, it is valid to interpret that Korean leaners of English could perceive word-final stress as phrase final stress, which facilitated segmentation due to the similarity to the prosodic structure of Korean and universal saliency on final stress.

4.1.2.3. Learning effect 3: learning types

The test result of Group C also revealed learning effect, mean score improved from 19.8 to 23.1 ($p < .001$), the effect size r (.404) and Cohen's d (-.86) was large. Comparing with the result of Group A whose stimuli also had word-initial stress and the large effect size, the learning effect by a way of learning is pronounced. The participants of Group C learned word-initially stressed new pseudo-words by mimicking the intonation of the words in the phase 2. Judging from the result of Group C, articulation practice could facilitate the segmentation and perception of the words.

This results implies the same opinions that production can affect L2 perception in recent L2 studies (Cason et al., 2015; Kato & Tanaka, 2015; Linebaugh & Roche, 2015). The study of Carson et al. (2015) indicated that matching rhythmic prime facilitated phoneme detection as well as rhythmic audio-motor training improved the priming effect. Audio-motor learning meaned repeating the rhythmic pattern, thus, producing aloud the rhythmic contour is helpful for L2 learners to recognize phonemes.

This repeating aloud is a useful way in case of distinguishing problematic L2 sounds. In the study of Linebaugh & Roche (2015), Arabic speaking learners of English can improve specific L2 sound perception by articulation training. Further, Kato & Tanaka (2015) revealed that reading aloud (RA) and listening performance have strong correlation, which suggests production accuracy/fluency played an important role in listening of both less-proficient and advanced learners.

These study results implies that the Motor Theory (Liberman, 1996; Liberman et al., 1967; Liberman & Mattingly, 1985) is actually helpful. The claims of the motor theory are (1) perception of speech entails the representation of the gestures related to the certain articulation, (2) speech perception and speech production share the same set of invariants which are linked innately.

With above noted research, the results of the present study can also imply the effect of the motor theory, which indicates that production consisting of practicing gestures related to articulation and intonation makes it easy to segment the speech clearly and perceive relatively more accurate words. Liberman et al. (1967) interpreted this mechanism in a sequence that listeners analyzed the acoustic input, guessed how the sound was created, synthesized a virtual acoustic signal based on the assumption, and met the virtual to the actual signal. In this sequence, the articulation memory can expedite the synthesizing virtual signals and matching them to actual ones.

Moreover, there are several neuro imaging studies supporting this theory. For example, Hickok et al. (2003) tried to identify human auditory regions with both sensory and motor response properties using fMRI. In the study, they tested the activation of a lateral site in the posterior superior temporal sulcus (STS) and a more dorsal site in the left posterior Sylvian fissure at the parietal-temporal boundary (Spt) when participants just listening the speech or musical stimuli as well as listening and rehearing what they heard for several seconds. The results revealed that Spt area was more activated when auditory + rehearsing stimuli were given than only auditory stimuli were given, which indicated it is part of a large auditory-motor integration circuit, and this activation was also under the action of working memory (Cohen et al., 1997). They suggested that the circuit enables acoustic-phonetic input to guide the acquisition of language-specific articulatory-phonetic gesture. In addition, more fMRI studies have supported there is overlap between the cortical areas activated during speech production and passive listening to speech (Pulvermüller et al., 2006; Wilson et al., 2004).

The participants of Group C could memorize the pitch and duration information of the first syllable in a word and try to segment following speech by the pattern. Moreover, they might memorize the gesture of the relevant articulation organs at the moment they produced the new words in the learning section. That experience could facilitate the learning effect. These are related to working memory since the participants had to retain the production involved information. And this information bridged between production and perception.

Taken together, it is feasible to say that production can expedite and make it more accurate to perceive and segment words in a fluent speech and it is also useful to learn methods not only for speaking itself but for efficient listening.

4.2. Experiment 2

4.2.1. Test Results and Discussion

In order to ascertain whether Korean learners show preferred stress position in word level, Experiment 2 was performed. Depending on the preferred stress position, their segmentation choice and result would be different. Fig 4.2 indicates the expected segmentation results according to their stress pattern preference in word-level.

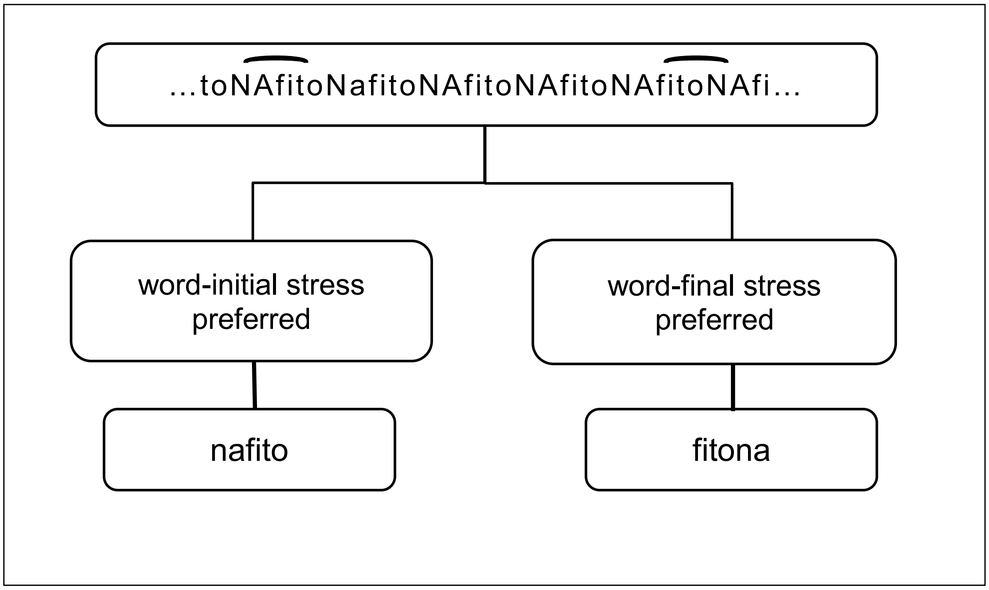

Fig. 4.2 Word segmentation by preferred stress position

In a string of one bi-syllabic or tri-syllabic word stressed on the first syllable (word-initial stress), participants had 16 forced choice tests in the option of “A-b” or “b-A” (bi-syllabic 8 words), “Abc” or “bcA” (tri-syllabic 8 words) each.

Fig 4.2 indicated how the participants would recognize the words by stress position preference and the capitalized syllables meaned the stressed syllable. In a continuous stressed word speech, the word segmentation occurs differently depending on the preference of the stress pattern by the listeners. For example, if the listeners prefer word-initial stress, they tend to recognize the stressed syllable as the onset of a word. On the other hand, if the listeners prefer word-final stress, they are apt to recognize the stressed syllable as coda of a word.

As illustrated in Fig 4.3, mean scores of word-initial stressed word spotting test were 2.6 out of 8 bi-syllabic words, 2.1 out of 8 tri-syllabic words and 4.7 out of total 16 words, which indicates 29.375 % opted for word-initial stress. In other words, approximately 70% recognized the word based on word-final stress pattern.

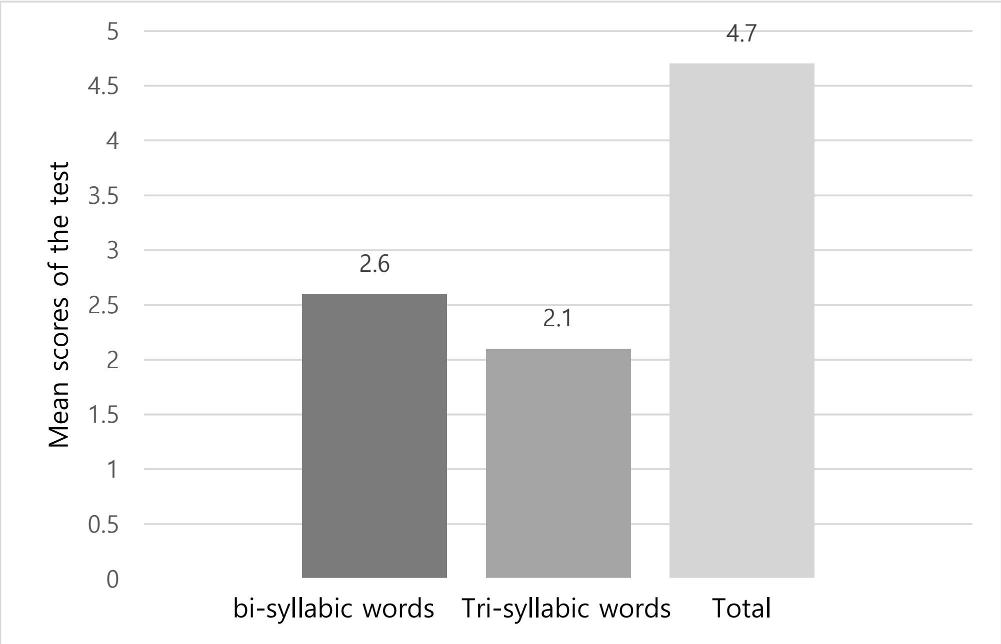

Fig. 4.3 Mean scores of word-initial stress preference test in Experiment 2

Analyzed by the number of word-initial stress proponents who scored more than 8 out of 16, it appears that approximate 10 % of Korean participants (10 out of 97) segmented the continuous speech by word-initial stress, as showed in Fig. 4.4. In other words, roughly 90 % recognized words by word-final stress. Particularly in tri-syllabic words, they heavily tended to perceive the word-final stress syllable as a signal of the end of a word. This results were aligned to the result of Group B in Experiment 1 in which participants easily learned to segment by the word final stressed syllable.

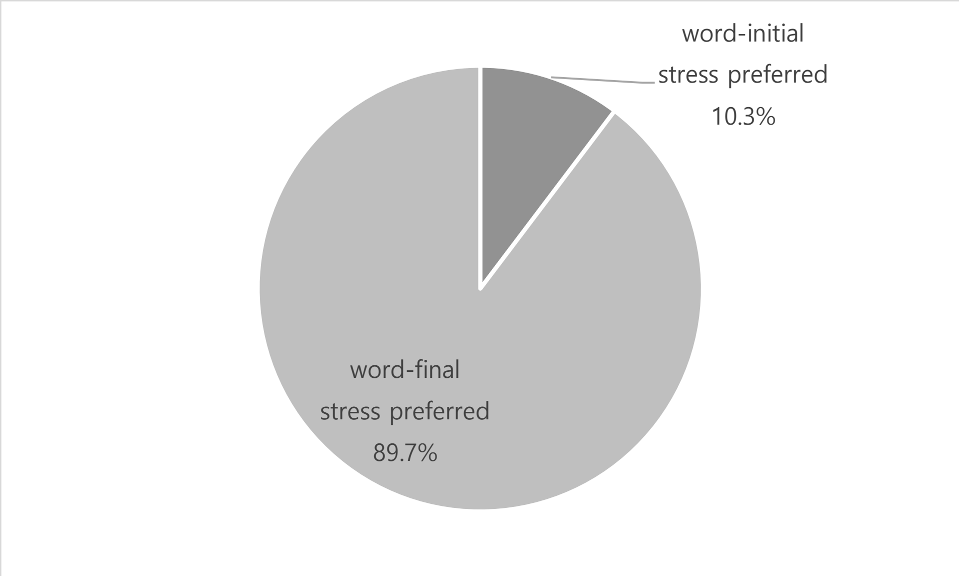

Fig. 4.4 Preference ratio by stress patterns for Korean learners

Although it is known that word-level stress doesn't exist in Korean, in case that only given cue is word-level stress pattern, they appeared to segment by word-final stress analogous to prosody of their native language, L-H phrasal level pattern and it is known universal phenomenon. This supports the learning effect of Group B in Experiment 1, which indicates the influence of L1 prosody on new language perception.

Chapter 5.

CONCLUSION

This chapter is composed of three sections. Section 5.1 summarizes the major findings of the present study. In Section 5.2, the implications are presented on English vocabulary education and on the involvement load hypothesis. Finally, Section 5.3 reports the limitations of the present study and makes suggestions for the further research.

5.1 Major Findings and Pedagogical Implication

This study was aimed at addressing the question of (1) Can Korean L2 learners sense statistical regularities in a string of speech for segmentation? (2) Do Korean L2 learners have preference of stress pattern for segmentation? (3) Can Korean L2 learners learn stress pattern cues for segmentation? (Especially word-initial stress pattern) (4) Can the experience of articulation in learning affect the enhancement of the perception ability?

According to the results of phase 1 in Experiment 1, Korean learners appeared to use statistical cues (in the current study, TP) for segmentation, showing above chance level scores. In addition, Korean learners seemingly used word-final stress as a signal of the end of a word. Experiment 2 revealed approximate 90% of Korean participants preferred word-final stress position as a signal for segmentation. Under the condition of word-final stress cues (Group B), the participants showed significant learning effect by listening new pseudo-words and applying the stress pattern to segmentation for another stream of new pseudo-words.

Although a large volume of research on Korean prosody suggested Korean does not have word-level stress but has phrase-level intonation (Jun, 1993, 1998; Kim & Cho, 2009) which interferes with Korean learners' learning stress in word-level (Guion, 2005), it appears that Korean learners can use analogy of the first language prosody (L-H pattern in phrase level) to new language’s prosody processing. This could be interpreted by PAM (Best et al., 1988) or the universal phenomenon of final lengthening as a fundamental saliency of speech.

Finally, the experience of articulation also provided learning effect with the most significant effect size. Since the motor theory was introduced, a considerable amount of research has been conducted to investigate the relationship between speech production and perception. Not only behavioral studies but also neuro imaging studies have supported the bilateral influences. The results of the current study also add evidence that production can facilitate more accurate and expeditious segmentation and perception as well as it helps L2 learners to acquire the prosodic feature which does not exist in the first language, Korean.

These findings imply that practice of accurate articulation plays a key role in clear production as well as efficient perception. In particular, production with correct articulation can easily be ignored in EFL (English as a Foreign Language) environment such as Korea. In spite of a growing number of foreigners who speak in diverse languages comparing with the past, Korean learners of English still have limited listening input since official language is Korean as well as the condition of reading-centered curriculum of English classes. In fact, it is common that students just listen to recorded conversation and check the contents in English classes of Korean classrooms.

Prosody is a key element to comprehend words by segmenting accurately in fluent speech. Therefore, speaking based on accurate prosody including stress and rhythm is necessary for Korean students who are not familiar with English prosodic structure in order to perceive speech more accurately as well as product more naturally.

5.2 Limitations and Suggestions

This study was conducted with three groups, group A and B had the same way of learning and different stress patterns. Groups C had the same stress pattern as Group A but different way of learning. It would have showed more robust results between the stress patterns and ways of learning if it was designed 2(stress pattern) x 2(ways of learning). However, the main question was whether Korean learners of English can learn word-initial stressed pattern and adapt it to segmentation. Accordingly, word-initial stress pattern was used for two ways of learning (listening and repeating aloud) in the current study.

With regard to exposure time, from 6 to 10 minutes might be too short to fully familiarize with the stimuli and to generalize the learning effect. However, considering that the age of the participants was about 12 years old, it was estimated that more than 10 minutes would be too difficult to have them listen and attend during the experiments. Therefore, expanding learning period in the future study would provide more robust evidence for the motor learning effect.

REFERENCES

Albin, D. D., & Echols, C. H. (1996). Stressed and word-final syllables in infant-directed speech. *Infant Behavior and Development, 19*(4), 401-418.

Altenberg, E. P. (2005). The perception of word boundaries in a second language. *Second Language Research, 21*(4), 325-358.

Altmann, H. (2006). *The perception and production of second language stress: A cross-linguistic experimental study.* University of Delaware.

Bagou, O., Fougeron, C., & Frauenfelder, U. H. (2002). *Contribution of prosody to the segmentation and storage of" words" in the acquisition of a new mini-language.* Paper presented at the Speech Prosody 2002, International Conference.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of experimental psychology: human perception and performance, 14*(3), 345.

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot international, 5*(9/10), 341-345.

Cason, N., Astésano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta psychologica, 155*, 43-50.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics, 29*(2), 155-190.

Clopper, C. G. (2002). Frequency of stress patterns in English: A computational analysis. *IULC Working Papers Online, 2*(2).

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task.

Cole, R. A., Jakimik, J., & Cooper, W. E. (1980). Segmenting speech into words. *The Journal of the Acoustical Society of America, 67*(4), 1323-1332.

Cunillera, T., Gomila, A., & Rodríguez-Fornells, A. (2008). Beneficial effects of word final stress in segmenting a new language: evidence from ERPs. *BMC neuroscience, 9*(1), 1.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of memory and language, 31*(2), 218-236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language,* *2*(3), 133-142.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance, 14*(1), 113.

Echols, C. H., & Newport, E. L. (1992). The role of stress and position in determining first words. *Language acquisition, 2*(3), 189-220.

Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of memory and language, 60*(3), 351-367.

Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General, 134*(3), 406.

Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: first language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition, 108*(2), 477-499. doi:10.1016/j.cognition.2008.04.002

Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. *Crosscurrents in second language acquisition and linguistic theories, 2*, 249-289.

Galantucci, Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic bulletin & review, 13*(3), 361-377.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109-135.

Gow Jr, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human perception and performance, 21*(2), 344.

Guion, S. G. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition, 27*(04), 503-533.

Hayes, & Clark, H. H. (1970). Experiments in the segmentation of an artificial speech analog. *Cognition and the development of language*, 221-234.

Hickok, G., Buchsbaum, B. R., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Cognitive Neuroscience, Journal of, 15*(5), 673-682.

Ito, K., & Strange, W. (2009). Perception of allophonic cues to English word boundaries by Japanese second language learners of English. *The Journal of the Acoustical Society of America, 125*(4), 2348-2360.

Jenny, S. (2003). Statistical language learning mechanisms and constraints. *Current directions in psychological science, 12*(4), 110-114.

Johnson, & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of memory and language, 44*(4), 548-567.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody.* The Ohio State University.

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology, 15*(02), 189-226.

Jusczyk, Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive psychology, 39*(3), 159-207.

Kato, & Tanaka, K. (2015). Reading Aloud Performance and Listening Ability in an L2: The Case of College-Level Japanese EFL Users. *Open Journal of Modern Linguistics, 5*(02), 187.

Kim, & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *The Journal of the Acoustical Society of America, 125*(5), 3373-3386.

Kim, S.-M., & Nam, K.-C. (2011). Strong (stressed) syllables in English and lexical segmentation by Koreans. *Journal of the Korean society of speech sciences, 3*(1), 3-14.

Kim, S. (2004). *The role of prosodic phrasing in Korean word segmentation*.

Unpublished Ph.D. dissertation. University of California, Los Angeles.

Kim, S., Broersma, M., & Cho, T. (2012). The use of prosodic cues in learning new words in an unfamiliar language. *Studies in Second Language Acquisition, 34*(03), 415-444.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America, 118*(2), 1038-1054.

Langus, A., Marchetto, E., Bion, R. A. H., & Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of memory and language, 66*(1), 285-306.

Liberman, A. M. (1957). Some results of research on speech perception. *The Journal of the Acoustical Society of America, 29*(1), 117-123.

Liberman, A. M. (1996). *Speech: A special code*: MIT press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review, 74*(6), 431.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

Lin, C. Y., Wang, M., Idsardi, W. J., & Xu, Y. (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition, 17*(02), 316-346.

Linebaugh, G., & Roche, T. B. (2015). Evidence that L2 production training can enhance perception. *Journal of Academic Language and Learning, 9*(1), A1-A17.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77-80.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General, 134*(4), 477.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of memory and language, 39*(1), 21-46.

Mechsner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature, 414*(6859), 69-73.

Mo, Y. (2008). *Duration and intensity as perceptual cues for naïve listeners' prominence and boundary perception.* Paper presented at the Proceedings of the 4th Speech Prosody Conference, Campinas, Brazil.

Nakatani, L. H., & Dukes, K. D. (1977). Locus of segmental cues for word juncture. *The Journal of the Acoustical Society of America, 62*(3), 714-719.

Nespor, M. _& Vogel, I. 1986. Prosodic phonology. *Foris, Dprdrecht*.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 21(5), 1209.

Oh, M. (1998). The prosodic analysis of intervocalic tense consonant lengthening in Korean. *Japanese/Korean Linguistics, 8*, 317-330.

Ordin, M., & Nespor, M. (2013). Transition Probabilities and Different Levels of Prominence in Segmentation. *Language Learning, 63*(4), 800-834.

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences, 103*(20), 7865-7870.

Repp, B. H. (1987). The sound of two hands clapping: An exploratory study. *The Journal of the Acoustical Society of America, 81*(4), 1100-1109.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27-52.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language, 35*(4), 606-621.

Saffran., Newport., & Aslin. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language, 35*(4), 606-621.

Sanders, L. D., Neville, H. J., & Woldorff, M. G. (2002). Speech Segmentation by Native and Non-Native SpeakersThe Use of Lexical, Syntactic, and Stress-Pattern Cues. *Journal of Speech, Language, and Hearing Research, 45*(3), 519-530.

Shin, J., & Speer, S. (2012). *English lexical stress and spoken word recognition in Korean learners of English.* Paper presented at the Proceedings of the 6th International Conference on Speech Prosody.

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology, 54*(1), 1-32.

Sohn, H.-M. (2001). *The Korean Language*: Cambridge University Press.

Strange, W., & Broen, P. A. (1981). The relationship between perception and production of/w/,/r/, and/l/by three-year-old children. *Journal of Experimental Child Psychology, 31*(1), 81-102.

Thiessen, & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language learning and development, 3*(1), 73-100.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental psychology, 39*(4), 706.

Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America, 126*(1), 367-376.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological science, 9*(4), 325-329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of memory and language, 40*(3), 374-408.

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of memory and language, 52*(2), 193-204.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of memory and language, 38*(2), 133-149.

Weber, A. (2000). *The role of phonotactics in the segmentation of native and non-native continuous speech*.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America, 119*(1), 597-607.

White, L., Melhorn, J. F., & Mattys, S. L. (2010). Segmentation by lexical subtraction in Hungarian speakers of second-language English. *The*

Quarterly Journal of Experimental Psychology, 63(3), 544-554.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature neuroscience, 7*(7), 701-702.

APPENDIX A

- The options of word spotting test in phase 1, Experiment 1

word : part-word (A-B-C)		word : non-word		word:part-word (C-A-B)	
1	batinu - bitofa	**13**	firabo – baruna	**25**	nobito – batinu
2	natuno - furina	**14**	batinu – turibu	**26**	rutafi – firabo
3	tafira – firabo	**15**	niruta – tatuno	**27**	tunobi – nobito
4	tofabu – bitofa	**16**	tofabu – baruna	**28**	batinu – rutafi
5	tafira – tunobi	**17**	tunobi – taribu	**29**	fabuni – tofabu
6	batinu – natuno	**18**	tatuno – batinu	**30**	nobito – rutafi
7	furina – tafira	**19**	turibu – furina	**31**	furina – tofabu
8	bitafo – tunobi	**20**	baruna – tunobi	**32**	fabuni – firabo
9	natuno – tofabu	**21**	turibu – firabo	**33**	niruta – fabuni
10	niruta – tafira	**22**	furina – tatuno	**34**	nobito – firabo
11	bitofa - niruta	**23**	niruta – baruna	**35**	tunobi – ruafi
12	natuno - firabo	**24**	tofabu - tatuno	**36**	fabuni - niruta

APPENDIX B

- The options of word spotting test in phase 2, Experiment 1

word : part-word (A-B-C)		**word : non-word**		**word:part-word (C-A-B)**	
1	bitafo - forufi	**13**	robafu – biruto	**25**	niboru – tunibo
2	tiroba – bunati	**14**	rufino – fibona	**26**	bafubi – torinu
3	robafu – nubita	**15**	bitafo – nituba	**27**	bitafo – natiru
4	tiroba – torinu	**16**	bunati – biruto	**28**	robafu – niboru
5	tunibo – nubita	**17**	tunibo – fibona	**29**	bunati – bafubi
6	rufino – toriba	**18**	torinu – nituba	**30**	tunibo – natiru
7	nubita – bitafo	**19**	tunibo – biruto	**31**	rufino – bafubi
8	forifi – robafu	**20**	bitafo – nituba	**32**	bunati – niboru
9	rufino – nubita	**21**	robafu – fibona	**33**	natiru – rufino
10	tunibo – forufi	**22**	torinu – biruto	**34**	bitafo – bafubi
11	tiroba – bitafo	**23**	rufino – nituba	**35**	niboru – torinu
12	torinu - forufi	**24**	bunati - fibona	**36**	natiru - robafu

국 문 초 록

영어 모국어 화자들은 강세 음절을 문장에서 단어의 시작 신호로 인지(Metrical Segmentation Strategy)한다고 알려져 있으나, 영어를 학습하는 한국어 화자들은 강세에 민감하지 못하며, 따라서 강세를 분절에 이용하지 못한다는 연구 결과들이 있었다.

본 연구는 한국인 영어 학습자가 통계적 규칙성, 특히 전이 확률(Transitional Probability)을 바탕으로, 연속되는 말에서 단어 간 분절(segmentation)을 위해 영어 모국어 화자와 같이 강세 음절을 단어의 시작으로 인지하여 분절하는 전략을 습득할 수 있는지 알아보고자 하였다.

실험은 서울 초등학교 및 중학교에 재학중인 만 10-13 세의 학습자를 대상으로 실시하였으며, 학습 방법에 따른 강세 습득의 효과를 알아보기 위한 실험 1 과, 한국인 학습자의 단어 강세 선호도를 알아보기 위한 실험 2 로 진행되었다. 실험 1 에서는 3 음절(CVCVCV)로 이루어진 18 개의 인공단어를 3 세트로 나누어 자극이 제작되었으며, TP 만이 주어진 자극과 TP 에 첫음(word-initial) 강세 및 끝음(word-final) 강세 단서를 가진 자극을 세 가지 방법으로 학습 한 후 단어 인식 과제 (word-spotting test)를 실시하였다. 실험 2 에서는 모든 참여자들이 첫음절에 (word-initial)

강세가 있는 8 개의 2 음절어, 8 개의 3 음절어 연쇄를 듣고, 어떤 단어로 인식하는지를 통해 강세의 선호도를 파악하는 테스트를 실시하였다.

실험 1 의 결과, TP 가 유일하게 단서로 주어졌을 경우, 한국인 학습자들도 새로운 언어를 배울 때, TP 만으로 분절을 할 수 있음을 확인할 수 있었으며, 첫음에 강세가 있는 단어를 듣기만으로 학습한 경우 오히려 분절에 어려움을 보였다. 그러나 첫음에 강세가 있는 단어를 발성하면서 학습한 경우에는 분절의 정확도가 증가하였으며, 이는 끝음에 강세가 있는 단어를 듣기만으로 학습한 경우와 비슷한 정도를 보였다. 실험 2 는 한국인들이 word-level 강세에 선호도를 보이는지에 대해 알아보는 실험이었으며, 약 89.7%의 참여자가 word-final 강세에 친숙하게 반응하여 분절하는 경향이 있음을 보여주었다.

종합하면, 이번 연구는 직접 발화하는 경험이 정확한 인지와 분절을 촉진한다는 Motor Theory 에 의해 발성하면서 학습한 집단의 학습 효과를 설명할 수 있으며, 특히 영어 모국어 화자가 분절 전략으로 사용하는 단어 첫음절 강세에서의 분절(MSS) 전략이 발성을 통해 학습될 수 있다는 가능성을 제시하였다. 또한 L1 인 한국어의 운율구조(L-H 패턴)를 새로운 언어의 운율 구조에 적용하여 단어 끝음의 강세를 분절 경계로 이용하고

있음을 알 수 있는데, 이는 PAM 모델이나 끝음 강세(word-final lengthening)의 두드러지는 특징을 통해 해석 될 수 있다.

그동안 한국인이 습득하기 어렵다고 알려진 강세의 사용 전략에 대해, 이러한 연구 결과는 학습자의 발화 경험이 말하기뿐 아니라, 듣기를 위해서도 필요한 학습과정이라는 것을 시사한다. 특히 한국어에 없는 '단어 단위의 강세'를 학습하기 위해서는 직접 조음을 하고 그 특징을 기억하는 것이 정확한 분절 및 인지에 있어서 중요한 역할을 한다고 할 수 있으며, 이는 듣기 경험(input)이 매우 제한적인 한국의 EFL 환경에서 영어 듣기 능력을 촉진시킬 수 있는 효과적인 방법이 될 것이다.

주요어: 말소리 지각, 단어 분절, 통계적 학습, 운율 분절 전략 (MSS), 모터 이론

학번: 2014-20899