Prosody in English Stop Representation by Korean Adult Learners of English

Hansook Choi
(Haskins Laboratories)


This study examines the patterns of prosodically induced variation in the English stop contrast produced by Korean learners of English. The acoustic realization of the laryngeal contrast in English stops is observed in utterance-initial, -medial, and –final positions either with or without contrastive focal accent. The speech data in English collected from 4 adult male Korean speakers are compared with the result of a previous study by 6 adult male speakers native in American English (Choi, 2005). Overall, statistical results show that Korean adult speakers do not specify variation conditioned by focus and position. Focus-induced acoustic variation in the stops, observed in the speech by native speakers of English, is not significantly replicated in the speech of the Korean learners in the present study. Additional difference in using the acoustic correlates to mark the contrast is detected in the Korean adult speakers. It suggests the adult learners have problem in specifying the postlexical prosodic structure and segmental variations, and thus, rather active instruction of the phrasal prosody is also needed in educational settings.

I. INTRODUCTION

The eventual goal in learning another language may be to attain the native-like accuracy and fluency. The very sensitive issue in achieving accuracy for the adult learners in foreign languages is how to overcome phonological accent in a second language. Phonology or phonetics has been suggested as a very age-sensitive feature (Flege et al., 1995, 1999; Scovel, 1988), and adult who learn second language are hardly able to speak the language without accent. Adult learners are found to learn the syntactic structures more readily than infants whereas they still maintain phonological accents (Flege, 1991; Ioup, 1984 for example).
Phonological accent or foreign accent includes any deviations from the native standards, which are shared by learners from same L1 background (Ioup, 2008). Out of many factors, foreign accent is discussed to be primarily influenced by the sound system of the native language. Foreign accent can thus relate to different L1 phonemic and phonetic patterns at the segmental level (Flege, 1991; Flege & Eefting, 1988), or the dissimilar prosodic systems at the suprasegmental level (Archibald, 1997; Zhang et al., 2008). These factors for the adult learner’s accent are intensively discussed at the lexical level, and it is still uncertain how the adult L2 learners specify the postlexical or phrasal phonological patterns in L2, such as the structures of focal accent, phrases, and utterances. The phrasal prosodic structures are found to condition systematic variation in segmental realizations. There is systematic strengthening or weakening in segments that reflects the hierarchical prosodic structure. In other words, prosodic structures are not simply suprasegmental features that are independent from segments, but are integrative aspects of speech marked by suprasegmentals and segments at the same time. The question here is how successfully the L2 learners are specifying the prosodic variation at the segmental representation, and whether it is better than the lexical features. This is a particularly important and valuable topic in English pronunciation instruction, given that the language instruction is goaled for successful speech production and comprehension. The speaker’s well-structured prosodic expression is essential for successful speech comprehension. A large amount of research on speech recognition has been valued the role of prosody as a key feature (Cutler et al., 1997 for a review) and in early language acquisition, prosody is suggested as a bootstrapping feature for segmentation and processing (Gleitman & Wanner, 1982; Jusczyk et al., 1992 for example). Previous studies on prosody in L2 are focusing only on suprasegmental aspects, as is mentioned above. To enhance speech comprehensibility, however, the optimal prosodic information is conveyed with the suprasegmental features and segmental variations as well.

The current study specifically explores how Korean adult learners who are exposed to English later than their puberty are depicting the English prosodic structure at the segmental representation. The acoustic values measured for different prosodic contexts are compared with the acoustic measures for those structures by native speakers of American English speakers. Any systematic deviations from the patterns by native speakers of English will be discussed as the features building up foreign accent and an educational implication on the English pronunciation instruction will be further discussed.
II. LITERATURE REVIEW

1. Prosody and Segmental Realization

Prosody was traditionally identified with suprasegmentals (Lehiste, 1970), and has been newly defined as the organizational framework that breaks down speech into units with various prominent items by Pierrehumbert and Beckman (1988). Under this notion, prosodic categories are specifications of the organization of the phonetic content, which are fundamentally different from segmental specification. Syllables, words, phrases, utterances, and (phrasal) stresses are broadly included in their notion of prosody. Prosody is easily understood to concern the features that are independent from the segments such as consonants and vowels.

However, in the field of prosodic phonology, the strong interaction between prosody and segmental realization has been reported and the hierarchical structure of prosody has been suggested as contexts for phonological changes (Beckman, 1996; Jun, 1998; Selkirk, 1986). In articulation and acoustic studies, prosodic prominence is found to condition marked gestural realizations and enhanced acoustic values. For example, at the lexical level, stressed vowels in English are found to exhibit greater opening of the vocal tract and featural enhancement, and are produced with longer and faster articulation (Beckman & Cohen, 2000; Beckman & Edwards, 1994). At the post-lexical level, prosodic domain boundaries are reported to condition segmental lengthening and gestural strengthening. Vowels are lengthened in phrase-final position (Crystal & House, 1990; Edwards & Beckman, 1988; Turk, 1999), and consonants are produced with longer and more extreme articulation in phrase-initial position, with enhanced VOT distinctions (Cho, 2001; Cho, & Keating, 2001; Fougeron & Keating, 1997; Pierrehumbert & Talkin, 1992).

Further, the different structures with prosodic prominences such as the accented and domain boundary positions are also marked differently in articulation and acoustics. In general, the accentual effects have a global effect extending the syllable with the stress or accent, whereas the domain boundary position reveals more localized effect on the initial Cs and final Vs (Crystal & House, 1990; Turk, 1999; Turk & Sawusch, 1997; Turk & White, 1999; White, 2002; Wightman et al., 1992). The prominent accented or stressed positions are marked with longer articulatory duration with somewhat bigger displacement and a general increase of peak velocity, whereas domain-final syllables do not reveal any significant changes in displacement but slower gestural movement during the acoustically longer part (de Jong, 1995; Edwards & Beckman, 1988; Edwards et al., 1991).

In short, the prosodic structures are found implemented from the suprasegmental features such as intonations but also the segmental variation with different acoustic realization. These different prosodic variations are understood to assist processing of the
continuous speech stream with the combination of strengthened and weakened acoustic cues.

2. Prosody-induced Acoustic Variation in Native Speakers of English

A systematic comparison over different prosodic prominence types at the post-lexical level shows that the prosodic prominence of constituents under contrastive focal accent or in domain-boundary position is found to condition variation in the acoustic correlates of the stop laryngeal contrasts and vowel contrasts in English (Choi, 2005). However, statistical results indicate that the effect of focal accent on acoustic measures is generally more distinctive and consistent than the effect of prosodic boundary. The distribution of acoustic measures also shows a greater distinction between the contrastive stops and vowels when accented. Variation due to domain boundary position, on the contrary, is not very great. There is no evidence that stop consonant or vowel contrasts are enhanced in domain-initial position. Domain-final vowels in English are lengthened but do not show featural enhancement or result in more distinctive vowel contrasts. Another study using the Boston Radio News Corpus confirms the findings from the previous study. Cole et al. (2007) reports that only the effect of accent on acoustic measures is found more distinctive and consistent compared to the effect of prosodic boundaries.

The present study is a replication of the previous study (Choi, 2005), but based on the speech tokens collected from the Korean learners of English. Using the same method under the identical setting, a direct comparison is tried to observe the variation between native and nonnative speakers of English. However, the current study focuses on the variation of English stops in different prosodic context, which shows a systematic effect only by the contrastive focal accent, not by the positional variation.

It will be revealed in this study whether the nonnative speakers of English are marking successfully the post-lexical prosodic structures in acoustic realization of the contrastive stop tokens, and what the differences are between the native speakers and the nonnative speakers. The differences, if any, will be further discussed for the reasons and the implication on English pronunciation instruction.

1 The subjects are not taking any language classes actively to learn English. Since they have not reached the native-like fluency and accuracy, and are struggling with their English skills, they will be identified as learners of English in the present study.
III. METHOD

1. Speech Material and Acoustic Measure

Target consonants were analyzed from the initial CV syllable of eight English words (with the target syllables underlined): pottery, botany, peter, beater, petter, bettor, pah, and bah. All target syllables bear the lexical stress in the present study. The first six words were presented in two prosodic contexts, located in the initial and medial positions of a carrier sentence. The last two words, pah and bah, were used only for the phrase-final context. The six words designed in phrase-initial position were used for the phrase-medial condition by adding modifiers in front. The edges of carrier sentences in the present study coincided with Intonational Phrase (IP) and Utterance (U) boundaries. The carrier sentences also varied in their accentual condition: with a contrastive focus on the target word (focused) or with a contrastive focus on another word in the sentence (nonfocused). (1) shows samples of the stimuli, and the complete stimuli are presented in Appendix A.

2. Subjects

Four males (HJ, IS, JH, KM), who are sequential Korean-English bilingual and native speakers in Korean, took part in the present experiment. The subjects were all born in Korea and were exposed to English when they came to the United States after their twenties either for work or for graduate studies, although they took English courses in Korea before coming to the states. All of them are in their thirties and have stayed in the states five to nine years. Their English is quite intelligible but with a strong feeling of foreign accent. The previous study with native speakers collected data from six males (AL, BD, HE, JH, MC, and RF) who were monolingual and native speakers in American English. Some results from the previous study are presented again for comparison. The English subjects were all from the Chicago area and undergraduate students in University of Illinois at Urbana-Champaign. All of the subjects had no phonetic training or knowledge and reported no speaking or listening impairment.
(1) Samples of the Speech Corpus.

i) Initial-Focused Condition
   Q: Botany is the main subject of your book?
   T: **Pottery** is the main subject of my book.

ii) Medial-Focused Condition
   Q: A yellow botany book was on the table?
   T: A yellow **pottery** book was on the table.

iii) Initial-Nonfocused Condition
   Q: Pottery is the title of your book?
   T: Pottery is the **main subject** of my book.

iv) Final-Nonfocused Condition
   Q: Bob’s nickname is Funny Pah?
   T: Pat’s nickname is Funny Pah.

3. Procedures

To control the testing materials varying in focal and positional conditions, sets of dialogues were designed with the target sentence as an answer to a question, represented as in (1). Target CV sequences which are underlined in (1) were placed in the three different positions in the prosodic domain and contained the contrastive focus when they were in the focused condition. (1) displays focused words with bold characters. The complete materials were presented to the subjects in separate blocks on the basis of the prosodic condition of the target word. The focused and nonfocused groups were presented in separate blocks in order to sustain reasonable prosody.

The blocked sets of dialogue were visually presented to the subjects with markings for the focused items to assist a clear understanding of the context. Subjects were supposed to take the role of the answerer in the minimized discourse situation. The question and answer parts were presented as a set in a similar way as in (1), and subjects were instructed to read the questions silently and answer with the given target sentences supposing the

---

2 In a dialogue with a question (Q) and the target sentence (T) as an answer, subjects took the role of an answerer. The target CV sequences are underlined, and the words with a contrastive focal accent are marked with bold characters in the target sentence. In the actual presentation of the materials to the subjects, there were no markings for the target CVs or the focused items, and the six conditions with different prosodic settings were presented in separate blocks.
discourse situation. Subjects practiced with whole set of materials until they were familiarized with all the situations. Test sentences were presented in quasi-random order after a rehearsal session. The test stimuli in a block were provided in fixed order in one round, and the order was totally mixed in the next round. The whole set was repeated again, which was the third and forth repetitions for each stimulus. Optional breaks were provided after each block upon the subjects’ request. Each block was repeated four times over the course of the experiment, varying within-block sentence order across repetitions, for a total of 450 tokens.

The recording procedure was done in a quite room. All sound stimuli were recorded through a head-mounted microphone and a Tascam DAT recorder at a sampling rate of 44100 Hz, and transferred to a PC for the analysis.

4. Measurement

Acoustic measurements of VOT and F0 at the onset of the following vowel are employed as acoustic correlates of stop voicing contrast. VOT is widely accepted as a feature to mark a laryngeal contrast of stop consonants in most languages including English (Lisker & Abramson, 1964), such that English voiceless stops show greater VOT values than their voiced counterparts. F0 at the following vowel onset is also reported as a cue to English stop voicing. Whalen et al. (1993) reports from their perception study that F0 at the vowel onset assists voiced vs. voiceless distinction of the preceding stops even with unambiguous VOT values.

The recorded sounds, transferred to a PC at a 44100 Hz sampling rate, are analyzed with the Praat program (Version 5.0.09). The duration from the stop release to the onset of the second formant in the following vowel was measured as VOT. Fundamental frequency values were manually calculated from the mean period over the initial three periodic cycles of the wave forms after the stop release, 3/D, where D is the duration of the 3 measured pitch period. The calculated values were compared with the results from the autocorrelation pitch analysis function in Praat, which showed similar values but reported some missing values that could be analyzed with the manual measurement.

5. Statistical Analyses

Repeated Measures Analyses of Variance (ANOVAs) was performed for the acoustic measures of VOTs and F0 in order for the statistical evaluation of the significant effects of diverse factors. The within-subject factors were Consonant (voiced vs. voiceless), Position (Initial, Medial, Final) and Focal Accent (accented vs. unaccented). To see more detailed effects of the considering factors, additional Univariate ANOVAs were conducted for
individual speakers. Three factor variables were generally employed for the ANOVA analysis in the present setting: Position (i.e., Initial, Medial, Final), Focal accent (i.e., Focused, Nonfocused), and Consonant (i.e., /p/, /b/). The factors which are not reported significant in the repeated measures ANOVAs can be further discussed as significant within a speaker. The patterns will be compared with the findings from the native speakers. For more detailed analysis of the observed factors, post hoc comparisons were performed at the significance level of .05. The SPSS statistical package (SPSS for windows, Standard Version, Release 11.0.1, Nov. 2001, SPSS Inc.) was used for the statistical measurement.

IV. RESULTS

1. Statistical Results

Repeated Measure (RM) ANOVAs on the VOT values report no significant effect of the within factors except for the type of consonants (F(1, 3)=2036.938, p<0.0001). Voiceless tokens are marked with greater VOT values consistently. The factor of consonant type shows variation for each subject, detected by the significant interaction of two factors (F(1,6)=78.221, p<0.0001). The subjects show different patterns in marking the two way contrast in English stops. There is no significant effect from Focal Accent.

RM ANOVAs also report a significant effect from the factor of consonant type in F0 variation (F(1,3)=264.225, p<0.0001). Just like the VOT values, F0 also shows the significant interaction between consonant types and speakers (F(1,6)=9.04, p<0.005). And a significant effect from the factor of POSITION is found (F(1,3)=24.081, p<.0001). However, the post hoc comparison shows only the significant distinction is between the final position and the other two positions, initial and medial. There are no clear distinctions between initial and medial, and thus, the initial consonants are not depicted with strengthened acoustic cues.

The results of the additional individual ANOVAs further show that focus is not a significant factor for most of the speakers. Only two speakers, IS and KM, show a significant effect of Focus in F0 variations (IS: F(1,114)=9.338, p<.005; KM: F(1,112)=10.089, p<.005). On the other hand, Position is found as a significant factors for all F0 variation across speakers (HJ: F(2, 112)=21.530, p<.0001; IS: F(2, 114)=14.752, p<.0001; JH: F(2, 112)=17.360, p<.0001 KM: F(2, 112)=60.662, p<.0001), and JH and KM’s VOT variations (JH: F(2, 112)= 3.711, p<.05; KM: F(2, 112)=4.080, p<.05). The Tukey and Scheffe Post hoc comparisons (α=0.05) reveal that the significant difference is in the final position marked with the significant low pitch. There is no variation between initial and medial, against to the expectation of the strengthening effect in the initial
position. In VOTs, JH provides significantly small durations for the final tokens without the distinction between initial and medial. In the same way, KM’s VOTs do not show any significant distinctions between initial and medial. Just like the case of native English speakers, U/IP-initial position is not marked with strengthened acoustic measures.

The lack of significant effects by focus in the present study, however, shows a clear contrast with the findings of a very significant focus-conditioned variation from the study with native speakers of English. The previous study on native speakers of English using the identical settings reported a very salient effect of focus on all speakers’ VOTs and most speakers’ F0s with a very low significance value. Tokens with the focal accent are marked consistently with higher VOTs for all native speakers of English and also higher F0 for most of the native speakers. Further comparison between native speakers of English and Korean learners of English bases on the distributional patterns follows in next sections.

2. Effects of Focal Accent on the Stop Tokens

![FIGURE 1](image)

**Distribution of VOTs in Focused and Nonfocused Contexts**

(a) Results from the Korean Learners of English

---

3 The figures on the left side are for voiced bilabial tokens and the right figures are for voiceless tokens. Boxes show means, and each error bar corresponds to mean \( \pm (1.0) \) standard error. The darker ones are for focused and the lighter ones are for nonfocused tokens.
The ANOVA analysis reveals that the contrastive focal accent is not a significant factor for all speakers’ VOT distributions in the present study with nonnative speakers of English. This result is very contradictory to the findings on focal accent-induced acoustic variation in the native speakers of English. The comparison is obvious with the following figure 1 for the VOT distributions. The clear separation detected from Figure 1(b) for the native American English speakers is not replicated in the patterns in Figure 1 (a), especially in the distribution of the voiced tokens. In voiceless tokens, a slight distinction is detected from JH and IS. Only IS shows a very marginal statistical significance though (p<.05).

**FIGURE 2**
Distribution of F0s in Focused and Nonfocused Contexts

(a) Results from the Korean Learners of English

(b) Results from the Native Speakers of English

---

4 The figures on the left side are for voiced bilabial tokens and the right figures are for voiceless tokens. Boxes show means, and each error bar corresponds to mean +/- (1.0) standard error. The darker ones are for focused and the lighter ones are for nonfocused tokens.
The variation of F0 conditioned by focal condition is depicted in figure 2. A marginally significant effect of focus in certain nonnative speaker’s F0 distributions can be observed from the higher F0, especially for the voiceless tokens, in figure 2(a). Again, the variation is not very obvious as in the patterns by native speakers. A consistent variation conditioned by focus is detected from the native English speakers’ F0 both in voiced and voiceless tokens in 2(b) with minor individual variations. Korean learners of English, on the other hand, reveal a higher F0 for voiceless tokens with the focal accent. Statistically, IS and KM mark significant effects (p<.01).

FIGURE 3
Distribution of VOTs and F0s in terms of Positions

(a) VOT distribution in the Korean Learners of English

(b) Results from the Native Speakers in American English

---

5 The left figures are for voiced bilabial tokens and the right figures are for voiceless tokens. Boxes show means, and each error bar corresponds to mean +/- (1.0) standard error. The darkest ones are for the initial, the lighter ones for medial, and the lightest ones for the final tokens.
Further, JH and KM’s VOT distributions with a statistically significant positional effect do not reflect a consistent pattern in figure 3(a). In voiceless tokens, JH has the initial tokens as the greatest, whereas KM marks the medial as the highest. The voiced tokens do not show a clear variation for different positions. Therefore, positional effects in JH and KM are due to individual uniqueness and the factor of position is not a significant within factor in the RM ANOVAs. There is no positional effect consistent for all the speakers in the present study. A consistent pattern due to the position in F0 is explained with the significant low values in the final tokens. The F0 values are always small in the final position across speakers. The inconsistent effects of Positional factors are, however, similar to the findings from the native American English speakers.

4. Contrast between English Stops

To observe the variation between the laryngeal contrast in stops under different prosodic contexts, tokens are presented in the 2-dimensional acoustic space by VOT and F0 at the onset of the following vowels. For the comparison of focal effects, figure 4 is prepared for Korean learners of English and figure 5 for native speakers of English.
Overall, the distribution of Korean learners’ English tokens is not so various according to the prosodic variations. The contrast under focus, which is depicted as a distance between the two darker symbols, does not provide a significant increase compared to the counterpart without focus, which is depicted with lighter symbols in figure 4. A slight increase is detected from IS and JH. English speakers in figure 5, on the other hand, present very different contrast marking in the focused and nonfocused contexts. In general, the contrast between voiced and voiceless stops is a lot greater in the focused tokens with enhanced VOT and F0 values. Comparatively, Korean speakers do not show a clear variation over all, especially in voiced tokens.
Stop Voicing Contrast in Focused/Nonfocused Condition by Native Speakers of English

The variation due to positional condition is depicted in figure 6. English speakers in the previous study didn’t show any acoustic enhancement in U/IP-initial position (Choi, 2005) and just in the same way, Korean learners of English in the current study also produce no strengthening effects in the tokens in U/IP-initial position. In figure 6, the darkest tokens for the initial contrast do not show the greatest distinction or the greatest distance compared to the other tokens in medial or final position except for the values in JH. The U-initial tokens show a kind of weakening patterns in HJ, KM, and IS, on the other hand.

In figure 4 and 5, symbol ‘B’ marks coordinates of the mean VOT and F0 of voiced stops over the three positions, whereas symbol ‘P’ marks the same value of voiceless stops. Bars are added to aid the visual comparison of a voicing pair, voiced and voiceless stops in each condition in figure 5. The tokens with focal accent are marked with darker symbols and the ones without focal accent are marked with lighter symbols.
6. Distribution of VOTs and F0s at the Onset of the Following Vowels

This section compares the range of acoustic values used by native speakers of English and Korean. The current study shows differences between native English speakers and Korean learners in the way how they depict the prosodic conditions in English stops. Only in the speech of the native speakers, the post-lexical prosodic structure of focal accent is marked with significantly different distributions in the acoustic correlates for the contrastive stops. As a final comparison, it is tested whether there are any differences between native and nonnative speakers in phonetic representation of the contrast segments. The phonetic influence of L1 in marking L2 segmental realization, especially in VOT ranges, has been intensively discussed, which is also greatly affected by the factor of age (Flege, 1991; Flege & Eefting, 1988). To observe any consistent deviation shared by the Korean adult learners of English, the overall distributions of the English production by native and Korean speakers are further compared with figure 7.
Figure 7 presents the mean values of six native speakers and four Korean learners of English, in the three different positions in phrases and for the two focal variations. Additional circles are added to delimit the size of variation. The solid ones are for the English speakers and the dashed ones are for the Korean learners.

One unique thing shared by Korean learners in figure 7 is the very restricted range of the voiced stops particularly in VOTs. Compared to the rather even size of variation for voiced and voiceless tokens in the English speakers, Korean learners’ voiced tokens have very small variation. In contrast to the active prevoicing in English speakers’ voiced tokens, which are marked with the negative VOTs, Korean learners rarely show prevoicing. However, voiceless tokens’ VOT range is not very dissimilar in the two different language.

The values are based on mean F0 (y-axis) and VOT (x-axis) values for individual speakers in three different positions with additional focal condition. The top left one is for final, the top right one for initial, and the bottom one is for medial.
groups. Another deviation in Korean learners from the English speakers is the overall higher F0 values and the very distinctive F0 distribution between the voiced and voiceless tokens if we ignore one speaker’s high F0s for voiced tokens in the medial position.

The observed deviations in acoustic values can be partially explained based on the stop laryngeal system in Korean. Korean laryngeal contrast does not include voiced phonemes, and voiced tokens are allophonic variation of the plain stop in the intervocalic context. In phrase initial, all tokens are voiceless phonetically, and are marked with positive VOTs. In addition, F0s at the onset of the following vowels are very actively used to mark the three way laryngeal contrast in Korean (Choi, 2005; Kim, 2000). Thus, the observed differences here can be understood due to the L1 phonology, and they feature the foreign accent in adult speakers. However, the patterns are not compared statistically and should be taken with caution. Further research is suggested.

V. DISCUSSION

The current study reveals difference between Korean learners of English and native speakers of English in presenting English laryngeal stop contrast conditioned by prosodic variation. Overall, Korean learners of English are not successfully marking the focal variation at the segmental representation in their speech production. The positional variation has been found as an insignificant factor to induce acoustic variation in the native or nonnative speakers. The Korean learners of English further show no variation induced by focal condition too, in contrast to the very obvious acoustic strengthening by focus in the speech by the native speakers of English. Korean learners of English in the present study are found with less variation to mark the prosodic structure. That is, the acoustic representation of segments is rather identical regardless the various prosodic contexts. Korean learners are producing very similar tokens in all contexts, which indicates the nonnative speakers’ greater attention on the segments. This very restricted acoustic variation in segments overrides the variation of phrasal prosodic rhythms marked by combination of relative weakness and strength. Therefore, the speech of the present Korean learners sounds rather monotonous and with foreign accent. This accent is not simply from the L1 system in that Korean also shows prosodically induced acoustic variation in stops (Choi, 2005). The current difference may due to the fact that the Korean learners pay great attention to lexical representation of segments ignoring the phrasal prosodic structures.

Further deviation of Korean learners from native English speakers involves segmental realization itself. Dissimilar use of acoustic correlates for the stop contrast is found in the speech by the native speakers of English and the Korean learners of English. The overall
VOT or F0 ranges of the Korean learners reveal a deviation from the ones of the native speakers. The deviation may induce further foreign accent in the adult Korean learners in this case, and the accent is due to the L1 phonetic and phonological system.

The implication to the English pronunciation instruction is clear from the current findings. Various efforts should be provided to reduce the effects of L1 sound system in L2 production in order for the minimized phonological accent. In addition, the attention on prosodic structure should be included in pronunciation instruction. The unsuccessful representation of phrasal prosodic structures at the segmental representation is particularly critical for the goal of speech communication. The lack of prosodic cues may impede the fast and successful speech processing and consequently, results in ineffective communication with the listeners, as is suggested by various speech recognition studies. English pronunciation instruction should thus focus on both lexical and post-lexical prosodic variation of speech as well as the accuracy of segmental representation. Too much attention on segmental representation may hinder appropriate production of suprasegmental features and prosodic variations on segments. Speech is a hierarchical structure, not a linear sequence of lexical items. The structure is marked not only by the suprasegmental features such as intonation, stress, and pause, but also strengthening and weakening of segments. The instructors should actively teach the aspects of variation to specify lexical and phrasal prosody. The learners should learn that the reduction of some segmental representation is a necessary strategy to mark the phrasal structures. The pronunciation instruction should not focus on segmental articulation or prosodic intonation separately, but should specify the interactive aspects rather overtly. Accuracy and fluency are suggested be achieved both at the segmental and the suprasegmental levels at the same time for the successful English learners.

VI. CONCLUSION

This study explores second language learners’ segmental realization of prosody in their L2 production. The dichotomy of segments and suprasegmentals is discarded, and prosodic influence on segmental realization is observed as a successful speech production. The Korean learners of English in the present study are found with limitation on segmental variation to mark prosodic structures. The deviation from the native speakers is understood as an indication of greater attention on the segmental representation. Therefore, the current findings point out the presumable priority on segmental articulation in the educational setting of English pronunciation and thus, the need of additional focus on prosody. The value of active instruction on phrasal prosody has been recognized, but within the features of intonation (McCarthy, 1991; Salter, 1999; Teschner, & Stanley Whitley, 2004 for
example), while the segmental variation as a post-lexical prosodic cue is not fully recognized nor observed. This initiative study is desired to induce future research with greater population on diverse prosodic effects in segmental representation, and more viable methods to instruct the integrative aspects in educational settings.

REFERENCES


Prosody in English Stop Representation by Korean Adult Learners of English


Appendix A

(a) Focused targets

i) initial position
A: Botany is the main subject of your book?
   B: *Pottery* is the main subject of my book.

   A: *Pottery* is the main subject of his book?
   B: Botany is the main subject of his book.

ii) medial position
A: A yellow botany book was on the table?
   B: A yellow *pottery* book was on the table.
   A: A yellow *pottery* book was on the desk?
   B: A yellow botany *book* was on the desk.

iii) final position

---

*Target syllables are underlined and the constituents under contrastive focus are marked with bold characters.*
A: Pat’s nickname is Funny Bah?
B: Pat’s nickname is Funny Pah.
A: Bob’s nickname is Funny Pah?
B: Bob’s nickname is Funny Bah.

(b) Nonfocused targets

i) Initial position
   A: Pottery is the title of your book?
   B: Pottery is the main subject of my book.
   A: Botany is the title of his book?
   B: Botany is the main subject of his book.

ii) Medial position
   A: A yellow pottery book was on the shelf?
   B: A yellow pottery book was on the table.
   A: A yellow botany book was on the table?
   B: A yellow botany book was on the desk.

iii) Final position
   A: Peter’s nickname is Funny Pah?
   B: Pat’s nickname is Funny Pah.
   A: Bonny’s nickname is Funny Bah?
   B: Bob’s nickname is Funny Bah.

(c) Targets with different vowels

i) Peter/Beater
   a) initial-focused
      A: Beater is what you call the guy?
      B: Peter is what I call the guy.
   
      A: Peter is what John calls the guy?
      B: Beater is what John calls the guy.

   b) medial-focused
      A: Lonely Beater is what you call the guy?
      B: Lonely Peter is what I call the guy.
   
      A: Lousy Peter is what John calls the guy?
      B: Lousy Beater is what John calls the guy.

   c) initial-nonfocused
      A: Peter is what John calls the guy?
      B: Peter is what I call the guy.
A: Beater is what Janet calls the guy?
B: Beater is what John calls the guy.

d) medial-nonfocused
A: Lonely Peter is what John calls the guy?
B: Lonely Peter is what I call the guy.

A: Lonely Beater is what Janet calls the guy?
B: Lonely Beater is what John calls the guy.

ii) Petter/Bettor
a) initial-focused
A: Bettor is the weirdo’s nickname?
B: Petter is the weirdo’s nickname.

A: Petter is your friend’s nickname?
B: Bettor is my friend’s nickname.

b) medial-focused
A: The Gloomy Bettor is the name of John’s new play?
B: The Gloomy Petter is the name of John’s new play.

A: The Sly Petter is the name of Billy’s new play?
B: The Sly Bettor is the name of Billy’s new play.

c) initial-nonfocused
A: Petter is your friend’s nickname?
B: Petter is the weirdo’s nickname.

A: Bettor is your brother’s nickname?
B: Bettor is my friend’s nickname.

d) medial-nonfocused
A: The Gloomy Petter is the name of your new play?
B: The Gloomy Petter is the name of John’s new play.

A: The Sly Bettor is the name of your new play?
B: The Sly Bettor is the name of Billy’s new play.

교육단계(applicable levels): all levels
주제어(key words): Prosody, VOT, F0, Acoustic variation
Hansook Choi
Haskins Laboratories
300 George St., Suite 900
New Haven, CT 06511
Email: choi@haskins.yale.edu

Received in December, 2009
Reviewed in January, 2010
Revised version received in February, 2010