

Acoustic Analysis for Natural Pronunciation Programs

Un Lim (Korea National University of Education)

<Contents>

- | | |
|----------------------------------|--|
| 1. Introduction | 3.3. Results Summary |
| 2. Methods | 3.4. Suggestion for Natural Pronunciation Programs |
| 3. Results and Discussion | 4. Conclusion and Suggestion |
| 3.1. Within and Below Word Level | |
| 3.2. Beyond Word Level | |

<Abstract>

Acoustic Analysis for Natural Pronunciation Program

Un Lim

Because the accuracy and the fluency are the essence in English speaking, both of them are very important in English teacher training and in-service English training programs.

To get the accuracy and the fluency, the causes and the phenomena of the unnatural pronunciation have to be diagnosed. Consequently, the problematic and unnatural pronunciation of Korean elementary and secondary English teachers should be analyzed with using Acoustic Analyzing tools like CSL, Multi-speech and Praat. In addition, an attempt to pinpoint what the causes of unnatural pronunciation was executed. Next a procedure and steps were proposed for in-service training programs that would cultivate the fluency and the accuracy.

In case of elementary teachers, the unnatural pronunciation of segmental features and suprasegmental features were found much. therefore segmental features should be emphasized in the begging of pronunciation training courses and then suprasegmental features have to be emphasized. In case of secondary teachers, the unnatural pronunciation of suprasegmental features were found much. Therefore segmental and suprasegmental features have to be focused at the same time. In other words, features in word level should be focused first for elementary English teacher, and features in word level and beyond word level should be trained at the same time for secondary English teachers.

* Keyword: Acoustic Analysis (음향분석), Postlexical Phonology (후어휘음운론)

1. Introduction

Although the fluency of speech may be directly influenced by the accuracy, which is more important in speaking? Is it accuracy and fluency in speaking? Answers may be difficult.

In order to speak English fluently, it is essential to speak accurately, but this does not mean that the accuracy is the essence in English speaking. It means that accurate pronunciation is to foster communicative effectiveness. The accuracy is concerned with segmental pronunciation, while the fluency is concerned with suprasegmental pronunciation, therefore thus both are very important in speaking.

Although English education in elementary schools started in 1997, elementary school text books have focused on listening and speaking. And the alphabet teaching is being taught after 5th grade. The trend of English education in Korea focuses mainly on “the fluency” and Korean English teachers are not focalizing on “teaching of pronunciation”. This doesn’t mean that pronunciation is not being taught.

In current teacher training course, there are not sufficient concentrations on “teaching pronunciation classes” and in the in-service teacher training course, the situation is the same. Therefore, it is impossible to say that Korean English teachers’ pronunciations are authentic and natural. Because of these problems, Korean English teachers face difficulties in speaking accurately and naturally, and they are hesitating to pronounce in the class.

Because students generally imitate their teacher’s pronunciation, teacher’s pronunciation is very important in the class. However, the unnatural pronunciation of Korean English teachers makes it difficult to provide students with a good models. To overcome their weakness, Korean English teachers are using video and cassette tapes. But, students don’t give the same constant of attention to video and cassette tapes, they would pay attention to their teacher’s lips.

In order to help Korean English teachers, The Ministry of Education and Human Resources Development has been inviting native English speakers since 1996 through the KORETTA (Korea English Teaching Training Assistant) and EPIK (English Program in Korea). Unfortunately these programs have not been effective (KEDI, 1998).

Therefore, this article will survey the problematic and unnatural pronunciations of Korean elementary and secondary English teachers using CSL, Multi-speech and Praat. And also try to pinpoint the causes of unnatural pronunciation. These research results could be basis for making the natural pronunciation programs.

2. Methods

To generate meaningful results, the sentences were surveyed and analyzed according to the Prosodic Hierarchy of Nespor & Vogel (1986). Sentences were divided into following categories: Syllable, Foot, Phonological Words, Clitic Group, Phonological Phrase, Intonational Phrase, and Utterance¹⁾.

The duration, energy, VOT (Voice Onset Time)²⁾ and phonological phenomena were examined in Syllable. Duration and phonological phenomena were also examined in Foot, Phonological Words, Clitic Group, Phonological Phrase, Intonational Phrase and Utterance.

Even though, more than 25 utterances of 4 Korean English teachers, presently teaching English in elementary and secondary schools were recorded, but just two utterances were chosen as data for this paper. Utterance 1 was "I guess she'd be eating chocolate cake" and utterance 2 was "She must have wanted to talk about moving in with us." These two sentences were then analyzed using CSL, Multi-speech and Praat.

From these two utterances, much information about the unnatural pronunciation of Korean English teachers could be gathered. In particular segmental features like aspiration, palatalization, flapping below words levels were centered on, and suprasegmental features within words and beyond words levels were centered on. And analyses were made of resyllabication below and within word levels.

3. Results and Discussion

3.1. Within and Below Word Level

3.1.1. Voice Onset Time

-
- 1) Phonological Word is composed of one Functional Word or Contents Word. Clitic Group is consisted of more than one Phonological Word. Phonological Phrased is consisted of more than one Clitic Group. In this Phonological Phrase, many phonological phenomena occur including flapping, palatalization, deleting, stress clash and stress shift, etc. Intonational Phrase is consisted of more than one Phonological Phrase and it is one breathing group. Utterance is consisted of more than one Intonational Phrase.
 - 2) VOT means the duration between the release of a closure and the start of the voicing. The VOT is measured in milliseconds from the spike indicating the release of the stop closure to the start of the oscillating line indicating the vibrations of the vocal folds in the vowel.

Abramson & Lisker (1964), and Klatt (1964) presented the VOT of English speakers as follows:

<Table 1> VOT of English speakers in msec

Phoneme Researchers	p ^h	p'	t ^h	t'	k ^h	k'
Abramson & Lisker	78	3	59	15	98	30
Klatt	47	12	65	23	70	30

Abramson & Lisker also presented the VOT comparison between English and Korean as follows:

<Table 2> VOT of English and Korean in msec

Phoneme Language	p ^h	p'	t ^h	t'	k ^h	k'
English	78	3	59	15	98	30
Korean	91	7	94	11	126	19

While though the Abramson and Lisker's study was implemented using words not sentences, this survey was implemented based on sentences or utterances. Because of research method, it is assumed that the duration would be changed according to the speech speed. If words and sentences would be spoken faster, the VOT would be shorter.

The results of VOT in this study are as follows:

<Table 3> VOT of Syllable and Foot in msec

	N	K1	K2	K3	K4
co	0.018	0.040	0.031	0.021	0.033
cake	0.030	0.072	0.108	0.041	0.056
talk	0.023	0.051	0.072	0.049	0.061

N: Native speaker, K: Korean English teacher

As previously mentioned, the VOT results were different. The VOT was shorter than those of Abramson & Lisker (1964). VOTs of stops appearing in the foot were longer than those in the syllable. In the utterance and discourse level, because people are usually speaking faster, the VOT becomes shorter. In other words, when the speech speed is increased, the VOT becomes shorter.

The VOT of the native English speakers was shorter than those of all other Korean English teachers. Moreover, the VOT of /c/ in 'co' and 'cake' is different, because 'co' composed syllable but foot, whereas 'cake' composed syllable and foot. The VOT of /c/ in 'co' is shorter than in 'cake'.

The Result of Standard Deviation on VOT based on <Table 3> are as follow <Table 4> below.

<Table 4> The Result of Standard Deviation on VOT

	N	Minimum	Maximum	Mean	Std. Deviation
co	5	.02	.04	.0286	.00902
cake	5	.03	.11	.0614	.03048
talk	5	.02	.07	.0512	.01823

According to <Table 4>, the mean of VOT is 0.286 msec in 'co', 0.614 msec in 'cake', and 0.0512 msec in 'talk'.

3.1.2. Flapping

Flapping occurs after a vowel and before an unstressed syllable. And it occurs after a vowel and before a syllabic consonant. And the consonant sequence /nt/ can both be realized as a nasalized flapping after deleting /t/.

Flapping seems to occur very often in spoken English, specially in North American English. Although flapping is not compulsory, without its usage, English seems unnatural. Additionally flapping can occur at various level such as Phonological Word, Clitic Group, Phonological Phrase, Intonational Phrase, and Utterance. Flapping occurred more freely at the low level than the high level.

North Americans do not usually use flapping when they address non-speaker of English, but when they are involved discourse, they appear to always apply flapping. Even though they seem to comprehend our English pronunciation not applying flapping, it would seem more natural and appropriate to make use of flapping.

In the utterance, "I guess she'd be eating chocolate cake," there is a single flapping opportunity, and in the utterance, "she must have wanted to talk about moving in with us," there is one nasalized flapping possibility.

As a result, three Korean English teachers used flapping and used the nasalized flapping. One elementary teacher didn't use the flapping and the nasalized flapping. However more interesting finding was that he/she aspirated /t/ like [t^h] aspirated. In

other words, the teacher resyllabicated /eating/ like [ea]₆ [ting]₆.

3.2. Beyond Word Level

3.2.1. Resyllabication

The prosodic hierarchy of “I guess she’d be eating chocolate cake” is shown in table 5 below.

<Table 5> Within and Beyond Word Level

U	I guess she'd be eating chocolate cake.						
I.P.	I guess			she'd be eating chocolate cake			
P.P.	I guess			she'd be eating chocolate cake			
C.G.	I guess		she'd be eating			chocolate cake	
Word	I	guess	she'd	be	eating	chocolate	cake

U : Utterance, I.P.: Intonational Phrase, P.P. : Phonological Phrase, C.G. : Clitic Group,
Word : Phonological Words

At the Clitic Group level ‘she’d be eating,’ two elementary school teachers and one secondary teacher inserted [ɨ] after /d/. In addition, at the Clitic Group and Phonological Word level ‘chocolate’, two teachers resyllabicated ‘chocolate’ like [chok]₆ [late]₆.

The prosodic hierarchy of “She must have wanted to talk about moving in with us” is presented below.

<Table 6> Within and Beyond Word Level

U	she must have wanted to talk about moving in with us										
I.P.	she must have wanted to talk about moving in with us										
P.P.	she must have wanted to talk						about moving in with us				
C.G.	she must have wanted				to talk		about moving in			with us	
Word	she	must	have	wanted	to	talk	about	moving	in	with	us

U : Utterance, I.P.: Intonational Phrase, P.P. : Phonological Phrase, C.G. : Clitic Group,
Word : Phonological Words

At the Clitic Group level 'she must have wanted,' one elementary school teacher inserted [ɪ] after /t/. This resyllabication maybe directly related to the Korean Syllabic organization, (C)V(C). This syllabic structure maybe causing some Koreans to insert [ɪ] after C.

3.2.2. Compound Stress Rule (CSR)

English is a stress-timed language. In other words, stress plays a important role in conveying meaning when conversing. The CSR and Nuclear Stress Rule (NSR) are important rules in English stress.

CSR :

$V \rightarrow [\text{stressed}] / [W \text{ ______ } W]_{\text{word}}$

Condition: Second word is not branching

NSR :

$V \rightarrow [\text{stressed}] / [W \text{ ______ } W]_{\text{phrase}}$

When non-native speakers of English don't apply English stress rules correctly, their speech appears artificial. In some severe situation, they may not be able to convey their intents.

The energy results of this study are as follows:

<Table 7> Energy in Syllable and Foot in dB

Syllable Person	Cho	co	late	cake
N	75.60	71.18	72.60	73.85
K1	53.97	39.10	56.66	48.83
K2	65.47	66.44	68.11	66.32
K3	58.21	58.29	70.23	64.05
K4	62.28	49.39	58.46	54.29

N: Native speaker, K: Korean English teacher

The results indicate that the native speaker appeared to place stronger stress on 'cho' than 'cake' by CSR. The native speaker, one elementary and one secondary school teacher also placed stronger stress on 'cho.' The other Korean English teachers

stressed 'cake.'

The Result of Standard Deviation on Energy based on <Table 7> are as follow <Table 8> below.

<Table 8> The Result of Standard Deviation on Energy

	N	Minimum	Maximum	Mean	Std. Deviation
Cho	5	53.97	75.60	63.1060	8.21292
co	5	39.10	71.18	56.8800	12.94315
late	5	56.66	72.60	65.2120	7.19179
cake	5	48.83	73.85	61.4680	9.93435

According to <Table 8>, the mean of Energy is 63.11 dB in 'Cho', 56.88 dB in 'co', 65.21 dB in 'late', and 61.47 dB in 'cake'.

3.2.3. Palatalization

Palatalization like flapping occurs very often in North American English. And like flapping, palatalization can occur at the levels of Phonological Word, Clitic Group, Phonological Phrase, Intonational Phrase, and Utterance. Palatalization occurred more naturally and freely at the low level than the high level.

a) Palatalization by neighboring palatal

- i) a bad joke [j] → [j]
- ii) whose shape [ʒʃ] → [ʃ]
- iii) worse shape [ʃʃ] → [ʃ]
- iv) is she [ʃʃ] → [ʃ]

b) Palatalization by high vowel

- i) coul[j] you guys
- ii) can'[č] you
- iii) unle[š] you do
- iv) ble[š] you

Even though, Nespov & Vogel (1986: 209) provided research examples of palatalization occurring at the Clitic group, one example of palatalization was found in Intonational Phrase in this study.

Native speaker, one elementary school teacher and one secondary school teacher

palatalized /s/ in 'guess.' The other teachers pronounced /s/ and /ʃ/ separately. Although two Korean English teachers palatalized /s/, duration of [ɛ] was longer, the pitch was higher than that of the native speaker's, whereas energy was lower than that of the native speaker.

3.2.4. Deletion and Reduction

At the Clitic Group level 'she must have wanted,' reduction and deletion occurred. And 'Have' was pronounced as a weak form and /h/ was deleted before a lax vowel. After using reduction and deletion, pronunciation of 'she must have wanted,' was changed like [ʃi mʌst hæv wɔnid]→[ʃi mʌst əv wɔnid].

At the Phonological Phrase 'she must have wanted to talk,' reduction, devoicing and deletion were applied like [ʃi mʌst hæv wɔnid tə tɔ:k] → [ʃi mʌst əv wɔnit tə tɔ:k]→ [ʃi mʌst əv wɔni tə tɔ:k].

Owing to reduction, devoicing and deletion at the Phonological Phrase, the duration of the two Phonological Phrases appeared nearly the same.

The duration of the two Phonological Phrases in this study are shown in <Table 9> below.

<Table 9> Duration in Phonological Phrase in sec

Duration	Native		K1		K2		K3		K4	
	P.P. 1	P.P. 2	P.P. 1	P.P. 2	P.P. 1	P.P. 2	P.P. 1	P.P. 2	P.P. 1	P.P. 2
	1.247	1.221	2.149	1.290	1.586	1.347	1.512	0.994	1.338	1.343

K: Korean English teacher

At the first Phonological Phrase, three elementary English teachers used nasalized flapping, and two of them used deletion. They didn't use contraction and reduction at all.

Two elementary English teachers pronounced 'she must have wanted' like [ʃi mʌst hæv wɔnid tu tɔ:k] or [ʃi mʌst hæv wɔnit tu tɔ:k]. It is also interesting to note that one teacher didn't use reduction, deletion and contraction, he/she pronounced 'she must have wanted' like [ʃi mʌst hæv wɔntid tu tɔ:k].

Because of the absence of reduction and/or contraction, the duration of the two phonological phrases may not equal.

The Result of Standard Deviation on Duration based on <Table 9> are as follow

<Table 10> below.

<Table 10> The Result of Standard Deviation on Duration

	N	Minimum	Maximum	Mean	Std. Deviation
PP1	5	1.25	2.15	1.5664	.35247
PP2	5	.99	1.35	1.2390	.14614

According to <Table 10>, the mean of duration is 1.57 msec of PP 1 and 1.24 msec of PP 2.

3.2.5. Mono Syllables rule

Nespor & Vogel (1986: 179) suggested the following Mono Syllable Rule in the Phonological Phrase as the following examples.

- 1) [The sluggers]_Φ [boxed]_Φ [in the crowd]_Φ
- 2) [the cops]_Φ [boxed in]_Φ [the crowd]_Φ

In 1), the preposition 'in' is the leftmost node of a Φ and is thus weak with respect to the strong node dominating 'crowd'. Hence, 'in' may undergo reduction. In 2), on the other hand, 'in' may not be reduced, since it is the rightmost node of the Φ containing 'boxed in,' and is thus labeled strong with respect to its sister, which is weak.

Like Nespor & Vogel's the preposition example, *in* of 'moving in' will be labeled strong with respect to Mono Syllables Rule.

The Energy results in "moving in" are as follows.

<Table 11> Energy in Phonological Phrase in dB

Syllable Person	move	ing	in	with	us
N	82.96	85.65	87.32	75.56	68.09
K1	71.30	71.70	75.15	67.16	61.74
K2	52.81	63.30	59.60	48.04	54.10
K3	64.28	66.48	65.17	57.30	61.86
K4	63.22	67.23	60.38	59.72	58.83

N: Native Speaker, K: Korean English teacher

<Table 11> seems to suggest that the native speaker put the strongest stress on 'in' but the Korean English teachers all appear to have placed the strongest stress on other words.

The Result of Standard Deviation on Energy based on <Table 11> are as follow <Table 12> below.

<Table 12> The Result of Standard Deviation on Energy

	N	Minimum	Maximum	Mean	Std. Deviation
move	5	52.81	82.96	66.9140	11.13797
ing	5	63.30	85.65	70.8720	8.78868
in	5	59.60	87.32	69.5240	11.71970
with	5	48.04	75.56	61.5560	10.38850
us	5	54.10	68.09	60.9240	5.09523

According to <Table 12>, the mean of Energy is 66.91 dB in 'move', 70.87 dB in 'ing', 69.52 dB in 'in', 61.56 dB in 'with' and 60.92 dB in 'us'.

3.3. Results Summary

Through this research, a variety of problems of Korean English teachers' pronunciation were be identified.

First, the VOT of Korean English teachers appeared longer than those of native speaker's.

Second, two elementary school teachers and one secondary school teacher appeared to flap /t/ as [r]. And one elementary teacher was viewed resyllabifying 'eating' as [ea]₆ [ting]₆, as well as aspirating /t/. Along with these finding, three Korean English teachers were found to use nasalized flapping.

Third, two elementary school teachers and one secondary school teacher inserted [i] after /d/. And at the Clitic Group and Phonological Word 'chocolate' was resyllabicated like [chok]₆ [late]₆ by two Korean English teachers.

Fourth, two elementary school teachers placed stronger stress on 'cake'. In word words, they might not know about Compound Stress Rule.

Fifth, all Korean English teachers didn't appear to have knowledge about the Mono Syllable Rule, and therefore they placed the strongest stress incorrectly.

Sixth, teachers appeared to understand the concept of devoicing and deletion, but they didn't seem to understand reduction and contraction. Because of these misunderstanding, they appeared not to make an isochronism.

Finally, one elementary, and one secondary school teacher palatalized in the /s/ of 'guess.'

From this analysis, it is concluded that elementary school teachers may have unnatural pronunciation below, within and beyond word level, and secondary teachers may have unnatural pronunciation within and beyond word level.

3.4. Suggestion for Natural Pronunciation Programs

With analysis, it claims that elementary school teachers may have difficulties in pronouncing below, within and beyond word level. And secondary school teachers even though they appear to may have unnatural pronunciation below word level, appear to have more difficulties in pronouncing within and beyond word level.

It may be that pronunciation training courses should focus mainly on pronunciation below word level for elementary school teachers at first, and then gradually move to within and beyond word level. Courses for secondary school teachers, although should be made below the word level, may need to concentrate on pronunciation within and beyond word level. This mean that several courses need to be integrated.

Pronunciation training courses have to first emphasize segment features, and then move to suprasegmental features for elementary school teachers. As for secondary school teachers, pronunciation training courses need to focus on pronunciation at the word level and then move to suprasegmental features, in other words beyond word level.

4. Conclusion and Suggestion

These results obtained in this study drawn from two utterances by four Korean English teachers, however post-study experiments on 24 other utterances will be needed.

Therefore current plans are underway to survey more than 200 Korean English teachers who will attend in-service training programs at various institutions. Pronunciation will be recorded before and after completion and a result greater headway in pronunciation development strategies will be generated.

References

- Abramson & Lisker (1964), A cross-language study of voicing in initial stops : acoustical measurements, *Word* 20.
- Altenberg, B. (1987), *Prosodic Patterns in Spoken English : Studies in the Correlation between Prosody and Grammar for Text-to-Speech Conversion*, Lund University Press.
- Avery, P. & S. Ehrlich (1997), *Teaching American English Pronunciation*, Oxford: Oxford Univ. Press.
- Chastain, K. (1971), *The Development of Modern Language Skills : Theory to Practice*, Philadelphia: Center for Curriculum Development.
- Giegerich, H. J. (1980), On Stress-timing in English Phonology, *Lingua* 51.
- Jeon, Sangbum (1993), *English Phonetics*, Seoul: Shina-sa.
- Jones, D. (1976), *An Outline of English Phonetics*, Cambridge University Press.
- KEDI (Korea Education Development Institute) (1998), The effectiveness of inviting program of Native Speakers, Research Paper of KEDI.
- Kim, Keyseop (1998), *Application of English Phonology*, Seoul: Hankook Press.
- Kim, Youngseok (1991), *English Phonology*, Seoul: Hanshin Press.
- Ladefoged, P. (1993), *A Course in Phonetics*, 4th ed., New York: Harcourt, Brace, Jovanovich, Inc..
- Nespor, M. & I. Vogel (1986), *Prosodic Phonology*, Foris Publication.
- Roach, P. (1985), *Metrical Phonology and Phonological Structure : German and English*, Cambridge University Press.
- _____. (1987), *English Phonetics and Phonology : A practical course*, Cambridge: Cambridge University Press.
- _____. (1993), *English Phonology*, Cambridge University Press.

접수일자: 2002년 4월 30일

수정일자: 2002년 10월 18일

게재결정: 2002년 12월 12일

▶ 임 운(Un Lim)

주소: 충북 청원군 강내면 월탄리 90 태암수정 102-106

소속: 한국교원대학교 영어교육과

전화: 043-232-4844

e-mail: knue087@hanmail.net