A Corpus-Based Analysis of Korean Segments Produced by Chinese Learners

Seung Hee Yang[1], Hyuksu Ryu[2], and Minhwa Chung[1,2]
[1] Interdisciplinary Program in Cognitive Science, Seoul National University, Seoul, REPUBLIC OF KOREA
E-mail: sy2358@snu.ac.kr
[2] Department of Linguistics, Seoul National University, Seoul, REPUBLIC OF KOREA
E-mail: {oster01, mchung}@snu.ac.kr

***Abstract*—The increasing demand for learning Korean as a foreign language yields a strong need for a CAPT system that is able to provide automatic tutoring. However, there is limited research on Korean pronunciation produced by non-natives. As a preliminary research towards developing a CAPT system for Chinese learners of Korean, we survey key findings of previous studies. And then, based on corpus analysis, we provide improved descriptions of segmental variation patterns of Korean produced by Chinese learners. The most salient variation is substitutions of liquid sounds: 33.0% of flap sounds were realized as lateral, and 35.0% of lateral sounds were realized as 3 major variation patterns. By quantifying all the variation patterns with statistical data, we resolve disagreements between previous studies, indicate new findings, which are important resources for developing a CAPT system, and lay the groundwork for Korean language learning for various L1 backgrounds.**

I. Introduction

Pronunciation variations in non-native speech are far more diverse than those observed in native speech. This poses a difficulty for Computer-Assisted Pronunciation Training (CAPT) systems to automatically recognize learners' speech and provide corrective feedbacks. To figure out the difficulties, many previous studies have analyzed segmental and context-dependent variation patterns in non-native speech [1][2].

Despite the growing interest in learning Korean as a foreign language (L2), no automatic tutoring system using ASR technology has been seriously considered or developed so far. To be able to benefit from the improvements in ASR technology, an objective and scientific analysis of Korean speech as L2 is necessary. However, research in the field still remains limited.

Some common variation patterns observed in Korean pronunciation learning include (1) final consonants deletion, (2) lenis, fortis, and aspirated consonants substitutions, and (3) additional phoneme transfers from the learner's mother language (L1). These phonetic and phonological variation patterns can be categorized and described systematically, and it is one of our research motivations to compare how the patterns vary among different L1 backgrounds.

In this paper, as a first step for developing a CAPT system for Chinese learners of Korean, we provide a quantitative analysis of a spoken Korean corpus produced by Chinese learners of Korean and show our new findings.

Contrastive analysis shows the pronunciation units which uniquely exist in L1 and L2, respectively, and allows us to predict the pronunciation variations in non-native speech. Previous studies have used this method to predict Chinese learners' Korean pronunciation variations, cross-checked the predictions with learners' data, proposed teaching methods based on the outcomes [3–8], and assessed its effectiveness in a classroom environment [9]. Contrastive analysis was also conducted at the supra-segmental level [10][11], and many of these studies carried out acoustic phonetic experiments to complement the analysis results.

In Section II, we describe a contrastive analysis of Korean and Chinese pronunciations. In Section III, based on the predictions of variation patterns found in previous studies, we organize their experimental results according to a consistent standard. To obtain a comparable perspective, we distinguish the areas where the previous studies agree and disagree. In Section IV, we then propose an experimental methodology and show a corpus-based statistical analysis of pronunciation variations.

II. Contrastive Analysis of Korean and Chinese Pronunciations

Speech Learning Model (SLM) [12] claims that L1 and L2 both influence the pronunciation of the learners. Given that L1 influences foreign language learning, comparing its similarities and differences with L2 will help predict the learners' pronunciation variation patterns. In this Section, we compare the phonemic inventories and syllable structures of Korean with those of Chinese, which provides grounds for predicting pronunciation variations in Korean segments produced by Chinese learners.

A. Phonemic systems for consonants

Tables I and II show Korean and Chinese consonants. The Chinese language discussed in this paper refers to Standard Chinese. There are 19 phonemes in Korean consonants including the approximants /w, j, ɰ/ [10], and 22 in Chinese consonants excluding the approximants /w, j, ɰ/ [11].

The stops and affricates in Korean can be grouped into lenis, fortis, and aspirated sounds, while in Chinese, they are grouped into voiced and voiceless distinctions. The lenis stops /b̥, d̥, g̊/ and lenis affricate /dʑ̥/ in Korean are slightly aspirated, while the aspirated stops /p^h, t^h, k^h/ and affricate /tɕ^h/ are heavily aspirated. The fortis stops /p$^=$, t$^=$, k$^=$/ and affricate /tɕ$^=$/ are laryngealized and not aspirated. Chinese affricates /ʈʂ/ and /ʈʂ^h/ do not exist in the Korean counterpart.

Fricatives are grouped into lenis and fortis in Korean, while they are grouped into voiced and voiceless distinctions in Chinese. The post-alveolar fricative /ʂ/ and labio-dental fricative /f/ do not exist in Korean. For approximants, Korean has /w/, /j/, and /ɰ/ which do not count as individual phonemes in Chinese.

TABLE I
KOREAN CONSONANTS. ADOPTED FROM [10]

	Bilabial	Alveolar	Palatal	Velar	Glottal
Stop	b̥ p^h $p^{=}$	d̥ t^h $^{=}$		g̊ k^h $k^{=}$	
Affricate			ʥ̥ $ʨ^h$ $ʨ^{=}$		
Fricative		s $s^{=}$			h
Nasal	m	n		ŋ	
Liquid		l			
Glide	w		j	w ɰ	

TABLE II
CHINESE CONSONANTS. ADOPTED FROM [11]

	Bilabial	Labio-dental	Dental	Post-alveolar	Velar
Stop	p p^h		t t^h		k k^h
Affricate			ts ts^h	ʈʂ $ʈʂ^h$	
Fricative		f	s	ʂ	x
Nasal	m		n		ŋ
Liquid			l		
Approximant				ɹ	

B. Phonemic systems for vowels

There are 8 phonemes in Korean vowels and 10 in Chinese vowels. Tables III and IV show Korean and Chinese vowel inventories. The two inventories share /i/, /u/, and /a/ in common. While /ɯ/, /e/, /o/, and /ɛ/ sounds of Korean do not exist in Chinese, /ə/ sound of Chinese does not exist in Korean.

TABLE III
KOREAN VOWELS. ADOPTED FROM [10]

	Front	Central	Back
Close	i		ɯ u
Close-mid	e		o
Open-mid	ɛ		
Open		a	ʌ

TABLE IV
CHINESE VOWELS. ADOPTED FROM [11]

	Front	Central	Back
Close	i y		u
Mid		ə	
Open	a		

C. Syllable structures

A syllable in Korean is composed of (C)(j/w/ɰ)V(C), a consonant in the onset, a monophthong or diphthong in the nucleus, and a consonant in the coda. The onset and coda consonants are optional. A syllable in Chinese is composed of initials and finals, and the former is composed of optional consonant and the latter is composed of a monophthong or a diphthong, followed by optional /n/, /ɻ/ or /ŋ/. Chinese is an open syllable language in general, whereas Korean adopts both open and closed syllables.

The differences in syllable structures show that /n/ and /ŋ/ are the only consonants that can be realized as the final segment in Chinese, whereas a Korean syllable allows /g̊, n, d̥, l, m, b̥, ŋ/ as the final segment.

III. SURVEY OF PREVIOUS STUDIES

Previous studies have conducted experiments to check their predictions on pronunciation variation patterns occurring in Korean produced by Chinese learners. These experiments were conducted by analyzing read speech that is composed of words [5][7][9] or sentences [8] from Korean as a foreign language textbooks for Chinese learners of beginner [7] [9], intermediate [5][7][9], and advanced [7][8] level, respectively. Table V compares their results.

TABLE V
A SURVEY OF EXPERIMENTAL RESULTS IN PREVIOUS WORKS

	Agreements	Disagreements
Stop	Substitutions found in lenis [3–9]	Lenis realized as fortis [3][6] or aspirated [7][8]
	Substitutions found in fortis [3–8]	Fortis realized as lenis [3–7] or aspirated [9]
Affricate	Substitutions found in lenis [3][5][6][8][9]	Lenis realized as aspirated [3][5][6][8] or fortis [5]
Fricative	Substitutions found in lenis [3–5][7–9]	Lenis realized as aspirated [7] or fortis [3–5][8][9]
Liquid	Substitutions found in flap [3][5][8][9]	Flap realized as lateral, retroflex [3][5][8][9] or not mentioned [6][7]
Monophthong	Substitutions found in /ɰ/ [3–5][7–9]	
	/o/ realized as /ʌ/ [5][9]	/ʌ/ realized as /o/ [3][4][9], /o/ realized as /u/ [5], or else [6]
	Substitutions found in /ɯ/ [5][9]	/ɯ/ realized as /ʌ, u/ [5][9] or diphthongs [5]
Diphthong		Variations in diphthongs mentioned [5][9] or not [3]
		Realized as other diphthongs [9], or monophthongs [5][9]
Final consonants	Variation in all final consonants - deletion, and substitution with other consonants [3–9]	Consonant insertion of /n/, /l/ [9]
		Closed syllables showed less variation than open syllables [8]

Their experiments showed agreements and disagreements in the results, which are organized in Table V. For example, it is not yet clear whether variations in lenis stops are realized more frequently as fortis or aspirated stops. The number and nature of disagreements pose difficulty in achieving a consensus over salient variation patterns.

Possible causes of disagreement are differences in the purposes of the experiments, types of data used, and research methods used in phonetics and in education fields. In addition, their methodologies have been limited to contrastive analysis, which sometimes limit their scopes and variables of the experiment, and could not offer holistic description of the phenomenon. In the following Section, we propose an experimental methodology to overcome such limitations.

IV. CORPUS-BASED ANALYSIS

The disagreements shown in Table V open up the necessity for a corpus-based approach. This study proposes to conduct an experiment with a larger number of learners at all levels, consisting of all disagreed aspects mentioned in Section III, in order to find out the prominent variation patterns. Corpus-based approach not only resolves inconsistencies among previous studies, but also enables us to examine any patterns that were undiscovered in previous findings.

A. Corpus

We use L2KSC (L2 Korean Speech Corpus), a speech corpus for Korean as a foreign language spoken by Chinese learners [13]. The corpus was built to evaluate acquisition of phonetic and phonological sounds in Korean language by foreign learners of various L1 backgrounds.

From L2KSC corpus, we analyze 300 words read-speech produced by 53 male and female Chinese learners. The list of words was built for Korean segmental pronunciation learning as a foreign language, based on the vocabulary used in 8 mainstream textbooks [13]. The gender and proficiency, from novice to advanced levels, are balanced in the distribution.

TABLE VI
PLU SET FOR TRANSCRIPTION OF KOREAN SOUNDS

Consonants				Vowels			
PLU	IPA	PLU	IPA	PLU	IPA	PLU	IPA
K	g̊	KQ	k̚	AA	a	AX	ʌ
KK	k͈	KH	kʰ	OW	o	UW	u
T	d̥	TQ	t̚	IY	i	WW	ɯ
TT	t͈	TH	tʰ	EH	ɛ	EY	e
P	b̥	PQ	p̚	UI	ɥi	JA	ja
PP	p͈	PH	pʰ	JX	jʌ	JH	jɛ
Z	ʥ̊	ZZ	ʨ͈	JE	je	JO	jo
CH	ʨʰ	HH	h	JU	ju	WA	wa
S	s	SS	s͈	WH	wɛ	WX	wʌ
M	m	MM	m̚	WE	we	WI	ɰi
N	n	NN	n̚				
NX	ŋ	L	l				
R	ɾ						

TABLE VII
ADDED PLUS FOR TRANSCRIPTION OF KOREAN SOUNDS SPOKEN BY CHINESE LEARNERS

PLU	IPA	PLU	IPA	PLU	IPA
CS	ʦʰ	ZS	ʦ	F	f
SH	ʂ	ZH	ʈʂ	RR	ɹ

B. Transcription

Transcribers only use the units available in PLU (phone-like unit) set shown in Tables VI and VII to transcribe the speech data [14]. We propose to add six Chinese phonemes /ʦʰ/, /ʦ/, /f/, /ʂ/, /ʈʂ/, and /ɹ/ to the Korean PLU set that only contains Korean phonemes. We force-align the data and generate canonical transcriptions using this extended PLU set [14]. We obtain auditory transcription results, which use the extended PLU set to mark the differences compared to automatically-generated canonical transcriptions. Three graduate students with knowledge in phonetics from the Department of Linguistics of Seoul National University performed the transcription.

In order to confirm that our transcribers transcribed reliably, correction rate and pair-wise agreement are calculated. Correction rate is calculated by dividing the number of corrected phones by the total number of phones in force-aligned result. Pairwise agreement is calculated by dividing agreement pairs by the sum of agreement and disagreement pairs between annotators. Correction rate and pairwise agreement are 0.143 and 0.860, respectively. Comparing these figures with those of previous studies' rates, 0.105 and 0.881 [15][16], respectively, we verify the reliability of transcription results in this study.

C. Results and Discussions

TABLE VIII
SALIENT VARIATION PATTERNS OF KOREAN PRODUCED BY CHINESE

	Target Segment	Count	Variation Rate(%)	Realized Target & Corresponding Rate(%)		
Consonants	R	1,604	36.10	L (33.0)	N (1.0)	
	L	4,969	35.10	R (15.0)	NN (5.0)	del.[a](6.6)
	PP	942	27.49	P (20.2)	PH (6.9)	
	ZZ	613	23.49	Z (11.9)	ZS (7.2)	CH (2.9)
	Z	1,694	22.67	ZS (7.9)	ZZ (6.9)	CH (3.7)
	TQ	1,792	21.94	del. (15.6)	KQ (3.8)	
	KK	1,607	21.22	K (15.6)	KH (4.7)	
	T	3,584	17.75	TT (14.0)	TH (3.2)	
	TT	1,840	16.37	T (13.1)	TH (2.8)	
	KQ	2,086	14.96	del. (11.0)	L (1.4)	
Vowels	WI	230	47.16	IY (30.1)	EY (7.9)	UI (3.5)
	JH	47	25.53	JE (19.1)	EH (6.4)	
	WX	240	16.31	UW (4.7)		
	WE	474	14.76	UI (9.1)		
	WH	235	10.74	EH (3.4)		
	JA	570	9.12	AA (5.1)	JX (3.7)	
	JE	331	9.06	EY (4.2)		
	UI	143	7.69	WE (2.8)	UW (2.1)	WI (2.1)
	JU	239	7.53	UI (2.5)	UW (2.1)	JO (2.1)
	JX	1,746	7.10	IY (1.8)	JO (1.6)	
	AX	3,316	6.12	OW (2.8)		
	WA	573	5.76	AA (1.4)	WE (1.2)	
	EH	1,275	4.94	AA (1.5)	AX (1.1)	
	WW	2,078	4.19	AX (1.6)		

[a] del. = Deletion

A confusion matrix is generated to quantify the relation between the canonical and the actual pronunciation. When the realized segment is different from the canonical, it is considered as a variation. The average variation rates for consonants and vowels are 13.74% and 3.35%, respectively. In order to identify salient variation patterns, the phoneme whose variation rate is larger than the average variation rate is considered as major variation pattern and shown in the Target Segment column of Table VIII.

The Realized Target column shows how the overall variation rate is broken down. For example, we see that 33.0% of flap /ɾ/ is realized as lateral /l/, while 15.0% of lateral is realized as flap, and 5.0% of lateral sounds are realized as nasalized coda /n̚/, to mention a few. The ratios quantifying the likelihood of realized phonemes are meaningful because they can be used to model where variation is likely to occur, and when it does, how they are likely to be realized, and thus enabling ASR system to detect the pattern in Chinese learners' speech.

The results help resolve the disagreements as well as discover new variation patterns. Regarding the disagreements mentioned earlier, we clarify that:

- Variations in lenis stops are more likely to be realized as fortis than aspirated stops.
- Variations in affricates are more likely to be realized as fortis than aspirated stops.
- Variations in diphthongs are more likely to be realized as monophthongs than other diphthongs.

For the phonemes with the highest variation rate, /ɾ, l/, Table IX presents context-dependent patterns with possible explanations. Quantitative results of L1, L4, L5, Ins1, and Ins2

TABLE IX
CONTEXT-DEPENDENT VARIATION PATTERNS FOR LIQUIDS PRODUCED BY KOREAN LEARNERS OF CHINESE

No.	Variation Pattern[a]	Rate (Freq.)	Example
R1	R → L/V_V[b]	33.0 (529)	K AX **R** IY → K AX **L** IY
L1	L → R/V_L	15.0 (727)	PP AA **L** L EH → PP AA **R** L EH
L2	L → NN/V_L	5.0 (244)	K WX **L** L IY → K WX **NN** L IY
L3	L → – /V_L	6.6 (243)	N OW **L** L AA T AA → N OW L AA T AA
L4	L → RR/L_V	1.42 (73)	M AX L **L** IY → M AX L **RR** IY
L5	L → RR/V_	1.17 (60)	JX NN PH IY **L** → JX NN PH IY **RR**
Ins1	– → RR/V_R	45.14 (739)	N AA R AA → N AA **RR** R AA
Ins2	– → L/V_R	2.5 (53)	JO R IY → JO **L** L IY

[a] The patterns are represented as [canonical phoneme → realized phoneme / left context_right context]
[b] V = vowels

patterns are new findings of our study that were not discovered in previous studies.

- R1: Since Chinese inventory does not have Korean flap sound, it is realized as lateral.
- L1: Since there is no lateral final in Chinese, it is realized as flap when it occurs in a row at the coda and the onset.
- L2: Korean phonological rule assimilating final nasal with the lateral sequence is not realized.
- L3: Since there is no lateral final in Chinese, it is omitted in pronunciation.
- L4, L5: Chinese retroflex influences the Korean lateral pronunciation, which is not a final in Chinese.
- Ins1, Ins2: Since Chinese inventory does not have Korean flap sound, retroflex and lateral are inserted before realizing the flap, possibly to enable easier articulation by introducing similar sounds of mother tongue in between.

V. CONCLUSION

Despite the increasing demand for learning Korean as a foreign language, there is limited research that systematically analyze Korean speech produced by non-natives. In this paper, as a preliminary study for CAPT systems that use ASR technology, we have provided a quantitative analysis of Korean corpus produced by Chinese learners. We examined contrastive analysis of Chinese and Korean phoneme inventories, surveyed previous studies' predictions and experimental results, and proposed an improvement in experimental methodology.

The disagreements found among previous studies confirmed that while there are common variation patterns for Korean speech produced by non-natives, a systematic approach is necessary. In order to resolve the disagreements, we conducted an experiment with a larger number of vocabulary and learners at all levels. We added L1 phonemes to the PLU set to evaluate their influence in language learning. Aspirated sounds were found to influence more frequently than fortis sounds, thereby resolving some major disagreement.

In addition to resolving the disagreements, we quantified all variation patterns. The most prominent pattern was substitutions among liquid sounds, as flap and lateral sounds showed variations rates of 36.1% and 35.1%, respectively. Context-dependent analysis was conducted to further clarify where the variations occur.

These patterns with statistics can be used for pronunciation modeling in the ASR system to be able to recognize Chinese learners' Korean speech and provide corrective feedback in the CAPT system. This is one of the key contributions we achieved in this paper that can be further extended to future works. Furthermore, it will be interesting to compare how the variation patterns vary among different L1 backgrounds, by applying our methodology to other L1 systems of Korean learning.

ACKNOWLEDGMENT

This work was supported by the ICT R&D program of MSIP/IITP. [10035252, Development of dialog-based spontaneous speech interface technology on mobile platform] This paper is made possible through the help of professor Seok-Chae Rhee from Yonsei University who provided L2KSC corpus for this research.

REFERENCES

[1] H. Hong, S. Kim, and M. Chung, "A corpus-based analysis of Korean segments produced by Japanese learners," in *Speech and Language Technology in Education (SLaTE) 2013*, (Grenoble, France), pp. 189–192, 2013.

[2] F. N. Chen, V. Shivakumar, M. Harikumar, B. Ma, and H. Li, "Large-scale characterization of Mandarin pronunciation errors made by native speakers of European languages," in *INTERSPEECH 2013*, (Lyon, France), pp. 2370–2374, 2013.

[3] J. Cui, "The education of Korean pronunciation for students speaking Chinese," *Bilingual Research*, vol. 20, pp. 309–343, 2002.

[4] J.-Y. Han, J.-S. Choi, H.-Y. Lee, J.-Y. Park, K.-M. Lee, H.-Y. Cho, J. Cui, and S.-W. Yi, *Teaching Korean Pronunciation*. Seoul: Hollym, 2003.

[5] Y. Qin, "A study on the plans for teaching Korean pronunciaiton to Chinese learners," thesis, Konkuk university, Seoul, Korea, 2010.

[6] H. Cho, "A study on the teaching method of Korean pronunciation for foreigners - focused on Chinese-speaking learners of Korean," *Korean Journal of General Education*, vol. 7, no. 6, pp. 531–559, 2013.

[7] M. Hwang, *A Study on Errors in Consonant Pronunciation by and an Educational Approach for Korean Language Learners from Chinese-Speaking Regions*. Thesis, Chungnam National University, Daejon, Korea, 2012.

[8] L. Leng, *Chinese Pronunciation of Consonants Education of Korea Students*. Thesis, Chungnam National University, Daejon, Korea, 2014.

[9] K.-D. Kim, *A Study on Teaching Korean Pronunciation to Chinese Speakers*. Thesis, Dankook University, Chonan, Korea, 2008.

[10] H. Lee, *Korean Phonetics*. Seoul: Taehaksa, 1996.

[11] Y.-H. Lin, *The sounds of Chinese*. New York: Cambridge University Press, 2007.

[12] J. E. Flege, "Second language speech learning: Theory, findings, and problems," *Speech perception and linguistic experience: Issues in cross-language research*, pp. 233–277, 1995.

[13] S.-C. Rhee, J.-A. Kim, and J.-W. Chang, "Design and construction of speech corpus for Korean as a foreign language (L2KSC)," *The Journal of Chinese Language and Literature*, vol. 33, pp. 35–53, 2005.

[14] H. Hong and M. Chung, "Improving phone recognition performance via phonetically-motivated units," in *INTERSPEECH 2009*, (Brighton, UK), pp. 1663–1666, 2009.

[15] H. Ryu, K. Lee, S. Kim, and M. Chung, "Improving transcription agreement of non-native English speech data transcribed by non-natives," in *Speech and Language Technology in Education (SLaTE) 2011*, (Venice, Italy), pp. 61–64, 2011.

[16] S. Goddijn and D. Binnenpoorte, "Assessing manually corrected broad phonetic transcriptions in the spoken Dutch corpus," in *15th ICPhS 2003*, (Barcelona, Spain), pp. 2981–2984, 2003.