INTERSPEECH 2020
October 25–29, 2020, Shanghai, China

Perception and Production of Mandarin Initial Stops by Native Urdu Speakers

Dan Du[1], Xianjin Zhu[2], Zhu Li[1], Jinsong Zhang[1]

[1]Beijing Advanced Innovation Center for Language Resources
Beijing Language and Culture University, Beijing, China
[2]The Robotics Institute, Harbin Institute of Technology, Harbin, China
danmaxine@163.com, zhuxianjin@sina.cn, lz980213@163.com, jinsong.zhang@blcu.edu.cn

Abstract

In the area of second language (L2) acquisition, studies of stop consonants have focused on those first language (L1) and L2 with a two-way stop contrast and a three-way stop contrast, few are about languages with a four-way stop contrast. The current study, mainly concerning VOT, investigates how native speakers of a language with a four-way stop contrast acquire the two-way stop contrast in L2. Mandarin presents a two-way stop contrast, which is primarily differentiated by VOT, whereas Urdu presents a four-way stop contrast. Speech perception and production experiments are designed to explore L2 learners with a more complex language system learning a relatively simple language system, and the results show that the speech perception of Mandarin initial stops by native Urdu speakers has no significant difference compared with those by Chinese speakers while the speech production by native Urdu speakers is significantly different from those by Chinese speakers. It demonstrates that L1 exerts different influences on L2 perception and production separately and sometimes good L2 perception doesn't mean good L2 production. This study has made some contribution about the four-way stop contrast in L2 acquisition and will shed light on L2 learning.

Index Terms: two-way stop contrast, four-way stop contrast, L2 speech perception and production, VOT

Introduction

Stop consonant (or plosive) has played a significant role in languages since it exists in all world languages [1]. Previous research has shown that VOT, CF0 (consonant-induced F0) [5, [6] and voice quality [7] are relevant acoustic cues of stop contrast. A great many of cross-language studies [8-10] have provided ample evidence that VOT is the main acoustic correlate of laryngeal distinction in voicing, and it has been used to separate initial stop consonants of different languages into two-way stop contrast, three-way stop contrast, four-way stop contrast and more than four-way stop contrast [1, 8, 9]. Such phonological categories, to some extent, have represented challenges in the area of L2 acquisition [10]. This study, mainly based on VOT, aims to explore the perception and production of Mandarin initial stops by native Urdu speakers.

Considerable studies over the past few decades have provided strong evidence that L1 exerts an influence on the L2 acquisition, especially when L1 and L2 possess different groups of stop category. Numerous studies have been focusing on the L2 acquisition of a language with a two-way stop contrast by L1 speakers of languages that also have a two-way stop contrast. VOT mismatch and compromise of stop consonants are two main performances of unsuccessful L2 acquisition. On the one hand, mismatch often happens when subjects make clear-cut substitutions of one phoneme for another. For example, Simon [11] examined the acquisition of the English laryngeal system by native Dutch learners. Both languages have a two-way laryngeal system, but while Dutch contrasts pre-voiced with short-lag stops, English has a contrast between short-lag and long-lag stops. Dutch learners often transfer their L1 pre-voiced stops into short-lag stops in L2 English speech production. Similarly, Flege found that Spanish monolinguals consistently identified long-lag English [t^h] tokens as /t/ in a two-alternative forced-choice test [12]. On the other hand, compromise appears when subjects' speech production is a kind of phonetic distortion of the target phoneme. For instance, native English speakers often tend to produce voiceless Spanish stops with VOT values that are too long (i.e., with too much aspiration) [13] while Saudi Arabians predictably under-aspirate English /p^h, t^h, k^h/ even they have lived for more than two years in the United States [14]. Flege and Hillenbrand also found that adult French learners of English produced English /t^h/ with compromised VOT values that were intermediate to the short-lag values observed for French monolinguals and the long-lag values observed for English monolinguals [15]. Similarly, the French subjects realized English /t^h/ with significantly shorter (French-like) VOT values than English monolinguals; and the native English subjects realized /t/ in French words with longer (English-like) VOT values than French monolinguals [16]. All in all, previous studies have shown that many adult learners produce English /p^h, t^h, k^h/ with significantly shorter VOT values than English monolinguals, but with longer VOT values than monolingual native speakers of the learners' L1 [17-19].

In addition, there are also some studies concerning the L2 acquisition of languages with a three-way stop contrast. Jeffrey J. Holliday [20] indicated that some learners were unable to produce the Korean contrast in a native-like way even after one year of intensive L2 instruction through a longitudinal study of Chinese speakers with a two-way stop contrast acquire Korean with a three-way stop contrast. Korean is a language with short-lag, medium-lag and long-lag VOT respectively, and in some perceptual assimilation studies that have tested L1 listeners from languages with a two-way stop contrast, such as Japanese [21], English [22], and Mandarin [23], it was found that Korean fortis stops were perceived as unaspirated stops (or voiced or short-lag VOT) in the vast majority of trials, and both Korean lenis and aspirated stops were perceived as aspirated (or voiceless or long-lag VOT), with lenis stops sometimes being a slightly less good category fit depending on the place of articulation and vowel context. Ying Ma [24], through acoustic experiments, has discovered that the VOT of Mandarin voiceless stops produced

http://dx.doi.org/10.21437/Interspeech.2020-1921

by native Korean speakers, influenced by Korean lenis and fortis, is between Mandarin voiceless stops and Korean lenis.

As we can see, there are voluminous academic work contribute to L2 acquisition of languages with a two-way stop contrast and some work about languages with a three-way stop contrast, but few are about languages with a four-way stop contrast. Therefore, the current study investigated the case of L1 speakers of Urdu, a language with a four-way stop contrast among negative long-lag, negative short-lag, short-lag and long-lag VOT stops (i.e. voiced aspirated, voiced unaspirated, voiceless unaspirated and voiceless aspirated stops), learning Mandarin, a language with a two-way stop contrast between short-lag and long-lag stop contrast. As shown in Table 1, there are 18 stops in Urdu, and they are all four-way stop contrast except the uvular /q/ and glottal /ʔ/. There are two noteworthy features of the inventory of consonant phonemes in Urdu, among which the distinction between voiced aspirated and unaspirated stops contribute to its difference from its European cousins [25]. Concerning the topic of this study, only bilabial, dental and velar stops in Urdu are considered. There are 6 stops in Mandarin as shown in Table 2, and they are differentiated by aspiration into two groups, voiceless aspirated and unaspirated stops. Aspiration of Mandarin stops represents a great challenge for those who learn Mandarin as an L2. It's said that Kenneth L. Pike can't pronounce Mandarin aspirated stops well even though he is a famous American linguist with excellent articulatory skills and profound theoretical knowledge [26]. Aspirated stops are also one of the difficulties for Japanese speakers [27] and European language speakers [28]. In the process of annotation, it's also found that aspiration is difficult for Urdu speakers, for instance, they often pronounce Mandarin /pʰ/ as /p/. Although both languages share similar voiceless stop category, there are still differences between them, such as orthography [29], subtle differences of VOT values, etc. The purpose of the study is to examine what challenges the differences between Mandarin and Urdu stops have given rise to in L2 acquisition.

Table 1: *Stops of modern standard Urdu* [29].

	Bilabial	Dental	Retroflex	Velar	Uvular	Glottal
unaspirated	p	t	ʈ	k	q	ʔ
unaspirated	b	d	ɖ	g		
aspirated	bʰ	dʰ	ɖʰ	gʰ		
aspirated	pʰ	tʰ	ʈʰ	kʰ		

Table 2: *Stops of standard Mandarin Chinese* [30].

	Bilabial	Dental	Velar
unaspirated	p	t	k
aspirated	pʰ	tʰ	kʰ

Design and Methodology

Two experiments have been conducted in this part, one perception and another production of Mandarin initial stops by native Urdu speakers. All experiments were programmed and presented to participants using Matlab.

1.1. Perception

2.1.1 Participants

The participants were 15 native Chinese-speaking adults (9 females, 6 males, age range: 22 to 24 years) and 15 native Urdu-speaking adults (12 males, 3 females, age range: 21 to 29 years) with no self-reported history of speech, language, or hearing problems. All Chinese participants are graduate students, and all Urdu participants consist of undergraduate and graduate students with HSK 5. Informed written consents were obtained from all participants before the experiment begins.

2.1.2 Stimuli

VOT values of stops vary even in the same category, thus the materials used in the experiment below were derived from naturally reading materials of Mandarin monosyllables. These monosyllables are chosen from the BLCU-SAIT speech corpus, including those speech pronounced by one male and one female who have passed the *Putonghua Shuiping Ceshi*, the national standard Mandarin proficiency test in China, with the first-level certificate. Speech materials are consonant-vowel (CV) syllables (as shown in Table 3), which were used to make continua from aspirated sounds to unaspirated counterparts by cutting back VOT values in equal steps. Every continuum has ten stimuli, and they are constructed through Praat script [31].

Table 3: *CV syllables used for speech continuum synthesis.*

	/a/
/pʰ/	[pʰa]
/tʰ/	[tʰa]
/kʰ/	[kʰa]

2.1.3 Procedure

A familiarization task was presented before each identification and discrimination task to make participants know how to manage the computer and the experiment procedure.

In the identification test, participants were presented with single tokens chosen randomly from the VOT continua. For each token presented to them, they had to decide which one they heard and mouth-clicked on the right answer shown on the computer screen (e.g. ba and pa; da and ta; or ga and ka). For each continuum, participants were presented with 3 repetitions of each token in the test phase. This made a total of 90 trails (three syllables × ten steps × three repetitions) presented in a random order. The intertrial interval was set to 1s both in the training phase and the testing phase. Participants were asked to label the target syllable with a forced choice between two Chinese syllables on the screen. They could listen to each language sample as often as they liked.

In the discrimination test, the method of Peng [32] were adopted, stimuli were presented in pairs. For each speech continuum, a total of 26 pairs, 16 consisted of two different stimuli separated by two steps on the speech continuum, in either forward (1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, 8-10) or reverse order (3-1, 4-2, 5-3, 6-4, 7-5, 8-6, 9-7, 10-8), and 10 consisted of the 10 stimuli on the speech continuum each paired with itself (same pairs). The above 26 pairs were repeated to each participant for three times randomly. After hearing each pair, participants were instructed to judge whether the two stimuli were identical or not, and to mouse-clicked on a "same" or "different" option on the computer screen. The identification and discrimination tasks took approximately 30 to 40 minutes.

1.2. Production

The production data in this part was also collected from the BLCU-SAIT speech corpus, and participants consist of 40 speakers (20 Chinese speakers and 20 Urdu speakers, age range: 22 to 29 years. Urdu speakers all have the certificate of HSK 5.), and none reported any speech or hearing impairment at the time of testing. Informed written consents were obtained from all participants before the recording begins. All the production materials have been labeled by three graduate students with phonetics background, and only correct pronunciation of initial stops have been adopted in the experiment. Thus, voicing errors or place of articulation errors were not included in the analyses. In a nutshell, 5841 Mandarin monosyllables beginning with stops in the initial position were considered in the production experiment. These syllables have covered almost all possible vowel context with stops in Chinese. Their VOT were measured by Praat scripts after adjusting all results of automatic segmentation carefully.

Results

The perception and production data of the above two experiments have been collected.

Figure 1 and Figure 2 are the results of speech production experiment. We can see from the Figure 1 that the mean VOT of Mandarin aspirated initial stops produced by Urdu speakers are smaller than those produced by Chinese speakers ($p<0.01$), while the mean VOT of Mandarin unaspirated initial stops produced by Urdu speakers are larger than those produced by Chinese speakers ($p<0.05$), and an independent sample *t*-test confirmed their significant difference. Urdu speakers have produced more outliers than Chinese speakers, which indicates that their VOT production of Mandarin initial stops has more disturbance.

From Figure 2 (a) and (b), we know that both Mandarin aspirated and unaspirated initial stops produced by Urdu speakers have wider VOT distribution than those produced by Chinese speakers. The VOT range of unaspirated stops is from 20ms to 160ms and those of aspirated stops 20ms to 220ms produced by Urdu speakers while the VOT range of unaspirated stops is from 20ms to 120ms and those of aspirated stops 60ms to 200ms produced by Chinese speakers. And the VOT between aspirated and unaspirated stops produced by Urdu speakers has much more overlap. The different peaks of aspirated and unaspirated stops in Figure 2 (a) and (b) indicate that the VOT of Mandarin aspirated initial stops produced by Urdu speakers are smaller than those by Chinese speakers, but unaspirated ones are the opposite. This result is in keeping with what shown in Figure 1.

The results of speech perception experiment are shown in Figure 3 (a) to (c). From the three figures, we can see that as for the continuum /pa/-/p^{h}a/ and /ta/-/t^{h}a/, the category boundaries for Chinese and Urdu speakers are both between stimulus 5 and 6, and the categorical boundaries of Urdu learners falls at a later point than those of Chinese speakers of these two continua. For continuum /ka/-/k^{h}a/, the categorical boundary is approaching stimulus 5 for native Urdu speakers and between stimulus 5 and 6 for Chinese speakers. Although the category boundaries of continuum /ka/-/k^{h}a/ between Chinese and Urdu speakers varies more than those of continuum /pa/-/p^{h}a/ and /ta/-/t^{h}a/, there is no significant difference of the category boundaries between Chinese and Urdu speakers ($p>0.05$). The discrimination peaks of continuum /ka/-/k^{h}a/ of Chinese and Urdu speakers differs more widely than those of continuum /pa/-/p^{h}a/ and /ta/-/t^{h}a/. The discrimination peaks of these three continua also have no significant difference between Chinese and Urdu speakers ($p>0.05$). Taken together, it can be speculated that Urdu learners have formed the ability of categorical perception of Mandarin aspirated and unaspirated initial stops, and such ability is nearly same with Chinese speakers.

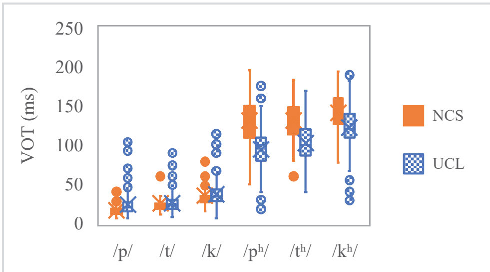
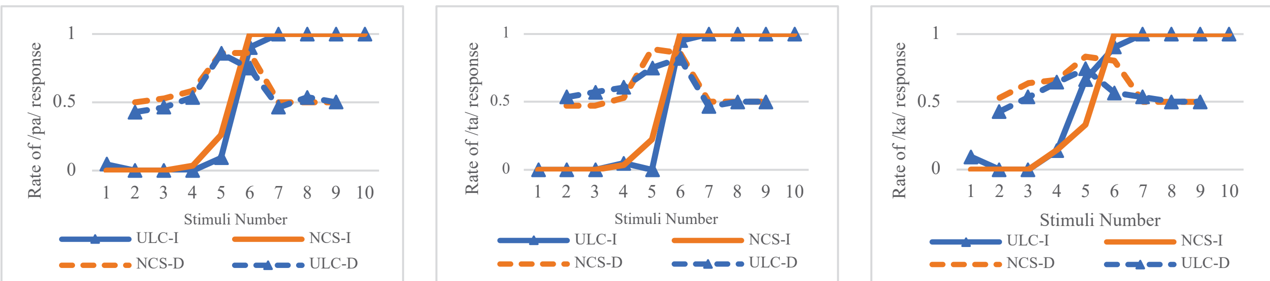


Figure 1: *Box plot of VOT values of initial stops in Mandarin monosyllables produced by native Chinese speakers (CNS) and Urdu learners of Chinese (ULC).*

(a) (b)

Figure 2: *The frequency of VOT values of Mandarin monosyllables produced by Chinese speakers (a) and Urdu speakers (b).*



(a) (b) (c)

Figure 3: *Identification (solid lines) and discrimination (dashed lines) curves of continuum /pa/ - /p^{h}a/ (a), /ta/ - /t^{h}a/ (b) and /ka/ - /k^{h}a/ (c). Blue lines represent Urdu learners of Chinese (ULC) and Orange lines serve for native Chinese speakers (NCS).*

Discussion

The current study has examined the perception and production of Mandarin initial stops by native Urdu speakers. Speech perception and production experiments have been applied to investigate their L2 acquisition. In this study, we have taken one step further toward the four-way stop contrast in the area of L2 acquisition.

The perception results indicate that the perception of Mandarin initial stops by Urdu and Chinese speakers is similar as there is no significant difference between identification and discrimination of stop continua by them, which demonstrates that it's relatively easy for Urdu speakers to perceive Mandarin initial stops and they can master the perception of these phonemes after a period of learning. As there are four-way stop contrast in Urdu, including the two laryngeal types in Mandarin, Urdu learners are at a higher start point in learning Mandarin initial stops than others whose languages have totally different laryngeal types from Mandarin. The stop consonant system of Urdu has already had voiceless aspirated and unaspirated stops, which enable Urdu speakers advantages to perceive Chinese counterparts. To a certain extent, the laryngeal system of Urdu facilitates their perception of Mandarin initial stops. The results, to a certain extent, also conform with a theory that those languages with a more complex language system will learn a language with a more simple language system easily [33].

In comparison, the production of Mandarin initial stops of Urdu and Chinese speakers is significantly different, which demonstrates that Urdu speakers are poor in producing Mandarin initial stops authentically. In other words, their production of Mandarin initial stops is accented. Although there are also voiceless aspirated and unaspirated stops in Urdu, they are different from those of Mandarin in the duration of VOT [34, 35]. As the VOT of Urdu voiceless stops (aspirated stops are shorter while unaspirated ones are longer) are different from Mandarin counterparts, Urdu learners of Mandarin tend to compromise VOT affected by their L1. Furthermore, we have evidenced that Mandarin initial aspirated stops produced by native Urdu speakers differ much more from those produced by Chinese speakers compared with the unaspirated ones. Because, on the one hand, there is a very strong favoring of voiceless unaspirated stops over the voiceless aspirated ones in Urdu [29]. Whereas the full utilization of the four-way opposition of stops in the phonological paradigm of Urdu is a response to the communicative need for distinctive units, the marked skew in the frequency of usage for the four stop types in the language system is a response to the physiological-acoustic and human behavior constraints in their production. On the other hand, compared with unaspirated stops, aspirated ones are more marked [36], which, to some extent, prevents the target-like speech production of L2 learners.

Finally, the results show something about the relationship between speech perception and production. Previous studies have demonstrated mixed results of the relationship between speech perception and production. Former studies have shown evidence of both association and dissociation between two modalities. Perceptual Assimilation Model (PAM) [37, 38] and Speech Learning Model [39] suggest a close relation between perception and production. They argued that those who have excellent L2 speech perception are also good at L2 speech production, and they indicated that L2 production ability can be reflected by L2 perception. However, some studies found L2 learners' perception and production accuracy of non-native sounds sometimes are not correlated or moderately correlated [40, 41]. According to Strange [42], L2 segmental production and perception by experienced speakers of an L2 may be "uncorrelated". Congruent with the latter opinion, this study showed little relationship between perception and production of Mandarin initial stops by native Urdu speakers.

Conclusion

We have examined the VOT perception and production of Mandarin initial stops by native Urdu speakers in this study. The question was explored through perception and production experiments. Overall, the results demonstrated that the perception of Mandarin initial stops of Urdu speakers is close to those of Chinese speakers while the production of Mandarin initial stops by native Urdu speakers has significant difference from those by Chinese speakers. These results suggest that when two languages use VOT to cue stop contrasts differently, L1 speakers of one language learning the other one as an L2 will often fail to produce L2 stops with native-like VOT. For learners of L1 with a four-way stop contrast learning L2 with a two-way stop contrast, the speech perception precedes production and there is little relationship between speech perception and production. This study has made some contribution about the four-way stop contrast in L2 acquisition and will shed some light on L2 acquisition.

Acknowledgments

This study was supported by Advanced Innovation Center for Language Resource and Intelligence (KYR17005), the Science Foundation and Special Program for Key Basic Research fund of Beijing Language and Culture University (the Fundamental Research Funds for the Central Universities) (20YJ040002, 16ZDJ03), Research Funds of Scientific Research Project (17YPY054), the Fundamental Research Funds for the Central Universities, and the Research Funds of Beijing Language and Culture University (20YCX134). Jinsong Zhang is the corresponding author.

References

[1] C. Henton, P. Ladefoged, I. Maddieson, "Stops in the world's languages," *Phonetica*, vol. 49, no. 2, pp. 65–101, 1992.

[2] J. L. Schwartz, L. J. Boë, P. Badin, T. R. Sawallis, "Grounding stop place systems in the perceptual-motor substance of speech: On the universality of the labial-coronal-velar stop series," *Journal of Phonetics*, vol. 40, no. 1, pp. 20–36, 2012.

[3] I. Maddieson, *Consonant Inventories*, The World Atlas of Language Structures Online, Available online at https://wals.info/, 2013.

[4] M. K. Gordon, *Phonological Typology*. Oxford: Oxford University Press, 2016.

[5] J. Kingston, R. L. Diehl, C. J. Kirk, et al. "On the internal perceptual structure of distinctive features: The voice contrast," *Journal of Phonetics*, vol. 36, no. 1, pp. 28-54, 2008.

[6] J. P. Kirby, D. R. Ladd, "Effects of obstruent voicing on vowel F0: Evidence from "true voicing" languages," *Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 2400–2411, 2016.

[7] R. P. Dixit, "In defence of the phonetic adequacy of the traditional term 'voiced aspirated'," *UCLA Working Papers in Phonetics*, vol. 67, pp. 103–111, 1987.

[8] L. Lisker, A. Abramson, "A cross-language study of voicing in initial stops: Acoustical measurements," *Word*, vol. 20, pp. 384–422, 1964.

[9] T. Cho, D. H. Whalen, G. Docherty, "Voice onset time and

beyond: Exploring laryngeal contrast in 19 languages," *Journal of Phonetics*, vol. 72, pp. 52-56, 2019.

[10] R. Lado, *Linguistics across Cultures: Applied Linguistics for Language Teachers,* Ann Arbor: University of Michigan Press, 1957.

[11] E. Simon, "Acquiring a new second language contrast: an analysis of the English laryngeal system of native speakers of Dutch," *Second Language Research*, vol. 25, no. 3, pp. 377-408, 2009.

[12] J. E. Flege, "Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language," *Journal of the Acoustical Society of America*, vol. 89, no. 1, pp/ 395-411, 1991.

[13] R. Stochwell, J. Bowen, *The Sounds of English and Spanish*, Chicago: University of Chicago Press, 1965.

[14] J. E. Flege, "Phonetic approximation in second language acquisition," *Language Learning*, vol. 30, pp. 117-134, 1980.

[15] J. E. Flege, J. Hillenbrand, "Limits on phonetic accuracy in adult foreign language speech production," *Journal of the Acoustical Society of America*, vol. 76, pp. 708-721, 1984.

[16] J. E. Flege, "The production of "new" and "similar" phones in a foreign language," *Journal of Phonetics*, vol. 15, pp. 47-65, 1987.

[17] R. F. Port, F. M. Mitleb, "Segmental features and implementation in acquisition of English by Arabic speakers," *Journal of Phonetics*, vol. 11. No. 3. pp. 219-229, 1983.

[18] G. Nathan, "On second-language acquisition of voiced stops," *Journal of Phonetics*, vol. 15, pp. 313-322, 1987.

[19] R. C. Major, "English voiceless stop production by speakers of Brazilian Portuguese," *Journal of Phonetics*, vol. 15, pp. 197-202, 1987.

[20] J. J. Holliday, "A longitudinal study of the second language acquisition of a three-way stop contrast," *Journal of Phonetics*, vol. 50, pp. 1-14, 2015.

[21] T. Yasuta, "Stop perception in second language phonology: Perception of English and Korean stops by Japanese speakers (doctoral dissertation)," University of Hawaii a, Manoa, 2004.

[22] A. M. Schmidt, "Cross-language consonant identification: English and Korean," *In O.-. S. Bohn & M. Munro (Eds.), Language experience in second language speech learning: In honor of James Emil Flege* (pp. 185–200). Amsterdam: John Benjamins, 2007.

[23] J. J. Holliday, "The perceptual assimilation of Korean obstruents by native Mandarin listeners," *Journal of the Acoustical Society of America*, vol. 135, no. 3, pp. 1585–1595, 2014.

[24] Y. Ma, "The experiment study of Mandarin unaspirated stops by native Korean learners (master thesis)," East China Normal University, 2019.

[25] B. Comrie, *The World's Major Languages,* Britain, Routledge, 2009.

[26] Y. C. Chen, "The position and learning of aspirated stops in foreign language teaching," *Language Teaching and Linguistic Studies*, no. 3, pp. 61-73, 1983.

[27] C. Zhu, "The comparison of Chinese and Japanese phoneme inventories," *Language Teaching and Linguistic Studies*, no. 4, pp. 77-90, 1981.

[28] N. F. Chen, D. Wee, R. Tong, B. Ma, H. Z. Li, "Large-scale characterization of non-native Mandarin Chinese spoken by speakers of European origin: Analysis on iCALL," *Speech Communication*, vol 84, pp. 46-56, 2016.

[29] W. Reid, R. Otheguy, N. Stern, *Signal, Meaning, and Message: Perspectives on Sign-based Linguistics.* Amsterdam, John Benjamins Publishing, 2002.

[30] T. Lin, L. J. Wang, *Phonetic Course*, Beijing: Peking University Press, 2013.

[31] P. Boersma, D. Weenink, Praat: Doing phonetics by computer. http://www.fon.hum.uva.nl/praat/, 2009.

[32] G. Peng, H. Y. Zheng, T. Gong, et al. "The influence of language experience on categorical perception of pitch contours," *Journal of Phonetics*, vol. 38, no. 4, pp. 616-624, 2010.

[33] B. G. Evans, W. Alshangiti. "The perception and production of British English vowels and consonants by Arabic learners of English," *Journal of Phonetics*, vol. 68, pp15-31, 2018.

[34] K. Shimizu, "A cross-language study of voicing contrasts of stops," *Studia Phonologia*, vol. 23, pp. 1-12, 1989.

[35] Z. J. Wu, M. C. Lin, *Experimental Phonetics*, Beijing: Higher Education Press, 2014.

[36] F. R. Eckman, "Markedness and the contrastive analysis hypothesis," *Language Learning,* vol. 27, pp. 315-330, 1977.

[37] C. T. Best, "The emergence of native-language phonological influences in infants: A perceptual assimilation model," In C. Goodman & H. Nusbaum (Eds), *The development of speech perception*, pp. 167-224, Cambridge: The MIT Press, 1994.

[38] C. T. Best, "A direct realist view of cross-language speech perception," *Speech Perception and Linguistic Experience: Issues in Cross-Language Research,* pp. 171-204, Baltimore: York Press, 1995.

[39] J. E. Flege, "Second language speech learning: Theory, findings, and problems," *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, pp. 233–277, Baltimore: York Press, 1995.

[40] S. Peperkamp, C. Bouchon, "The relation between perception and production in L2 phonological processing," *in INTERSPEECH 2011 – 12th Annual Conference of the International Speech Communication Association*, August 27-31, Florence, Italy, Proceedings, pp. 161-164, 2011.

[41] M. L. Zampini, "The relationship between the production and perception of L2 Spanish stops," *Texas Papers in Foreign Language Education*, vol. 3, no. 3, pp. 85-100, 1998.

[42] W. Strange, "Phonetics of second-language acquisition: past, present, future," *in ICPhS 1995 - 13th International Congress of Phonetic Sciences*, August 14-19, Stockholm, Sweden, Proceedings, pp. 76-83, 1995.