Cross-Linguistic Study on VOT of Chinese Trilingual Speakers

SUN Yue[a],[*]; Stefania Profita[b]

[a] Institute of Innovation & Entrepreneurship, Southwest University, China.
[b] College of International Studies, Southwest University, China.
[*] Corresponding author.

Abstract

In the field of third language acquisition (TLA) in China, only a few empirical researches were carried out to discuss the negative transfer from the mother tongue, so this work aimed to examine that if the cross-linguistic influence (CLI) from L2 exists. We measured and analyzed 8 Chinese college students' voice onset time (VOT) of stops /p t k b d g/ in English (L2) and Italian (L3). Two Italian native speakers' VOT values were taken as the reference group. The result shows that Chinese students can hardly distinguish unaspirated voiceless stops /p t k/ and voiced stops /b d g/ in Italian because students are affected by Chinese (L1)'s stop system which is characterized by aspiration. Pre-voicing was observed in voiced stops /b d g/ in both L2 and L3. The analysis of variance shows they are similar (P>0.05). Based on this result, we discussed the possibility to develop the Speech Learning Model (SLM) which was brought out on account of second language acquisition of phonetics and adopt it to explain the learning of L3 speech.

Key words: Third language acquisition; Plosives; Voice onset time; Speech learning model

INTRODUCTION

In the era of globalization and information technology, the opportunity to engage with other languages is greater than ever. People's demand for multiple language learning is growing and TLA is also gradually concerned. However, its mechanism is more complicated than SLA as trilingual learners have more knowledge and learning experience of languages than bilinguals. Any language variables may affect the learning of the target language, but are they equally activated in language learning and why? In many countries, most of the studies adopted phonetic experimental methods to study TLA among Indo-European languages. (Temblay,2006; Sypiańska & Olende,2013; Llama et al.,2010; Wunder,2010) In China, most of the empirical studies on pronunciation focus on the negative mother tongue (Chinese) effect on L3 (Zhang,2012) but rarely discuss the second language effect on L3. Voice onset time（VOT） can distinguish different plosives and it is easy to be measured. In this paper, we measured and analyzed the English (L2) and Italian (L3) VOT values of Chinese students who come from the same Italian class in a Chinese university to figure out the role of L2 in TLA and to test whether SLM can explain the result well.

1. LITERATURE REVIEW

1.1 The Concept of Third Language Acquisition

In recent years, multilingual acquisition has become an independent research field that gradually separated from SLA. Cenoz (2003, p.39) defined multilingual acquisition as the process of acquiring two or more languages. On the other hand, De Angelis (2007, pp.3-7) pointed out that the mechanism of multilingual acquisition is more complicated than SLA, because multilingual have more knowledge, learning experience than bilingual and more language variables could affect the learning of the target language. Herdina and Jessner (2000, pp.85-92) think that it had 5 features: Non-liner, language maintenance, individual variation, interdependence, and endogenous changes.

At present, the definition of TLA is controversial. The argument is whether the term 'L3' should cover all non-native languages. In this paper, we only discussed the third language other than L4 or L5, So we adopted Cenoz's (2001) definition of TLA :

TLA specifically refers to the acquisition process of the third language and the first two languages can be learned simultaneously or sequentially.

1.2 Current Studies of Cross-Linguistic Influence in Phonology

A large number of studies have shown that both L1 and L2 had some potential impacts on TLA, especially when learners' L3 is at low-proficiency level. As the level rises, the impact decreases. The cross-linguistic influence can be reflected in many aspects, such as vocabulary, grammar, pronunciation etc. Many factors are involved in TLA such as linguistic distance, psychotypology, L2 status, proficiency, recency, and exposure, etc. (cf. De Angelis, 2007; Cenoz, 2001).

The previous studies on TLA of phonetics were inconsistent, there are three opinions in major. Ringbom (1987) claimed that the mother tongue is the source language that has the greatest influence on non-native language in pronunciation, even though L1 and L3 do not share many similarities in typology. Many researchers support his point of view (Hammarberg & Hammarberg, 2009; Pyun, 2005). Some researchers believe that learners tend to activate the second language they have just learned to transfer, which is called the L2 status effect. (Hammarberg, 1998; Llama et al.,2010; Gut, 2010). Others hold the opinion that all component languages of multilingual subjects interact and influence one another in the global language entity which means L3 may also have an adverse effect on L1 and L2 (Sypiańska & Olende; Wrembel, 2014). There are different languages involved in trilingual learners but most of them are Indo-European languages, Chinese, Japanese, Korean and so on are seldom involved.

In China, the study of stops mainly focuses on dialects and EFL (English as a Foreign Language). EFL studies show that devoicing is the most frequent error of Chinese learners (Dai, 2012; Gao & Zhang, 2018). Although Chinese learners have a certain sense of distinguishing voiced and voiceless stops, their mother tongue still has a great impact on their pronunciation. There are a few studies on TLA of stops as well. For example, Zhang (2012) found that Chinese learners were still affected by the native phonetic categories in their L3 production by measuring and comparing the VOT between Chinese and French stops. But the effect of L2 was not explored.

In the field of second language speech acquisition, Flege (1995) proposed a speech learning model (SLM). He claimed that in the early stage of second language learning, it was difficult for learners to distinguish similar phonemes in L1 and L2, which may be that they do not perceive the differences between the similar phonemes,

or they perceive the differences but mistaken them as the same phoneme, resulting in pronunciation errors. In short, the more similar, the harder to acquire. But this model did not clarify how learners studied the different sounds absent from their phonetic categories. Recently, Flege (2018) claimed that the phonetic categories of interlanguage could move and got closer to native speakers by time due to the increase of output in SLM-r.

1.3 Voice Onset Time and Stops

VOT is the most commonly used acoustic parameter to distinguish stops. VOT refers to the time from the off-glide to the beginning of vocal cord vibration. If vocal cords vibrate first and then the obstruction is removed, then they are called voiced stops with negative VOT values. On the contrary, the stops which remove the obstruction first then the vocal cord begins to vibrate are voiceless stops and the VOT values are positive.

The stop is a consonant category shared in all languages (Jakobson, 1958). According to VOT, they are basically divided into voiced (with pre-voicing) and voiceless stops (without pre-voicing) and voiceless stops are divided into aspirated and unaspirated ones. However, the phonetic features of stops in each language are slightly different (Pierrehumbert, Beckman & Ladd 2000, p.285). Owing to individual and regional differences, there is no standard pronunciation of stops which are accurate to a certain value.

For foreign language learners, their learning object is blurry. Cho and Ladefoged (1999, p.223) explored a plausibly range of VOT values of stops but cannot give out accurate divines in the phonetic sense. We think native speakers' VOT values are in a comprehensible range which do not lead to misunderstanding, so should learners follow this rule and make sure their pronunciation meets the range.

In cross-linguistic studies, the level of learners' speech usually examined in two sides: pronunciation and content. In this paper we focus on learners' pronunciation and take native speakers' VOT range as reference. The VOT of L2 (English) and L3 (Italian) of trilingual speakers from China were measured and analyzed to answer the questions below:

• Is there any evidence shows that L1 (Chinese) or L2 (English) have CLI on L3 (Italian) speech production of learners?

• Can SLM explain the CLI phenomenon in TLA well?

2. EXPERIMENT

2.1 Participant

Participants included 6 trilingual speakers from the same Italian class of Southwest University in China (3 males and 3 females, with an average age of 20). They sequentially studied English and Italian and the average time of learning English is 10.5 years, among which 5

students have passed CET-4 (College English Test-4). The average time of learning Italian is three months and their teacher is an Italian native speaker. Two Italian speakers are in the control group, one from the North (Milan) and one from the South (Palermo).

2.2 Stimuli

The stimuli consisted of two words lists with 60 target words in English and Italian, in which 6 stops / p t k b d g / in stressed onset position were contained. Check the appendix for details. Before the experiment, the speaker had some time to get familiar with the word lists in which words were randomized and embedded in carrier phrases in the English or Italian (I am saying…, Io dico…), each word is required to read three times.

2.3 Method

Participants were required to read each word at a normal speed. Chinese speakers are required to read the English and Italian word lists once, while Italian speakers only need to read the Italian word list once.

In this experiment, the voice was recorded in a quiet room, professional recording software Cool Edit Pro 2.0 was used, sampling frequency 22050hz, 16-bit, mono. The voice analysis software PRAAT6.1.08 was used to measure VOT values. 829 tokens which had excluded mispronunciation were analyzed in SPSS (22.0) and mapped in Excel.

Table 2
Reference values of stops' VOT (ms)

Plosives	/p'/	/t'/	/k'/	/p/	/t/	/k/	/b/	/d/	/g/
L1 (Mandarin)	105.6	103.6	111.7	12.9	13.4	30.1	---	---	---
L2 (English)	58.0	70.0	80.0	---	---	---	1.0/-101.0	5.0/-102.0	21.0/-88.0
L3 (Italian)	---	---	---	23.2	31.2	41.8	-82.4	-84.8	-100.9

It can be seen from Table 2 that the VOT values of / p' t' k' / of Mandarin are around 100ms and English ones are around 70ms. The aspiration of Mandarin is stronger. Mandarin unaspirated voiceless stops /p t k/ are similar to Italian ones, with the VOT value around 30ms. Both English and Italian voiced stops are mainly around - 90ms. Italian voiced stops all have pre-voicing while English ones may not have pre-voicing, instead with short-lags.

In addition, the stops in Italian are divided into singleton stops and geminate stops. Geminate stops are longer than singleton stops and vocal muscles are tenser when pronouncing (Laver, 1994). This phonetic difference can make the meaning of words different. For example, "copia / kopia /" means "copy" in Italian, while "coppia / kop:ia /" means "couple". The closure duration is the distinctive feature of singleton stops and geminate stops instead of VOT (Pickett et al. 1999), so they will not be discussed in this paper.

2.5.2 Data And Analysis

According to Table 3 below, the English VOT data shows: Learners' English /p' t' k'/ are between 73.4~96.3ms, less aspirated than Mandarin and close to English native

2.4 Hypothesis

Based on the findings of Flege (1995, 2018) which demonstrated that low-proficiency learners can hardly distinguish the similar phonemes across languages and tend to group them into the similar category in their interlanguage, it can be hypothesized that both Chinese (L1) and English (L2) will have an impact on Italian (L3) production. Furthermore, it can be expected that Chinese learners will categorize / b d g / in Italian into Chinese / p t k / or English / b d g / and / p t k / in Italian as Chinese / p t k/.

2.5 Result

2.5.1 The Stop Sound Systems of Mandarin, English and Italian

Chinese stops are divided by aspiration while English and Italian stops are divided by pre-voicing. But in some cases, American English voiced stops actually do not show pre-voicing. Details are shown in Table 2. The data of Mandarin and English VOT is from Ran & Shi（2007）and Lisker & Abramson (1964). The data of Italian is from two native speakers in this study as shown in Table 1.

Table 1
The VOT Value of Italian Stops (ms)

| Plosives	/p/	/t/	/k/	/b/	/d/	/g/
M.	23.2	31.2	41.8	-82.4	-84.8	-100.9
Std.	14.6	11.5	13.8	28.4	28.7	24.6

speakers. The pre-voicing of English /b d g/ are between -54.5 and -48ms, which is much lower than the standard norm and the frequency is low as well. As for the /b d g/ without pre-voicing, the VOT is slightly higher than 5ms and close to Mandarin /p t k/.

The Italian VOT data shows: Learners' Italian /p t k/ are mostly in the range of 20-50ms, but there are a small number of voiceless plosives with pre-voicing. Most of the voiced stops /b d g/ were between 18.4 and 31ms, and a few of them are with pre-voicing.
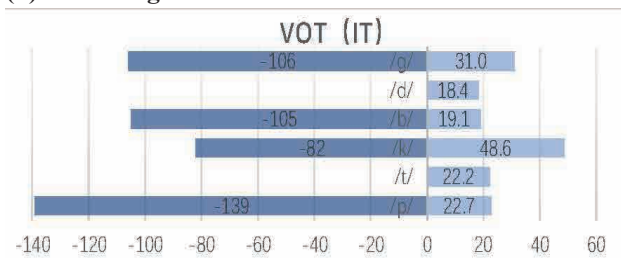
Table 3
The English and Italian VOT Value

EN	/p'/	/t'/	/k'/	/b/	/d/	/g/
M.	73.4	84.7	96.3	17.1	17.0/-54.5	32.1/-48
Std.	23.5	25.0	23.0	9.2	14.5	15.4
N.	59	58	58	59	56/2	59/1
IT (singleton)	/p/	/t/	/k/	/b/	/d/	/g/
M.	23.2/-139	22.45	49.0/-82	19.1/-105	18.4	31.0/-106
Std.	25.1	13.2	31.7	17.5	6.8	19.8
N.	59/1	60	59/1	59/1	60	58/1

(a) is the diagram form of Table 3
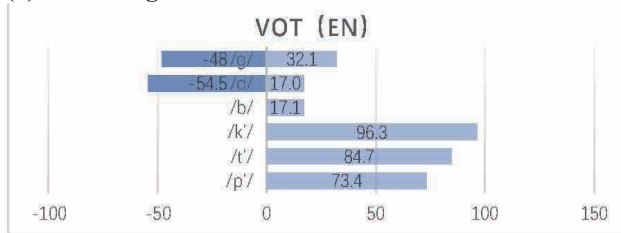
(b) is the diagram form of Table 3

This evidence shows that Chinese trilingual learners can distinguish between aspirated and unaspirated stops, but it is difficult for them to distinguish between voiced stops and unaspirated voiceless stops. Learners are still influenced by the phonetic system of their mother tongue, which is distinguished by aspiration. However, voiced stops with pre-voicing have been observed in both English and Italian. Chinese trilingual learners have begun to establish voiced stops categories. The differences between English and Italian interlanguage /b d g/ are not significant (/b/ , $P = 1.000 > 0.05$; / d /, $P = 0.151 > 0.05$; / g / $P = 0.436 > 0.05$), indicating that the learners have transferred the voiced stops from the L2 to L3.

DISCUSSION

The experimental results agree with the hypothesis that Chinese learners categorized Italian /b d g/ into English /b d g/. However, due to the mother tongue's negative effect on L2, learners did not fully grasp the feature of English voiced stops and transferred them into L3.

SLM focusses on the learning of position-sensitive allophones and the more similar the harder to acquire. But the term "similarity" is not precise and should be elaborated. Ringbom & Jarvis (2009, p.106) pointed out the actual similarities can accordingly be analyzed linguistically and the assumed similarities are hypothetically established by learners based on the perceived similarities. The similarity in SLM specifically refers to the wrong hypothetical assumption of learners which leads to mispronunciation.

In our study, /p t k/ in Italian shares actual similarities with /p t k/ in Chinese, which are unaspirated voiceless stops with short-lag around 30ms. Learners do not form a new phonetic category but fall into the original phonetic category of the mother tongue, which is easy for learners to perceive and acquire.

This can be seen as a positive transfer from L1. On the other hand, /b d g/ in Italian does not share the actual similarity with /b d g/ in English, because non-pre-voicing voiced stops do not exist in Italian. But learners still do not form a new phonetic category and falls into the category of L2 phonetic categories. This phenomenon can be explained by SLM. However, due to the fact that learners did not fully grasp the characteristics of voiced stops in the process of SLA, the inadequate pre-voicing appears in L3 as well.

CONLUSION

In this paper, VOT of stops in English and Italian of Chinese trilingual speakers were measured and analyzed. The results showed that learners could not distinguish /b d g/ and /p t k/ in Italian due to the wrong perceived similarities between L1 and L3 or wrongly assumed similarities between L2 and L3. But they acquired /p t k/ in Italian well as it shared actual similarities with Chinese. The SLM model can explain the learning of position-sensitive allophones in TLA, but cannot explain the learning of actual similar phones. To expend the application scope of the model, the term "similarity" should be distinguished more finely as actual similarities in phonetics, the perceived similarities of learners and assumed similarities of learners' minds. The difficulty level of learning phonetic features should be discussed in further researches.

REFERENCES

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*, *111*(45), 8-20.

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, *7*(1), 71-87.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207-229.

Dai, Z. Z. (2012). VOT Values of plosives Produced by Advanced Chinese English Learners. *Journal of XiDian University*, (6), 102-106. (In Chinese)

De Angelis, G. (2007). *Third or additional language acquisition* (Vol. 24). Multilingual Matters.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E. (2018). It's input that matters most, not age. *Bilingualism: Language and Cognition*, *21*(5).

Gao, Y. J., & Zhang, M. M.(2018). An Experimental Study on Chinese Students' Acquisition of English Plosives. *Journal of Northeast Normal University* (6), 13. (In Chinese)

Gut, U. (2010). Cross-linguistic influence in L3 phonological acquisition. *International Journal of Multilingualism*, *7*(1), 19-38.

Hammarberg, B. (1998*). The learner's word acquisition attempts in conversation*.

Hammarberg, B., & Hammarberg, B. (2009). Re-setting the basis of articulation in the acquisition of new languages: A third language case study. *Processes in Third Language Acquisition*.

Herdina, P., & Jessner, U. (2000). The dynamics of third language acquisition. *English in Europe: The acquisition of a third language*, *19*, 84.

Jakobson, R. (1958). 1971. Morphological inquiry for Slavic declension: structure of Russian case forms. Text in Russian, summary in English in 1971 version. *Jakobson, R.: Selected writings II: Words and Language*, *154*, 83.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384-422.

Llama, R., Cardoso, W., & Collins, L. (2010). The influence of language distance and language status on the acquisition of L3 phonology. *International Journal of Multilingualism*, *7*(1), 39-57.

Pierrehumbert, J., Beckman, M. E., & Ladd, D. R. (2000). Conceptual foundations of phonology as a laboratory science. *Phonological Knowledge: Conceptual and Empirical Issues* (pp.273-304).

Pyun, K-S. (2000). A model of interlanguage analysis –the case of Swedish by Korean speakers. *Journal of the Linguistic Society of Korea*, *26*(June), 199-222.

Ran, Q. B., & Shi, F. (2007). The analysis of VOT Values of monosyllable in Mandarin. *NanKai Linguistics,* (2), 21-31. (In Chinese)

Ringbom, H. (1987). *The role of the first language in foreign language learning* (Vol.34). Clevedon: Multilingual Matters.

Ringbom, H., & Jarvis, S. (2009). The importance of cross-linguistic similarity in foreign language learning. *The Handbook of Language Teaching*, 106-118.

Sypiańska, J., & Olender, A. (2013). Phonetic awareness in Polish learners of English. *Poznań Working Papers in Linguistics*, *1*(1), 1-9.

Tremblay, M. C. (2006). Cross-linguistic influence in third language acquisition: The role of L2 proficiency and L2 exposure. *Clo/opl*, *34*(1), 109-119.

Wrembel, M. (2014). VOT patterns in the acquisition of third language phonology. *Concordia Working Papers in Applied Linguistics, 5,* 750-770.

Wunder, E. M. (2010). Phonological cross-linguistic influence in third or additional language acquisition. In *New sounds 2010: Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech* (pp. 566-571).

Zhang, J. Y. (2012). The VOT analysis of French plosives of Chinese students. *Quarterly Journal of Chinese Studies*, (2), 174-181. (In Chinese)

APPENDIX

English Words List

/p'/	/t'/	/k'/	/b/	/d/	/g/
Paper	Take	Cake	Ball	Duck	Gold
Poor	Tool	Cabin	Boat	Die	Gate
Pose	Tomato	Cool	Bake	Dad	Goose
peach	Towel	Car	Butter	Did	Gank
Person	Tibet	Cure	Burst	Door	God
Pub	Time	Coat	Beyond	Desk	Guy
Pot	Tutor	Key	Bee	Disney	Girl
Pick	Toast	Couple	Beef	Dust	Go
Pattern	Touch	Cut	Bucket	Day	Gain
Power	Torture	King	Boot	Double	Gun

Italian Words List

/p/	/t/	/k/	/b/	/d/	/g/
Padova	Tavolo	Calcio	Bagno	Donna	Gatto
Padre	Tipo	Cura	Bevi	Dio	Gusto
Punto	Tutti	Come	Buono	Destro	Gamba
Pena	Tesi	Camera	Bevo	Dire	Galleria
Posso	Titolo	Capo	Bosco	Dopo	Guanto
Pago	Tardi	Cosa	Bacio	Danno	Gomme
Piedi	Tipico	Corsa	Bene	Dito	Gamba
Puro	Turno	Cane	Busta	Duro	Gatti
Pepe	Tono	Cuore	Buoi	Dentro	Gomma
Poco	Tomba	Comodo	Borsa	Dodici	Guardia

102