The Effect of Lexical Tones on Voice Onset Time in L2 Mandarin Production by English Speakers

Chiu-Ching Tseng[1]

[1]George Mason University, U.S.A.
[1]*ctseng2@gmu.edu*

Abstract

This study investigates the effect of lexical tones on VOT length for word-initial aspirated stops (i.e., /p^{h}a/, /t^{h}a/, and /k^{h}a/) in L2 Mandarin production.

Fifteen native English speakers who had studied Mandarin at George Mason University (GMU) and eight native Mandarin speakers were recruited. Results show that VOT values were significantly affected by tones in both groups. The results also show that these L2 learners used non-L1 VOT (i.e., 58~80 ms [1] vs. 88~93ms) for L2 production. Although their VOTs were longer than their L1 VOTs, they were shorter than the native Mandarin VOTs (90.78ms vs. 107.70ms). This may imply the process of L2 acquisition and that these learners are approaching native Mandarin like VOT.

Keywords: VOT, lexical tones, tonal effect, Mandarin and English, L1 transfer.

1. Introduction

Mandarin is a tonal language, which can present a great difficulty for its learners of non-tonal speakers, such as native English speakers. Although English is an intonation language, which employs prosodic variations at the sentential level, Mandarin uses tonal variations at lexical level [2].

Studies of tonal language phonologies have reported the significant effect of lexical tones on VOT values of aspirated stops in L1 [3][4][5]. However, a question remains as to whether or not L2 Mandarin learners also exhibit such an influence. Therefore, this study investigates the effect of lexical tones on VOT length for word-initial aspirated stops in L2 Mandarin produced by native English speakers. The paper is organized as follows: the next section review some of the necessary background information of VOT and lexical tone. It will then posit questions and state the hypotheses in the third section. Section 4 explains the method of the present study. Section 5 is dedicated to the results and the discussion. Finally, the conclusion includes possible future study plans and implications.

2. VOT and lexical tones

VOT is the interval between the release of a stop and the start of voicing for the following segment [6]. This short interval serves as a significant perceptual cue to distinguish voicing contrast and aspiration in Mandarin [3]. The delay of the vocal cord vibration is also referred to as "voice-lag" [6], where cross-linguistically, the amount of lag may be a short-lag; i.e., fewer than 30 milliseconds (ms) or a long-lag; i.e., greater than 30ms. The vocal cord configuration for the relative stop may be 1) fully voiced, 2) partly voiced, 3) voiceless unaspirated, 4) weak or no aspiration, or 5) strongly aspirated.

The stop consonants can be either in the voiced or voiceless configurations. Such a statement may be oversimplified because what accounts for being voiced in one language may be perceived as voiceless in another language [7]. For example, Yavas [6] reveals that while the voiced stops (e.g., /b, d, g/) in Romance languages are fully voiced, they may be subject to a certain amount of voicing loss (i.e., partially devoiced) in some Germanic languages (e.g., in English). Thus, some English voiced stops in some instances may be perceived by Romance language speakers as voiceless stops. For instance, Teschner and Whitley [8] state that "To speakers of Romance languages, an English speaker's *'bar'* [when the /b/ in *'bar'* is partially devoiced] resembles their *'par'*, whereas the aspiration in *'par'* may be ignored as wasted air rather than a reinforcement of voicelessness" (p.187). On the other hand, Mandarin stops are phonetically voiceless across the board [9]. Thus, the importance of VOT is evident, as it serves as an acoustic cue for voicing contrast in Mandarin.

"Lexical tone ... is the systematic modulations of pitch ... [carried by the syllable of a word]" [10]. A tonal language such as Mandarin uses distinctive pitch to contrast individual lexical units [11]. For example, there are four basic tones in Mandarin[1]; the syllable */ma/* can change for the different semantic representation it bears depending on which tone it carries[2].

10.21437/ISAPh.2018-22

2.1. English stops versus Mandarin stops

One of the main differences between English stops and Mandarin stops is the number of stops in each series. Although English maintains a two-way distinction between voiced and voiceless contrast, it also categorizes another dimension that predictably differentiate /p, t, k/ into aspirated and unaspirated [1][2][12][13]. The feature of aspiration is the allophonic variant for the voiceless set of stops, where /p, t, k/ are pronounced with aspiration at the beginning of stressed syllables [6]. On the contrary, Mandarin only contrasts between voiceless unaspirated and aspirated stops. In other words, Mandarin does not have a set of voiced stops [9][12][14]. Furthermore, VOT length may deviate between languages. For example, Lisker and Abramson [13] provide a set of mean VOT values of the aspirated stops for American English: 58ms, 70ms, and 80ms for /p^h, t^h, k^h/, respectively. Rochet and Yanmei [15] provide a different set of mean VOT values for the Mandarin aspirated stops (i.e., 99.6ms for /p^h/, 98.7ms for /t^h/, and 110.3ms for /k^h/). Thus, it is generally agreed that the Mandarin VOT length is significantly longer than that of English. From the given data, English exhibits a gradient increase of the values as the place of articulation moves from the front of the mouth to the back of the mouth. However, Mandarin has a different pattern; some sort of dipping at the alveolar stop. What may have caused this pattern of discrepancy is still unanswered in the literature. Additionally, Cho and Ladefoged [1] suggest that universally, the categorical set of VOT values can be generalized as: a) around 30ms for unaspirated stops, b) around 50ms for slightly aspirated stops, c) about 90ms for aspirated stops, and d) over 90ms for highly aspirated stops (p.223). Thus, Mandarin stops fall into the category of "highly aspirated".

2.2. Why different VOT values?

Many linguists have looked at voicing contrast in stops in many languages with respect to place of articulation [1][13], vowel context [12], and lexical tone [2][3].

2.2.1. Place of articulation effect

Cho and Ladefoged [1] propose that universally the further back the closure, the longer VOT. English follows this tendency [13], but Mandarin does not because Mandarin /p/ has slightly longer VOT than Mandarin /t/ [15]. Cho and Ladefoged [1] reveal that the stop closure duration is different between places, which may be due to different degrees of air pressure in the cavity that is behind the constriction; the smaller the cavity behind the constriction, the more rapidly intraoral air pressure builds up to reach the equilibrium point with the sub-glottal air pressure (p.212). In other words, if the air pressure between the cavity and the sub-glottal area is unbalanced, the closure interval is breached. Thus, the quicker the pressure builds up, the shorter the time the closure can hold; therefore the aspiration time is longer. According to them, this is an inverse relationship between the closure duration and the VOT length.

2.2.2. Vowel context effect

Chen et al. [12] report that the VOT values are longer when the following vowels are high and tensed. Rochet and Yanmei [15] also report that "the nature of the vowel had a significant effect on the VOT values of the preceding consonants" (p.105). Additionally, Chang et al. [16] suggest that the tense vowels provide greater resistance to the air from escaping the mouth. In other words, the muscle tension allows the cavity to offer greater force against trans-glottal airflow; thereby delaying the vibration, and thus, longer aspiration.

2.2.3. Lexical tone effect

VOT is also found to be affected by lexical tone [3]. Liu et al.[3] offer a reason that the rising component of a tone can cause an increase of tension in the voicing source, which may delay the onset of vibration. Their study results (shown in Figure 1) reveal that Tone 2 and Tone 3 have higher VOT values than that of Tone 1 and Tone 4.

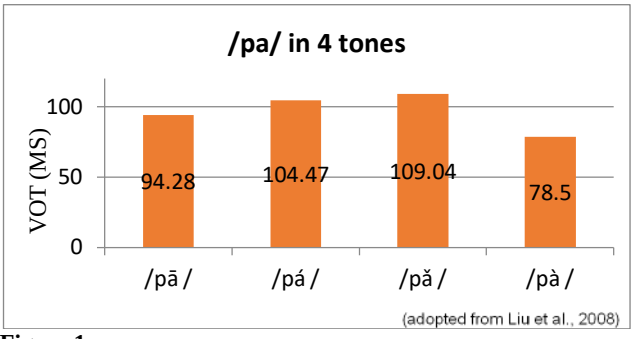

Figure 1.

The present study will take the position of Liu et al.'s [3] and Chang et al.'s [16] viewpoints, and assume that lexical tone affecting VOT values is a universal tendency, where the higher tone is correlated with more tense muscular movement. The rising part of producing a higher contour tone requires the raising of tongue root, which results in a smaller cavity chamber. The tensed up chamber muscle subsequently allows a larger holding of the air pressure; therefore, the higher pressure can build up [1]. As a result, the higher pressure leads to stronger aspiration.

3. Question, problem and hypotheses

As we can see, VOT has been widely studied. We have observed that firstly, the place of articulation alters VOT values [13], and it is attested in many languages worldwide. Secondly, vowel context has been shown to be influential as well [12]. However, these studies did not consider lexical tone effects. Many others (e.g., [2][10]) studied lexical tones of Mandarin, but they did not consider VOT in their research scope. Although studies have looked at the tonal effect on VOT, most of them focus on L1 [3][4]. The tonal effects on VOT may be conclusive in L1; however, there is no research found for the tonal effects on VOT in L2. Thus, it is still unknown whether or not the effect of lexical tones on VOT exists in L2 Mandarin produced by native English speakers. Another issue is that the literature has revealed that English and Mandarin exhibit significantly different VOT values. Therefore, it also seems valuable to inquire whether or not there is any correlation between L2 learners' L1 (English) VOT values and their L2 (Mandarin) VOT values, which may provide us some insight regarding their inter-language and the difference between theirs and the native Mandarin speakers' VOT values.

Additionally, since L2 learners' grammar has been found to be deeply influenced by their native language [17][18], we may predict English speakers to transfer their L1 VOT values in producing Mandarin. Secondly, if we assume tonal effects to be universal, we would expect the tone effects show not only in L1 but in L2 as well. Therefore, this study will make the first attempt to examine the tonal effects on L2 VOT values and hypothesize that:

A. Lexical tones affect VOT length in L2 Mandarin production.
B. L2 learners will substitute their native English VOTs for L2 production due to L1 transfer, leading to shorter VOT values than those of native Mandarin speakers.

4. Method

4.1. Participants

In order to test the hypotheses, an experimental study was designed. The production task elicited L2 learners' utterances for examination. Two groups of participants were recruited.

a. Native English learners of Mandarin: The experimental group consisted of 15 native speakers (6 males and 9 females) of English who have studied Mandarin for at least two semesters at GMU. The reason for controlling the length of learning was to ensure that these L2 learners are able to produce Mandarin tones correctly. All participants considered themselves native speakers of English. The average age of the participants was 27.8 years. The average age of onset of learning Mandarin was 20.3 years old; they had studied Mandarin for an average of 5.93 years prior to the experiment.

b. Native Mandarin speakers: The control group consisted of eight native Mandarin speakers (4 males and 4 females) who were born and resided in China for at least 15 years since birth. The reason for controlling the length of residency in China is to ensure that they had passed the Critical Period [18][19][20] before they came to the U.S. Their average age of arrival in the U.S. was 22 years. Their length of residency in the U.S. was 1.3 years.

All participants reported that they have no known history of speech or hearing impairment. The participation in the study was voluntary; although compensation in the form of a $5 coffee shop gift card was given to each participant.

4.2. Stimuli

This study used monosyllabic words controlled for three places of aspirated voiceless stops, / p^h, t^h, k^h /, one low-back-unrounded vowel, /a/, and four lexical tones, i.e., high-level (HL), mid-rising (MR), falling-rising (FR), and high-falling (HF). The pre-recorded tokens were spoken by a male native Mandarin speaker from Jiangsu province of China, whose native dialect is northern standard Mandarin[3]. The stimuli were normalized using Praat and then pre-programmed into the PsychoPy program [21].

Each stimulus was produced twice; once in isolation and once in a carrier phrase (i.e., **pā**, wǒ xiànzài shuō **pā**; "***pā**, I now say **pā***"). Each participant produced 24 tokens for the analysis. The stimuli were presented to the participants in a random order on a computer running the PsychoPy program version 1.83.04 [21]. This study chose to display *pinyin* system to the participants to avoid any possible inability or inaccuracy of reading Chinese characters because it is possible that at the current level of proficiency, these Mandarin learners might not have fully acquired the capacity of reading Chinese characters.

All recordings were done in the acoustic lab in the linguistics department of GMU. The speech samples were recorded using the equipments in the acoustic lab; a Mac Air 11' laptop running Audacity 2.1.3 with an external high-quality microphone (Apogee Mic Digital - H8309ZM/D) recorded in monotone at the CD quality rate of 44.1 kHz.

The participants heard the pre-recorded phrases and read the stimulus phrases one by one on the screen. Each phrase was played and displayed only one time. The participants then spoke each phrase into the microphone that was about one foot in front

of him/her. The procedures for the native Mandarin speakers were the same as for the Mandarin learners.

4.3. Data elicitation and analysis

A total of 552 tokens were recorded and measured. The measurements were obtained from the digitized signal by using Praat [22]. VOT values were measured based on a waveform and wideband spectrogram. The Excel data sheet was then imported into IBM SPSS 24 for data analysis.

The English VOT values were not elicited from the L2 participants because the present study adopted the values from what literature has found (i.e., approximately 58ms for /p^h/, 70ms for /t^h/, and 80ms for, /k^h/). The VOT values suggested for English by Lisker and Abramson [13] have been well-tested and attested by many other scholars [1][2][12][14]. Therefore, this study assumes their findings and uses them for comparison where necessary.

5. Results and discussion

VOT values from speech samples of Mandarin learners and native Mandarin speakers were measured and analyzed. Repeated-Measures (RM) ANOVA was used to analyze the data. The results indicated that the VOT values were significantly affected by the tone in both L2 learners *(F (3, 26) =16.069, p < .000)* and native Mandarin speakers *(F (3, 12) =12.457), p = .001)*. The place of articulation on the VOT values had marginal effect, but not statistically significant, for the L2 group *(F (2, 27) =2.605, p = .092)* and for the native Mandarin group *(F (2, 13) =1.437, p = .273)* as well. In terms of the environment, the results showed no significant effect from either isolated or embedded condition (*p=.468* for L2 group and *p=.874* for the native Mandarin group).

5.1. L2 group

There were 15 subjects in L2 groups. Figure 2 shows the mean VOT values for this group across three places and four tones. The results indicate that Tone 3 had the longest VOT values in tones, and /k^h/ was the longest with respect to the place of articulation.

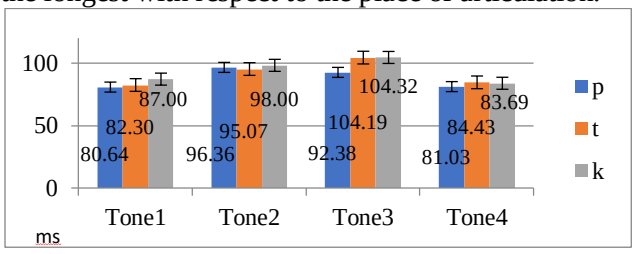

Figure 2. Mean VOT across 3 places & 4 tones for L2 group

The results show that averagely across three places, T3 was the longest (*M*=100.29ms), T2 the second-longest (*M*=96.48ms), and T1 and T4 were almost the same (*M*=83.31ms and 83.05ms). RM ANOVA revealed the statistically significant main effects found for the tone, but not statistically significant for the place. This may indicate that the tone had more significant effect than the place on the VOT value. Overall, VOT value of T3 was significantly longer than T1, and T4, but not as much to T2. Similarly, T2 was significantly longer than T1, and T4, but not to T3, and there was no difference between T1 and T4.

In terms of place of articulation, the means of /p^h/, /t^h/, /k^h/ were 87.06ms, 91.50ms, and 93.25ms, respectively. Although there were observable differences between places, RM ANOVA revealed that the place didn't have significant effect on VOT values with the current data. The *p*-value at 0.092 may indicate that there was only marginal main effect with respect to the place; it is possible to observe a statistically significant effect from the place if the number of sample size can be increased.

5.2. Native Mandarin group

There were eight subjects in the control group. Figure 3 shows the mean VOT values for the control group across four tones and three places. Similar to that of the L2 group, /k^h/ was longer than /p^h/ and /t^h/ with respect to the place of articulation, and T3 was longer than T1, T4 and slightly to T2.

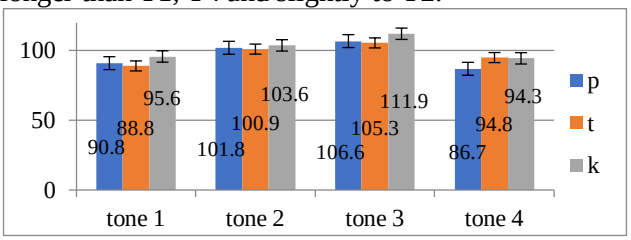

Figure 3. Mean VOT across 3 places & 4 tones for native Mandarins

RM ANOVA revealed that, in this group, the tone was also the only significant factor *(F (3, 12) =12.457), p = .001)*. There was no significant effect from the place for this group neither *(F (2, 13) =1.437, p= .273)*.

As far as within tone factor is concerned, the mean tone values were 91.739ms, 102.102ms, 107.970ms, and 91.947ms for T1, T2, T3, and, T4, respectively. The post-hoc pairwise-comparison was similar to the analyses of the L2 group that T3 was significantly longer than T1 and T4, but not to T2; and T2 was significantly longer than T1 and T4, but not to T3.

In terms of place of articulation, the means VOT values of /p^h/, /t^h/, /k^h/ were 96.51ms, 97.47ms, and 101.343ms, respectively. Although there were observable differences between places, the data analysis revealed that the place was not a significant

factor here *(F (2, 6) = .760, p = .508)*. In parallel to the place, the environment was also non-significant to the VOT values *(F (1, 7) = .027, p = .874)*.

5.3. Difference between groups

VOT values from speech samples produced by the native Mandarin group and the L2 group were analyzed to determine whether or not VOT values were significantly different between the two groups. Figure 4 shows the descriptive statistic values across four tones and three places between two groups.

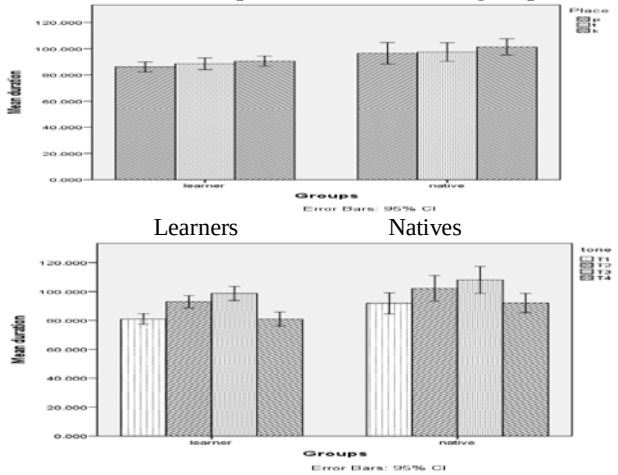

Figure 4. Comparison between groups for tones and places

There were observable different VOT values between two groups in tone and place; however RM ANOVA revealed that the difference was not at the significant level (*p = .154*). The tests of between-subject effects were not significant therefore we report that there was no significant difference found for the two groups. In other words, two groups produced similar values of VOT.

6. General discussions and conclusion

This study examined the effect of lexical tones on VOT in L2 Mandarin produced by English speakers.

We predicted that lexical tone does not only affect VOT values in native Mandarin production [3], but also affects on VOT values in L2 Mandarin production. In other words, that tone affects VOT values might be a universal tendency.

The present study showed these English speakers extended their native English VOT values while producing Mandarin. Generally speaking, English native speakers would produce VOT values at around 58ms to 80ms when speaking English. This study found their VOT were 87.06ms, 91.50ms, and 93.25ms for /p^h/ /t^h/ /k^h/, respectively, when producing Mandarin. Although the VOT values produced by L2 learners may not be as close to that of Mandarin group, the values certainly did not fall within normal English VOT ranges suggested in the literature.

Liu et al. [3] suggest that the rising part of the T2 and T3 may be the reason for longer VOT values. This was shown in this study as well: T2 and T3 were indeed longer than T1 and T4 for both L2 learners and the native Mandarin groups. Therefore, this tendency might be universal as the lengthening of VOT values (caused by tonal change) was shown in the Mandarin natives and the L2 learner groups of this study.

Secondly, RM ANOVA revealed that the VOT differences between two groups were not significant. This may suggest that the characteristic of "highly aspirated" VOT [1] in Mandarin is not only a special feature to the natives. It is that when one raises his/her tone of voice (e.g., advancing tongue root) when producing a voiceless stop, the vocal cord vibration may be delayed.

An additional concern about vowel duration affecting VOT value was not found in this study (i.e., where T4 has the longest VOT ratio at approximately 30:70% vs. T1, T2 & T3 all at approximately 25:75%, but T4 has the significant shorter syllable duration (281.56ms for the L2 group & 325.88ms for the native groups) than other three tones (431.77ms, 465.18ms & 515.16ms for the L2 group and 420.02ms, 430.15ms, & 496.18ms for the native group, respectively). If it would have been the case, then T3 should have the longest VOT ratio because of the longest syllable duration.

Many phonological studies examined the VOT and lexical tone separately. Both characteristics are seemingly mechanical. Thus, one would be misled by the fact that such mechanical processes should not show any evidence of variation for people with normal speech production ability. Apparently, this is not the case since many studies have shown that the L1 phonological grammar operates differently in various languages [1][2][5][13]. The results of this study provide evidence that the isolated acoustic properties, such as VOT, were involved with complex phonological categories, such as lexical tone. Mandarin learners' native VOT was not shown in their L2 production. In sum, hypothesis one was supported by the empirical data and hypothesis two did not hold, and further investigation is suggested.

7. Implications and future study

In many instances, L1 may involve in L2 speech productions for adult learners; however, the results of this study suggest that L1 transfer is not an isolated factor. Perhaps a universal tendency is also involved [23]. The present study was designed to answer whether or not there are tonal effects on

VOT values by L2 learners of Mandarin. It is possible to extend this study to various non-tonal languages that exhibit short native VOT values, such Spanish or Arabic. For example, Abramson & Lisker [24] reports that Spanish VOT falls under the range of 14ms to 24ms for /p, t, k/, which is very different than that of Mandarin. The observed patterns of lengthened VOT for the L2 group here may prompt possible further investigation in other languages such as Japanese, French, Arabic or Spanish whose VOT values are generally short or unaspirated.

8. Acknowledgements

This work was supported by the Linguistics program, in the Department of English at GMU. I am extremely thankful for my principle investigator, Dr. Steven Weinberger, and Dr. Douglas Wulf for their advice, guidance and suggestions, and my wife, who has been my biggest fan since day one of the process. Finally, I sincerely appreciate all participants who kindly and willingly spared their valuable time for this study. All errors are my own.

9. Selected references

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

[2] Ding, H., Jokisch, O., Hoffmann, R. (2010, May). *Perception and production of Mandarin tones by German speakers*. In Proc. of 5th Conference on Speech Prosody, Chicago, USA.

[3] Liu, H., Ng, M. L., Wan, M., Wang, S., Zhang, Y. (2008). *The effect of tonal changes on voice onset time in Mandarin esophageal speech*. Journal of Voice, 22(2), 210-218.

[4] Pearce,M. (2005). *Kera tone and voicing*. University College London Working Papers in Linguistics, 17.

[5] Tse, H. (2005). *The Phonetics of VOT and Tone Interaction in Cantonese* (Doctoral dissertation, University of Chicago).

[6] Yavas, M. S. (2011). *Applied English phonology* (2nd ed). Oxford ; Malden, MA: Wiley-Blackwell.

[7] Ashby, M., Maidment, J. (2005). *Introducing phonetic science*. Cambridge University Press.

[8] Teschner, R. V., Whitley, M. S. (2004). Pronouncing English: a stress-based approach, with CD-rom. *Georgetown University Press*.

[9] Iwata, R., Hirose, H. (1976). *Ann. Bull. RILP No. 10, 47-60 (1976) FIBEROPTIC ACOUSTIC STUDIES OF MANDARIN STOPS AND AFFRICATES*. Ann. Bull. RILP No, 10, 47-60.

[10] Sun, S. H. (1998). *The development of a lexical tone phonology in American adult learners of standard Mandarin Chinese (No. 16)*. University of Hawaii Press.

[11] Massaro, D. W., Cohen, M. M., Tseng, C. Y. (1985). *The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese*. Journal of ChineseLinguistics, 267-289.

[12] Chen, L. M., Peng, J. F., Chao, K. Y. (2009, December). *The effect of lexical tones on voice onset time*. In Multimedia, 2009. ISM'09. 11th IEEE International Symposium.

[13] Lisker, L., & Abramson, A. S. (1964). *A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements*. WORD, 20(3), 384–422. https://doi.org/10.1080/00437956.1964.11659830

[14] Chao, K. Y., Khattab, G., Chen, L. M. (2006, January). *Comparison of VOT patterns in Mandarin Chinese and in English*. In Proceedings of the 4th Annual Hawaii International Conference on Arts and Humanities (Vol. 840, p. 859).

[15] Rochet, B. L., Yanmei, F. (1991). *Effect of consonant and vowel context on Mandarin Chinese VOT: production and perception*. Canadian Acoustics, 19(4), 105-106.

[16] Chang, S. S., Ohala, J. J., Hansson, G., James, B., Lewis, J., Liaw, L., Urban, M. Yu, A. & Van Bik, K. (1999). *Vowel-dependent VOT variation: An experimental study*. The Journal of the Acoustical Society of America, *105*(2), 1400-1400.

[17] Chen, S. (2003). *Acquisition of English onset clusters by Chinese learners in Taiwan*. Paper presented at the Linguistics and English Language Postgraduate Conference. University of Edinburgh. Retrieved from http://www.lel.ed.ac.uk/~pgc/archive/2003/proc03/Szu-wei_Chen03.pdf

[18] Gass, S. M. (2013). *Second language acquisition: An introductory course*. Routledge.

[19] Johnson, J. S., Newport, E. L. (1991). *Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language*. Cognition, 39(3), 215-258.

[20] Lenneberg, E. H. (1967). *The biological foundations of language*. Hospital Practice, 2(12), 59-67.

[21] Peirce, J. W. (2009). *Generating stimuli for neuroscience using PsychoPy*. Frontiers in neuroinformatics, 2, 10.

[22] Boersma, P., Weenink, D. (2013): *Praat: doing phonetics by computer [Computer program].Version 6.0.21*, retrieved 25 September 2016 from http://www.praat.org/ on (pp. 552-557). IEEE.

[23] Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Routledge.

[24] Abramson, A. S., Lisker, L. (1972). *Voice-timing perception in Spanish word-initial stops*. Haskins Laboratories Status Report on Speech Research, 29(30), 15-25.

[25] Yip, M. J. (1980). *The tonal phonology of Chinese* (Doctoral dissertation, Massachusetts Institute of Technology).

[1] In addition to the four tones, there is a fifth tone, called "neutral tone". However, this tone is phonetically predictable. Therefore, it is not discussed in this study. Interested readers may consult in Yip [25].

[2] The four tones are high-level (HL), mid-rising (ML), falling-rising (FR), and high-falling (HF); they mean "mother", "hemp", "horse", and "scold" respectively (Sun [10]; Liu et al.,[3]).

[3] This particular speaker's speech was analyzed. The average values of /p^h/, /t^h/, /k^h/ were 95.38ms, 90.13ms, and 104.30ms, respectively. The mean VOT values of tones were 95.14ms, 83.59ms, 123.36ms, and 84.21ms.