Zeszyty Naukowe Towarzystwa Doktorantów UJ
Nauki Humanistyczne, Numer 21 (2/2018), s. 77–87
e-ISSN 2082-9469 | p-ISSN 2299-1638
www.doktoranci.uj.edu.pl/zeszyty/nauki-humanistyczne
DOI: 10.26361/ZNTDH.09.2018.21.05

Katarzyna Knoll

Uniwersytet Śląski
Wydział Filologiczny
Instytut Języka Angielskiego
Zakład Orientalistyki i Językoznawstwa Ogólnego
e-mail: katarzyna.knoll@us.edu.pl

Phonetic Imitation of Mandarin Stops. The Case of Polish Advanced Learners

Abstract

Mandarin Chinese, unlike Polish, has no voiced stops and it contrasts voiceless unaspirated and voiceless aspirated stops, whereas Polish contrasts voiced and voiceless unaspirated stops. A study on the perception of voicing contrasts in Mandarin showed that Polish advanced speakers of Mandarin consistently divide stops into lenis (/p t k/) and fortis (/p^h t^h k^h/). This paper is to show whether they are able to articulate these sounds in a native-like manner. A group of university students participated in our pilot study. They were all advanced learners of Mandarin (their L3). They were asked to imitate syllables they heard. Their performances were recorded and analysed by measuring VOT values of the target sounds. The results of the experiment reveal that long lag stops are generally produced with sufficient, native-like, aspiration, whereas short lag stops appear to be much more problematic and Polish advanced users of Mandarin, even though exposed to the model, do not always produce them correctly, i.e. they add voicing and articulate these sounds in a Polish-like manner.

Keywords

VOT, language acquisition, Mandarin stops, voicing contrasts, imitation

Introduction

Voice Onset Time has been widely used as one of the most reliable acoustic features for investigating voicing contrasts in stops, and it has been applied in the studies of many languages, including Mandarin Chinese and Polish. A considerable number of such studies have shown that Polish divides up the VOT continuum with two categories: voicing lead vs. short lag, whereas Mandarin contrasts short lag vs. long lag.[1]

Being an academic teacher (and previously a student) of Mandarin, I observed that learners, even though they do have sophisticated knowledge in Chinese phonetics and are aware of the above-mentioned difference, seem to articulate Mandarin stops in a Polish-like manner. This observation lead me to the following question: Are Polish advanced speakers of Mandarin able to produce Mandarin stops in a native-like manner?

This study is an attempt to answer this question.

1. Voice Onset Time

Voice Onset Time (VOT) is "the time interval between the burst that marks release and the onset of periodicity that reflects laryngeal vibration."[2] In other words, it can be defined as the time between the burst of air and the initiation of a vowel.

In most languages, stops can be characterized as produced with (i) voicing lead, (ii) short voicing lag, or (iii) long voicing lag.

(i) Voicing lead (negative VOT): voicing starts before the release of the stop, (approximately -30 ms or more VOT).
(ii) Short voicing lag (zero onset): voicing begins at or just after the release of the plosive (approximately 0 to +30 ms or up to +35 ms.[3]
(iii) Long voicing lag (positive VOT): voicing begins well after the release of the stop (approximately +50 ms or more VOT). It is either accompanied by silence[4] or aspiration which is heard "if the vocal tract resonates to turbulent air passing through the open glottis."[5]

---

1 P. A. Keating, M. J. Mikoś, W. F. Ganong III, *A Cross-language Study of Range of Voice Onset Time in the Perception of Initial Stop Voicing*, "Journal of the Acoustical Society of America" 1981, No. 70; Z.-J. Wu, M.-C. Lin, 实验语音学概要 [*Shiyan Yuyinxue Gaiyao*], Pekin 1987.

2 L. Lisker, A. S. Abramson, *A Cross Language Study of Voicing in Initial Stops: Acoustic Measurements*, "Word" 1964, No. 20, p. 422.

3 P. A. Keating, *Phonetic and Phonological Representation of Stop Consonant Voicing*, "Language"1984, No. 60.

4 D. H. Klatt, *Voice-onset Time, Frication, and Aspiration in Word-initial Consonant Clusters*, "Journal of Speech and Hearing Research" 1975, No. 18.

5 L. Lisker, A. S. Abramson, op. cit., p. 416.

1.1. VOT for Polish and Mandarin stops

Mandarin and Polish exploit the VOT continuum differently. It is well known that Mandarin has no phonetically voiced stops, and it is aspiration that is the only distinctive phonetic feature differentiating voiceless unaspirated and voiceless aspirated stops.[6] Polish, on the other hand, "contrasts prevoiced stops with voiceless unaspirated or slightly aspirated stops, which corresponds to a contrast of voicing lead with short-lag VOT."[7]

In Table 1 measurements of Polish VOT means, as reported in Kopczyński[8] and Keating *et al.*,[9] are presented, whereas Table 2 presents measurements of Chinese VOT means, as reported by different scholars: Shi and Liao,[10] Wu *et al.*,[11] Rochet and Fei[12] and, more recently, by Ran,[13] Chao *et al.*,[14] and Chao and Chen.[15]

Tab. 1. Mean VOT values for Polish stops; in milliseconds

	pʰ	**tʰ**	**kʰ**	**p**	**t**	**k**	**b**	**d**	**g**
Kopczyński				+37.5	+33	+49	-78	-72	-61
Keating et al.				+21.5	+27.9	+52.7	-88.2	-89.9	-66.1

Source: own work.

[6] Z.-J. Wu, M.-C. Lin, op. cit.

[7] P. A. Keating, M. J. Mikoś, W. F. Ganong III, op. cit., p. 1261.

[8] A. Kopczyński, *Polish and American English Consonant Phonemes: A Contrastive Study.* Warszawa 1977.

[9] P. A. Keating, M. J. Mikoś, W. F. Ganong III, op. cit.

[10] F. Shi, R.-R. Liao, 中美学生汉语塞音实质对比分析 [*Zhongmei xuesheng hanyu seyin shizhi duibi fenxi*], "Language Teaching and Linguistic Studies" 1986, No. 4.

[11] Z.-J. Wu, M.-C. Lin, op. cit.

[12] B. L. Rochet, Y. Fei, *Effect of Consonant and Vowel Context on Mandarin Chinese VOT: Production and Perception*, "Canadian Acoustics" 1991, No. 19 (4).

[13] Q.-B. Ran, 基于普通话的汉语阻塞辅音实研究 [*Jiyu putonghua de hanyu zuse fuyin shiyanjiu*], Nankai 2005.

[14] K.-Y. Chao, G. Khattab, L.-M. Chen, *Comparison of VOT Patterns in Mandarin Chinese and in English*, "Proceedings of the 4th Annual Hawaii International Conference on Arts and Humanities" 2006.

[15] K.-Y. Chao, L.-M. Chen, *A Cross-linguistic Study of Voice Onset Time in Stop Consonant Productions*, "Computational Linguistics and Chinese Language Processing" 2008, No. 13 (2).

Tab. 2. Mean VOT values for Mandarin stops: DS = disyllables; in milliseconds

	p^h	t^h	k^h	p	t	k	b	d	g
Shi and Liao	+94	+100	+103	+7	+7	+18			
Wu et al.	+72	+100	+85	+7	+9	+19			
Rochet and Fei	+99.6	+98.7	+110.3						
Ran	+106	+104	+112	+12	+13	+30			
Chao et al. DS	+82	+81	+92	+14	+16	+27			
Chao and Chen DS	+77.8	+75.5	+85.7	+13.9	+15.3	+27.4			

Source: own work.

As can be seen, Polish voiceless stops are produced with moderate positive VOTs, which are slightly higher than mean VOT values for Mandarin voiceless unaspirated stops, but, at the same time, much lower than VOTs for Mandarin voiceless aspirated stops.

2. Experimental study

2.1. Objectives

Due to the fact that Polish and Mandarin exploit the VOT continuum in a completely different manner, we can expect that both perception and production of Mandarin stops may pose a considerable difficulty for Poles learning Chinese.

A study on the perception of Mandarin stops by native Polish speakers[16] showed that adult advanced learners of Mandarin Chinese have a strong categorisation effect, and consistently divide Mandarin stops into fortis (/p^h t^h k^h/) and lenis (/p t k/), which is not the case for Poles who do not speak Mandarin. However, to the best of our knowledge, there are no studies which investigate pronunciation of Mandarin stops by Poles.

Thus, the present study is an attempt to show how Polish advanced learners of Chinese articulate Mandarin stops, and to check whether they are able to produce them in a native-like manner.

[16] K. Knoll, op. cit.

2.2. Subjects

A total of 20 Polish Advanced Learners of Mandarin (as their L3) participated in the study: 4 males and 16 females. They were all fifth-year students of English Philology following the English-Chinese, translation programme of studies at the University of Silesia, Poland. Their skills in Mandarin had been repeatedly confirmed by annual practical examinations. They ranged in age from 22 to 25 years (Mean = 22.6, Std. Dev. = 0.75). They had 4 years of experience with learning Mandarin Chinese (1200 hours), and they all received substantial amounts of native--speaker input (1140 hours). At the initial stages of study they were taught by native Mandarin speakers exclusively. All subjects volunteered and were not paid for their participation. None of the participants reported any speech or hearing disorders. They were all naive to the object of the study.

2.3. Stimuli

We selected 6 Mandarin syllables whose onsets (声母 *shēng mǔ* 'initials') were /p p^h t t^h k k^h/. CV(V) and CVC patterns were used, i.e. each item had three sounds: (i) a stop consonant, (ii) a vowel (the low vowel /a/ or mid-low vowel /æ/), (iii) a consonant or another vowel (i.e. second part in a diphthong).

The stimuli used in the experiment were recordings of the selected lexical items. We used recordings from popular course books and textbooks, so as to make sure that speakers' pronunciation could be considered to be standard Mandarin pronunciation. The items were all produced by male speakers. Two stimuli were produced in high-level (HL) tone, one in mid-rising (MR), two in falling-rising (FR), and one in high-falling (HF) tone. Then, VOT values of word-initial stops in each stimuli were measured using Praat 5.4.06 speech-analysis software package (Boersma and Weenink 2015) by means of a spectrographic display and waveforms. VOT in each stimulus was measured between the first peak of the release burst to the onset of the second formant of the following vowel.[17]

The following items were given to the subjects:

(1) 排 *pái* 'a row': /p^haɪ/, /p^h/ +166 ms VOT
(2) 百 *bǎi* 'a hundred': /paɪ/, /p/ +10 ms VOT
(3) 毯 *tǎn* 'a blanket': /t^han/, /t^h/ +118 ms VOT
(4) 单 *d*ān 'a bill': /tan/, /t/ 0 ms VOT
(5) 看 *kàn* 'to look': /k^han/, /k^h/ +155 ms VOT
(6) 干 *gān* 'dry': /kan/, /k/ +20 ms VOT

[17] P. A. Keating, M. J. Mikoś, W. F. Ganong III, op. cit.

2.4. Experimental procedures

The experiment took place in a quiet room. Prior to the experiment, each subject was instructed in Polish about the methodology of the study.

As it is a general ability to produce particular sounds that is examined, we decided to ask participants to repeat syllables after the model instead of reading out loud words or phrases.

The stimuli were presented via high-quality headphones built in the headset at a comfortable listening level. Special care was taken to provide the same acoustics for all subjects. Each stimulus was presented once and each presentation was followed by a four-second pause.

After the experiment, each subject was asked if they found Mandarin stops difficult to imitate.

2.5. Measurements

We measured VOT values of word-initial stops using Praat 5.4.06 speech-analysis software package[18] by means of a spectrographic display and waveforms, and categorised them as:

(i) voiced (approximately -30 ms or more VOT);
(ii) voiceless unaspirated (approximately 0 to +35 ms VOT);
(iii) voiceless aspirated (approximately +50 to +80 ms VOT);
(iiii) voiceless highly aspirated (approximately +90 ms or more VOT).

Mean VOT values, standard deviations (SD), and graphical representation were made using EXCEL. ANOVA and t-Test were used for all statistical analyses, including the comparison of results and calculation of significance.

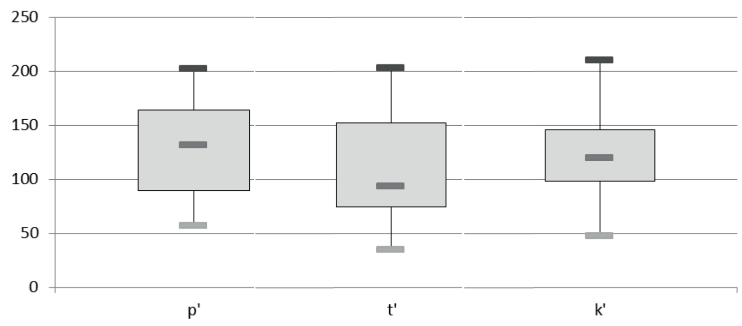
3. Results

Figure 1 shows the VOT distribution for all Mandarin aspirated stops. As can be seen, VOT ranges for /p^h t^h k^h/ were relatively wide for all three stops and centered around +87 to +159 ms, +73 to +139 ms, and +96 to +140 ms, respectively. The values of standard deviation presented in Table 3 also imply that /p^h/ (SD = 67.7 ms) allowed some more variation than /t^h/ (SD=45.4 ms) and /k^h/ (SD = 53.4 ms).

The ANOVA test showed that the differences between /p^h t^h k^h/ did not meet the criteria of statistical significance.

[18] P. Boersma, D. Weenink, "Praat: doing phonetics by computer", computer program: 5.4.08 version, [online] http://www.praat.org/ [accessed: 24.03.2015].

Fig. 1. Boxplot for aspirated stops, VOT in milliseconds



Source: own work.

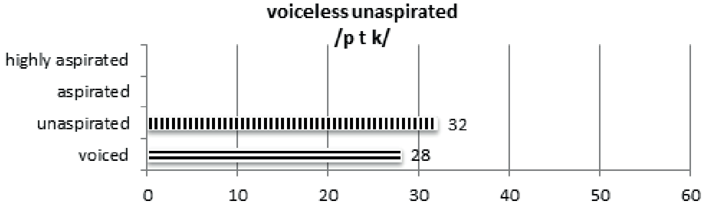
Tab. 3. Mean VOT values for highly aspirated stops (in milliseconds)

	p^h (+166 ms)	**t^h (+118 ms)**	**k^h (+155 ms)**
Mean	+127	+107	+124
Std. Dev.	67.7	45.4	53.4

Source: own work.

Productions of voiceless unaspirated stops were divided: they were produced correctly (i.e. with short voicing lag) 32 times (53%) and in a Polish-like manner (i.e. with voicing lead) 28 times (47%) (Figure 2).

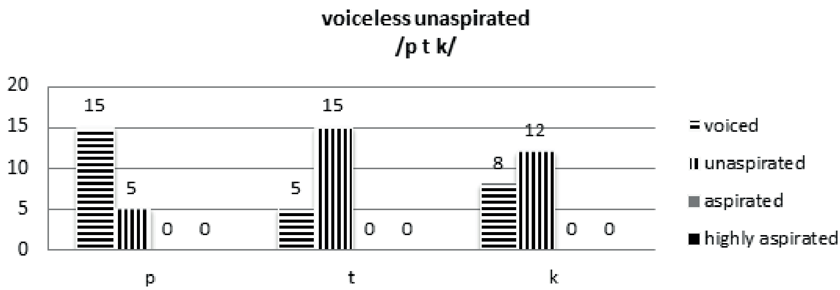
Fig. 2. Imitation of voiceless unaspirated stops – overall results



Source: own work.

As shown in Figure 3 and Table 4, +10 ms VOT /p/ was the most difficult to articulate, as 75% of the subjects produced it with voicing (mean VOT of -57 ms). Productions of +20 ms VOT /k/ were divided: 12 subjects (60%) articulated it correctly and 8 as voiced (mean VOT of -25 ms). 0 ms VOT /t/ appeared to be the easiest to imitate as 75% of the group articulated it with short voicing lag (mean VOT of -4 ms).

Fig. 3. Imitation of voiceless unaspirated stops



Source: own work.

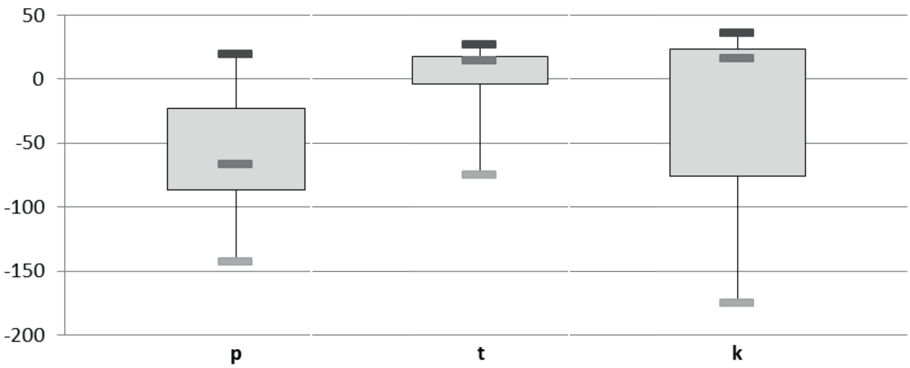
Tab. 4. Mean VOT values for unaspirated stops (in milliseconds)

	p (+10 ms)	t (0 ms)	k (+20 ms)
Mean	-57	-4	-25
Std. Dev.	45.4	36.3	66.6

Source: own work.

Figure 4 shows the VOT distribution for unaspirated stops. As can be seen, VOT ranges for /p t k/ centered around -55 to -15 ms, +10 to +18 ms, and -77 to +22 ms, respectively. It can also be seen that the range for /t/ was relatively small, whereas for the velar stop it was particularly wide. The values of standard deviation presented in Table 4 also imply that /k/ (SD = 66.6 ms) allowed much more variation than /t/ (SD = 36.3 ms) and /p/ (SD = 45.4 ms). The ANOVA test showed that the differences between /p t k/ were significant ($F(2,57) = 3.159$; $p=0.007$), and the t-Test showed that it was only the difference between /p/ and /t/ that reached statistical significance ($t(36) = 2.028$; $p<0.05$).

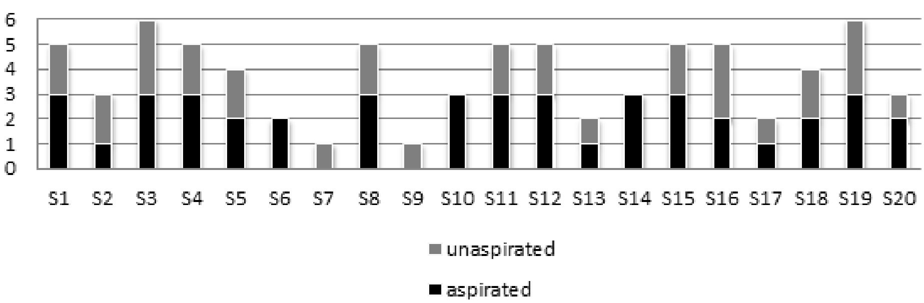
Fig. 4. Boxplot for unaspirated stops, VOT in milliseconds



Source: own work.

What is interesting, although all subjects claimed that Mandarin stops were easy to imitate, only a half of them managed to imitate them well during the experiment. As shown in Figure 5, one person (S2) had problems with adding aspiration, four subjects found imitating unaspirated sounds challenging (S6, S10, S14, S20), and for four subjects both aspirated and unaspirated sounds turned out to be very difficult (S7, S9, S13, S17).

Fig. 5. Number of stops imitated correctly by each subject



Source: own work.

Conclusions

Two conclusions may be derived from the present study: (i) Polish advanced users of Mandarin believe that pronunciation of Mandarin stops is not challenging, and (ii) pronunciation of Mandarin stops is challenging for Polish advanced users of Mandarin.

The results of the study indicate that what is particularly difficult for Poles is the production of voiceless unaspirated stops, and that there are many advanced learners whose pronunciation of Mandarin stops is incorrect (even though being exposed to the model).

As shown, Poles tend to produce Mandarin unaspirated stops in a Polish-like manner, i.e. they add voicing. The results of the study show that the only exception is 0 ms /t/, as it was generally produced correctly.

What probably makes Mandarin stops so difficult is the fact that they are similar (but not identical) to Polish stops, and Poles who start learning Chinese may tend to assimilate Mandarin categories into their pre-established L1 categories, which affects the further process of language acquisition.[19]

As mentioned, although a half of the group did poorly in the experiment, all learners believed that they did well. Thus, it seems that teachers should pay more attention to the proper pronunciation of stops, and, most importantly, they should increase students' awareness of the existence of between-language differences from the initial stages of studying Mandarin. In order to develop learners' good pronunciation habits in the course of learning Mandarin phonetics, it may be a good idea to use speech-analysis software during the classes and practise pairs of Polish-Mandarin syllables, e.g. Polish /gan/ vs. Mandarin /kan/, etc. Also, more theoretical instruction should be given to the learners.

Another, larger, study (including involving a larger group of students from different Polish universities and more stimuli) should be conducted in the future so as to verify present results.

[19] See also: C. Best, *A Direct Realistic Perspective on Cross-language Speech Perception*, [in:] *Speech Perception and Linguistic Experience: Issues in Crosslanguage Research*, Timonium 1995; J. E. Flege, *The Production of 'New' and 'Similar' Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification*, "Journal of Phonetics" 1987, No. 15; J. Tambor, *Nauczanie wymowy języka polskiego jako obcego. Warsztaty*, [in:] *Inne optyki: nowe programy, nowe metody, nowe technologie w nauczaniu kultury polskiej i języka polskiego jako obcego*, red. R. Cudak, J. Tambor, Katowice 2001.

Bibliography

1. Best C., *A Direct Realistic Perspective on Cross-language Speech Perception*, [in:] *Speech Perception and Linguistic Experience: Issues in Crosslanguage Research*, Timonium 1995.
2. Boersma P., Weenink D., "Praat: doing phonetics by computer", computer program: 5.4.08 version, [online] http://www.praat.org/ [accessed: 24.03.2015].
3. Chao K.-Y., Chen L.-M., *A Cross-linguistic Study of Voice Onset Time in Stop Consonant Productions*, "Computational Linguistics and Chinese Language Processing" 2008, No. 13 (2).
4. Chao K.-Y., Khattab G., Chen L.-M., *Comparison of VOT Patterns in Mandarin Chinese and in English*, "Proceedings of the 4th Annual Hawaii International Conference on Arts and Humanities" 2006.
5. Flege J. E., *The Production of 'New' and 'Similar' Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification*, "Journal of Phonetics" 1987, No. 15.
6. Keating P. A., Mikoś M. J., Ganong III W. F., *A Cross-language Study of Range of Voice Onset Time in the Perception of Initial Stop Voicing*, "Journal of the Acoustical Society of America" 1981, No. 70.
7. Keating P. A., *Phonetic and Phonological Representation of Stop Consonant Voicing*, "Language"1984, No. 60.
8. Klatt D. H., *Voice-onset Time, Frication, and Aspiration in Word-initial Consonant Clusters*, "Journal of Speech and Hearing Research" 1975, No. 18.
9. Knoll K., *The Perception of English, Mandarin and Polish Word-initial Stops by Polish Schoolchildren and Adults*, "Theoretical and Applied Linguistics" 2015, No. 1 (3).
10. Kopczyński A., *Polish and American English Consonant Phonemes: A Contrastive Study.* Warszawa 1977.
11. Lisker L., Abramson A. S., *A Cross Language Study of Voicing in Initial Stops: Acoustic Measurements*, "Word" 1964, No. 20.
12. Ran Q.-B., 基于普通话的汉语阻塞辅音实研究 [*Jiyu putonghua de hanyu zuse fuyin shiyanjiu*], Nankai 2005.
13. Rochet B. L., Fei Y., *Effect of Consonant and Vowel Context on Mandarin Chinese VOT: Production and Perception*, "Canadian Acoustics" 1991, No. 19 (4).
14. Shi F., Liao R.-R., 中美学生汉语塞音实质对比分析 [*Zhongmei xuesheng hanyu seyin shizhi duibi fenxi*], "Language Teaching and Linguistic Studies" 1986, No. 4.
15. Tambor J., *Nauczanie wymowy języka polskiego jako obcego. Warsztaty*, [in:] *Inne optyki: nowe programy, nowe metody, nowe technologie w nauczaniu kultury polskiej i języka polskiego jako obcego*, red. R. Cudak, J. Tambor, Katowice 2001.
16. Wu Z.-J., Lin M.-C., 实验语音学概要 [*Shiyan Yuyinxue Gaiyao*], Pekin 1987.