Oral Motor Exercises For CSL Learners to Master Productions of Retroflex And Non-Retroflex Consonants

Yixin Zhang and Jinsong Zhang*
*Beijing Language and Culture University, Beijing, China
E-mail: blcuzyx@163.com Tel: +86-18810018901
jinsong.zhang@blcu.edu.cn Tel: +86-13522706878

Abstract— **Retroflex and non-retroflex consonants in Mandarin (z[ts]、c[ts‘]、s[s]、zh[tʂ]、ch [tʂ‘]、sh[ʂ]) are difficult to produce for many Chinese as a second language (CSL) learners with different mother tongues. Aiming at developing an efficient way to help them to master the sounds, this study adopts the idea of Oral Motor Exercise (OME) to devise a speech production training method. The method consists of two steps: the first is non-speech OME, in which muscle groups of jaw, lips and tongue, essential to the production of target consonants, are intentionally exercised; the second is speech OME, in which monosyllables with target consonants are practiced. 30 participants took part in the experiment. The results showed that both non-speech and speech OMEs could effectively reduce subjects' consonant errors in a short time. When the two steps were accomplished, the trainee achieved the best performances. These indicated the effectiveness of the proposal.**

I. INTRODUCTION

Learning non-native speech sounds can be a challenging task. In particular, the production of the retroflex and non-retroflex consonants in Mandarin (z[ts], c[ts‘], s[s], zh[tʂ], ch [tʂ‘], sh[ʂ]) is exceptionally difficult for many foreign speakers who learn Mandarin Chinese as a second language(CSL) [1],[2]. For example, the literature has well documented that CSL learners who speak English, Japanese, Korean, Lao and other languages can hardly speak these consonants as well as natives, even if they have lived in the target language environment for a long time[3-13].

Guided by behaviorism, the process of speech production can be regarded as the acquisition of motor skills [14-16]. It mainly depends on the mastery of operational skills. Previous studies showed that the stages of motor learning are the cognitive phase, the associative phase, and the autonomous phase[17]. One requirement for accurate articulation is the correct and coordinated actions of related organs[18-20]. In the CSL teaching field, Liu pointed out the principles of teaching Chinese as a foreign language: to imitate and practice primarily, with informative knowledge as a supplement, and to focus on the combination of mechanical exercise and meaningful practise[21]. Therefore, it is crucial to help CSL learners to construct the correct movement modes of the muscles (groups), which involved in the articulation organ produced by the target language. Correct behavioral habits should be developed.

There are two assumptions in motor learning field. It is hypothesized that there is a common set of motor control principles and neural anatomical representation in the human nervous system for speech and non-speech activities that involve the same structures. Another hypothesis, which suggests that learning could be facilitated by breaking down complex movements into sub-components. Because this allows "the motor system to plan simpler movement patterns and gradually develop skilled control of more complex movement patterns"[22]. Guided by these hypotheses, Oral motor Exercises (OME), which are defined as an alternate approach for managing developmental speech sound disorder as a mechanical behavioral training method, are widely used in speech pathology[23-25]. Although still controversial, many clinical experience showed a positive effect on different types of dysarthria[26-29]. While there is still no research on similar training methods in L2 learning field. So it will be interesting to optimize the training method and to see if OME can master speech errors in CSL learners.

In order to better help CSL learners to correct the pronunciation of retroflex and non-retroflex consonants in Mandarin, an oral motor program were designed in this paper. The program is based on OME and crucial action modes of z[ts], c[ts‘], s[s], zh[tʂ], ch [tʂ‘], sh[ʂ]. In the training, two kinds of exercises are involved: non-speech OME (NS-OME) and speech OME (S-OME). NS-OME do not involve the practice of speech sound articulation. Instead, it focus on non-speech movements of the speech mechanism. S-OME are phonetic and phonemic treatments, in which monosyllables with target consonants are practiced. The feasibility and efficacy of the training method were also examined by an empirical research. The exploration may not only provide further evidence for the hypothesis mentioned above, but also provide a new perspective on the training methods of L2 learning.

II. METHOD

A. *Training program design*

1. NS-OME Program design

In this experiment, we designed the training scheme based on the NS-OME in speech therapy, and considered the physiological characteristics of retroflex and non-retroflex consonants. First of all, the production of retroflex and non-retroflex consonants is related to organs including the jaw, lips and tongue[19]. The movement of the tongue is crucial to the pronunciation of z[ts], c[ts‘], s[s], zh[tʂ], ch [tʂ‘], sh[ʂ]. Therefore, according to previous studies, the design of the training movement refers to two basic modes of tongue movement related to retroflex and non-retroflex consonants: the upward movement of both edges of the tongue (zh[tʂ], ch [tʂ‘], sh[ʂ]) and the slight upward movement of the blade z[ts], c[ts‘], s[s]. Second, considering the difference between dysarthria patients and L2 learners, we have screened and improved the original method. Most of the L2 learners are healthy adults. They have developed a mature sensory and perceptual system, as well as basic control over the modulating organs. Therefore, all passive oral motor therapy that required assistive devices were removed in this research. Finally, computer aided technology is applied to the training to avoid the instability of the experiment which may be caused by human intervention. The main objectives of non-speech OME are: 1. To strengthen the overall movement of the oral cavity, so as to improve the quality of speech production; 2. To enhance the flexibility and coordination of the tuning organs; 3. To break the original motion modes of the articulation organs and to establish new motion modes.

The training includes jaw, lip and tongue exercises:

(1) Jaw training, including the open jaw wide, jaw side to side and other exercises. Due to the page limitation, the training action screenshot are as follows:

Fig. 1 Screen shot examples of jaw training.

(2) Lip training, including the lip pucker, retracted smile, upper lip stretch, lower lip stretch, smile and pucker and so on.

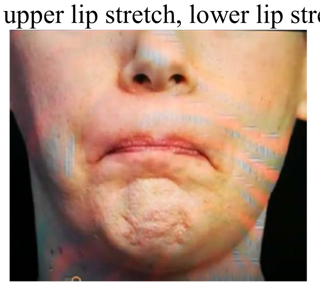

Fig. 2 Screen shot examples of lip training.

(3) Tongue training, including the tongue portrusion, tongue lift and lower, tongue sweep right and left, tongue sweep combo, tongue lateralization, tongue circles right and left and so on.

Fig. 3 Screen shot examples of tongue training.

The whole training process was recorded by speech therapy professionals with video. In the video, the key movements are explained to the subjects to ensure that they can understand and master the main points of movement. We asked all participants to do an imitation and provided them with feedback so that they could self-monitor and correct themselves at any time. All training moves are practiced three times in each part, with 1 minute rest between each part. Total training duration is about 15 minutes.

2. S-OME Program design

S-OME are phonetic and phonemic treatments, in which monosyllables with target consonants are practiced.It consists of two parts, including sounds movement training of retroflex and non-retroflex consonants and monosyllables training, which finals are “a、i[ɿ]/ i[ʅ]、u”.

(1) . Dynamic graphic diagrams of target consonants.

In order to ensure that the subjects can have a basic understanding of the movement of the key articulation organs of these six consonants, sagittal plane flash were produced according to the movement characteristics of the related organs. Based on Attention Mechanism [30], we made special color hints on the contact position of tongue and palate. Except for tongue and lips, all other organs are shown in gray scale. The purpose of this is to eliminate the possible interference of irrelevant information to the subject. This part requires participants to observe and imitate, and each flash is played 5 times. Take “z[ts]” and “zh[tʂ]” as an example, the flash screen-shots are shown as follows:

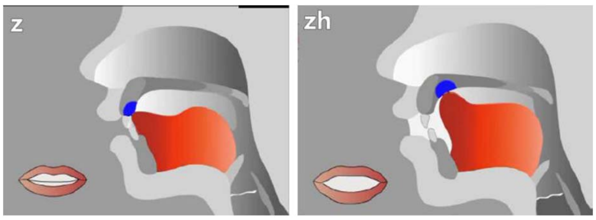

Fig. 4 Screen shot examples of sagittal plane flash.

(2) . Monosyllables training.

All monosyllables have target consonants as initial, and “a、i[ɿ]/ i[ʅ]、u” as final. The tone is high and level. During the training, these syllables are arranged and combined to form monosyllabic, disyllabic and trisyllabic sequences.Participants are asked to imitate carefully.

III. STIMULI

A. Perception stimuli

We designed a perceptual test for the pre-test and post-test of the experiment. The purpose of the perceptual experiment is to investigate the perception ability of target sounds by subjects. We used 54 monosyllabic stimuli, with high and level tone. All participants were asked to listen every sound and to make a choice of 6 or 1 on the computer keyboard. The accuracy and response time (RT) were recorded.

B. *Production stimuli*

There are three parts of production stimuli were involved in the experiment. It is composed of 118 monosyllabic and disyllabic natural syllables. The initial and final categories are as balanced as possible. The generalization test is three - and four-syllable words that meet the above conditions. There are 98 syllables in the generalization test The production of subjects were recorded using the *x-recorder* for further analysis. In addition to this, There are also phase tests at the end of each day. It is the same as the pre-test, and the material is 52 double-syllable words with target consonants. The purpose of the phase test is to examine the participants' production after daily training.

IV. PARTICIPANTS

Thirty second language subjects participated in the experiment. All the students were studying Chinese in Beijing universities or colleges, with an average age of 22. They came from Laos, Bangladesh, Thailand, Cuba, Pakistan, Japan and other countries. They studied Chinese for an average of one year, with the level of elementary. All the subjects had no physiological and psychological abnormalities and no experimental phonetics background. They will be paid after the experiment. All subjects were randomly divided into three groups, and the gender distribution was basically balanced.

The three groups were group A, which participated in NS-OME and S-OME training. Group B, which only participated in NS-OME training. Group C, which only participated in pre - and post-test and generalized test, and did not participate in any training. Each group consisted of 10 people and was randomly assigned by country.

Beside this, we invited three experts major in experimental phonetics to judge the test results of the subjects, and labeled all syllables of L2 learners with PET, and finally form the standard of the production.

V. PROCEDURE

This experiment adopts the design process of pre-test, training, phase test and post-test. The experiment lasted 5 days, about 40 minutes a day. Before the first training, the subjects were measured by pre-test of perception and production. After each day's training, take a phase test. After 5 days of training, the post-test was conducted, and the generalization test was conducted after the post-test. Taking group A as an example, the specific process of the experiment is arranged as follows:

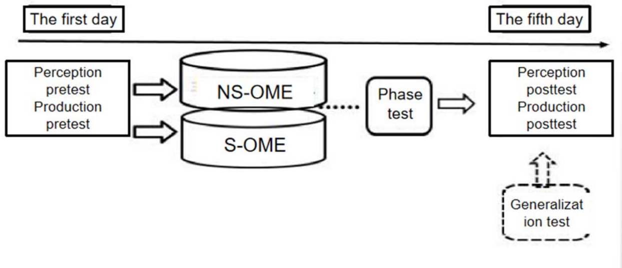

Fig. 5 Experimental process (for group A and B).

VI. RESULTS

A. *Perception tests*

Figure 6 and 7 shows the mean error rate of pretest and posttest perception under three groups.

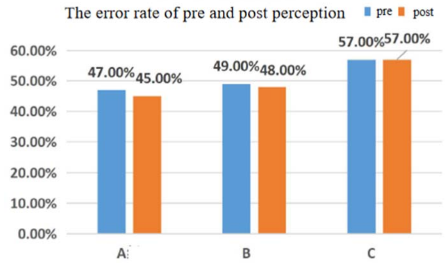

Fig. 6 The error rate of pre and post perception tests.

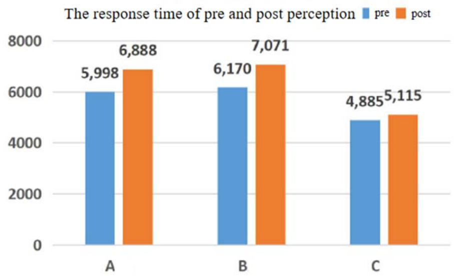

Fig. 7 The response time of pre and post perception tests.(/ms)

The results demonstrated that the perceived error rate of the three groups is relatively high, more than 45%. There are no significant differences in error rate of all groups($p=0.29$, $p>0.05$), which indicated that the three groups of subjects had difficulty in identifying the target consonants. This indicates that the training scheme has no significant effect on the perception of target consonants. In response time (RT), we found that the RT in post test increased in all groups compared with the pretest. The difference in RT between group A and group B was larger than that of group C.($p=0.012$, $p<0.01$). This may be because after training, subjects tend to adopt more careful experimental strategies than group C, and they may think and decide more carefully.

B. *Production tests*

The pretest, post test and generalization results of group A, B and C are shown in the following figure:

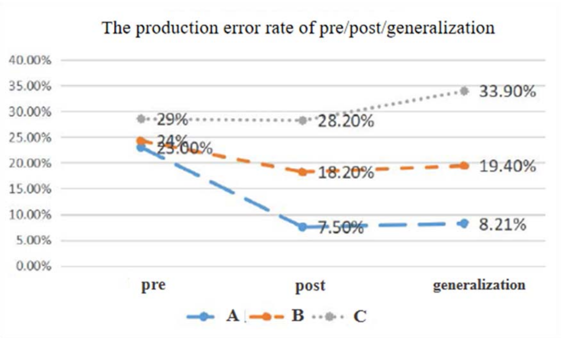

Fig. 8 The production error rate of pre/post/generalization tests.

As showed in the figure, the post-test error rate of group A was significantly reduced, which was about 15.5% lower than that of the pre-test (p=0.002, p<0.05). The error rate in the generalization test is equal to that in the post-test, which indicates that the training effect of the post-test has been maintained. Group B, which only participated in NS-OME, also had A lower error rate of 6.1% (p=0.046, p<0.05), but the decline was smaller than that of group A. The training result of group B was nearly maintained in the generalization test. For untrained group C, the average error rate of pre and post tests only slightly decreased by about 1%, while the generalized error rate increased by about 5% compared with the latter error rate. The error rate increased in the generalization test of group C.

Because the level of participants before taking part in the experiment is different, we used the relative progress rate to investigate the training effect. It is the rate at which the participants' errors on the later test are reduced compared with the previous test. The calculation method is as follows:

Relative progress rate = (post-test error rate - pretest error rate) / pretest error rate × 100%.

The relative progress rates of group A and group B are shown as follows:In order to explore whether there is a certain correspondence between production and perception in the training, we examined the confusion characteristics of the participants in production and perception. According to the result of group A and B's production and perception, Multidimensional scaling (MDS) figures were created and shown as follows:

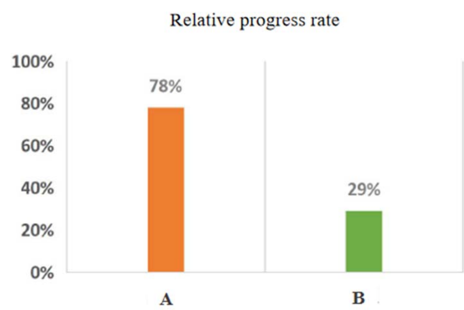

Fig. 9 The relative progress rates of group A and group B .

The results showed that the average relative improvement rate of group A after training was relatively high (78%). This shows that the training method is feasible and effective. The relative progress rate in group B was little, but also statistically significant.The results indicate that NS-OME also has a training effect on retroflex and non-retroflex consonants error in CSL learners. It is worth mentioning that the combination of S-OME and NS-OME training regimen can achieve the best training results.

VII. DISCUSSION

First of all, it can be found that there is no significant difference in the perceptual error rate of target consonants in all groups pre and post test, which indicates that the training program will not improve learners' perception of target consonants. Interestingly, after training, the subjects' production improved significantly. This shows that training can improve the production error of the subjects in a short time, and also proves the feasibility and effectiveness of the training method.

Second, after training, compared with group B, the production of group A participating in NS-OME and S-OME improved significantly. This indicates that the combination of NS-OME and S-OME training programs can achieve better results. The reason why the training effect of group B is not as good as that of group A may be that NS-OME does not focus on speech articulation. It lacks pertinence in training. Even so, as the result of group B is still better than control group C, it may suggest that NSOME may play a certain auxiliary role in the training, which is similar to some conclusions obtained from speech therapy.

In order to explore whether there is a certain correspondence between production and perception in the training, we examined the confusion characteristics of the participants in production and perception. According to the result of group A and B's production and perception, Multidimensional scaling (MDS) figures were created and shown as follows:

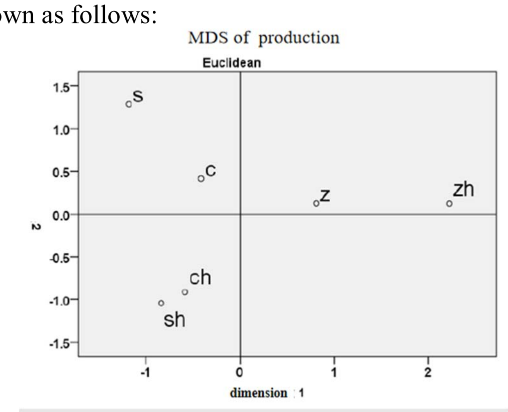

Fig. 10 The MDS of production from group A and B .

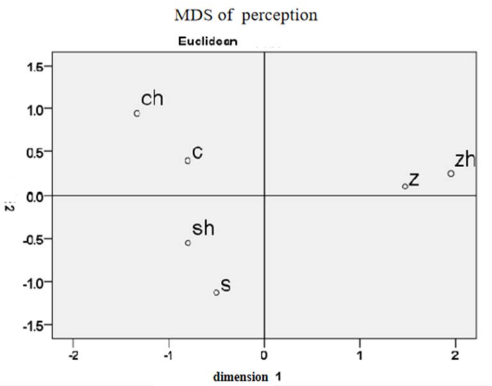

Fig. 11 The MDS of perception from group A and B .

As can be seen from the figures, there is a certain correspondence between perceptual confusion pairs and output confusion, and there are also differences. In the perceptual experiment, subjects tend to confuse the supradental and blade-palatal sounds. That is the confusion of the place of articulation. Meanwhile, It is easy for them to confuse affricate and fricative in the production experiment. Only z[ts] and zh[tʂ] are similar in the case of perceptual and production confusion.

VIII. Conclusion

Aiming at developing an efficient way to help CSL learners to master the retroflex and non-retroflex consonants in Mandarin, this study adopts the idea of Oral Motor Exercise (OME) to devise a speech production training method. The method consists of two steps: NS-OME, in which muscle groups of jaw, lips and tongue, essential to the production of target consonants, are intentionally exercised; S-OME, in which monosyllables with target consonants are practiced. 30 participants took part in the experiment and were divided into three groups. Group A, which participated in NS-OME and S-OME training. Group B, which only participated in NS-OME training. Group C, which as a control group, and only participated in pre\post-test and generalized test.

As a preliminary study, the results showed that both NS-OME and S-OME could effectively reduce subjects' consonant errors in a short time. When the two steps were accomplished, the trainee achieved the best performances. The results verified the feasibility and efficacy of the training method. This study not only supports the hypothesis that learning could be facilitated by breaking down complex movements into sub-components, but also provides a new perspective on the training methods of L2 learning.

IX. Acknowledgment

This study was supported by Advanced Innovation Center for Language Resource and Intelligence (KYR17005), the Fundamental Research Funds for the Central Universities (16ZDJ03, 18YJ030006, 19YXC126), and the project of "Intelligent Speech technology International Exchange" . Jinsong Zhang is the corresponding author.

References

[1] Zhao jinming. The possibility and basis of simplifying the teaching of phonology [J]. Language teaching and research, 1985. (In Chinese)
[2] Wu zongji. Introduction to modern Chinese phonetics [J]. Chinese teaching press, 1992.(In Chinese)
[3] Zhu yongping. Prediction of second language acquisition difficulty and teaching strategies [J]. Language teaching and research, 2004.(In Chinese)
[4] Zhu chuan, Ye jun, Cui liang, et al. Strategies for learning Chinese phonetics for foreign students [M]. Language and culture press, 1997.(In Chinese)
[5] Wang xiuzhen. Analysis of phonetic difficulties and errors in Korean learning Chinese [J]. World Chinese teaching, 1996(4):107-109.(In Chinese)
[6] Li dandan, Zhou xiao-bing. Perception and pronunciation relationship of Korean students in mandarin consonants learning [J]. Journal of Chinese teaching, 2005.(In Chinese)
[7] Si tong. Study on difficulties in learning Chinese consonants for Laotian students [D]. Guangxi university for nationalities, 2012.(In Chinese)
[8] Fu guo-qiang. Pronunciation errors in Chinese acquired by Lao students [J]. Modern Chinese (language research edition), 2013(11):114-116.(In Chinese)
[9] Cai chunying, Cao wen. Analysis of phonetic errors of Thai students [J]. Chinese teaching in the world, 2002.(In Chinese)
[10] Chen chen. Analysis of Chinese phonetic errors and countermeasures for Pakistani foreign students [D]. Jilin university, 2010.(In Chinese)
[11] Han shuang. Common problems analysis and coping strategies of Chinese phonetic vowels for Indian and Pakistani students whose mother tongue is Hindi or Urdu [D]. Xi 'an international studies university, 2017.(In Chinese)
[12] Ni weiman, Lin mingxian. Investigation on Chinese phonetics of Indonesian Chinese students and corresponding teaching countermeasures [J]. Journal of huaqiao university (philosophy and social science edition), 2000(2):70-73.(In Chinese)
[13] Wang gongping. Experimental research on Chinese initial perception among Indonesian students [J]. Language teaching and research, 2008(5):32-38.(In Chinese)
[14] Adams, J.A. A closed loop theory of behavior[J]. Journal of Behavior, 1971(3): 111-149.
[15] Schmidt, R.A. Motor control and learning. Champaign, IL; Human Kinetics, 1982.
[16] Maas E, Robin D A, Hula SNA, et al. Principles of motor learning in treatment of motor speech disorders[J]. American Journal of Speech-Language Pathology, 2008, 17(3): 277-298.
[17] Fitts, P.M. and Posner, M.I. (1967) Human Performance. Brooks/Cole, Belmont.
[18] Li shengli. Speech therapy [M]. Huaxia press, 2013.(In Chinese)
[19] Huang zhaoming, zhu qunyi, lu hongyun. Speech therapy [M]. East China normal university press, 2017. (In Chinese)
[20] Shao ruizhen. Educational psychology (revised edition)[M]. Shanghai: Shanghai education press, 1997 (In Chinese)
[21] Xun liu. Introduction to teaching Chinese as a foreign language [M]. Beijing language and culture university press, 2000.(In Chinese)
[22] Clark HM. Nonspeech oral motor intervention. In:Williams AL, McLeod S, McCauley RJ editor(s).Interventions for Speech Sound Disorders in Children.Baltimore, MD: Paul H. Brookes Publishing, 2010:579–99.

[23] Lee A S , Gibbon F E . Non-speech oral motor treatment for children with developmental speech sound disorders.[M]// The Cochrane Library. John Wiley & Sons, Ltd, 2011.
[24] Ruscello DM. Treating Articulation and Phonological Disorders in Children. St. Louis: Elsevier, 2008.
[25] Ruscello DM. Nonspeech oral motor treatment issues related to children with developmental speech sound disorders. Language, Speech, and Hearing Services in Schools 2008;39(3):380–91.
[26] Qiu weihong. Evaluation of articulation disorder and speech therapy [J]. China clinical rehabilitation, 2004. (In Chinese)
[27] Li yong-xue, zhang tao, Chen shao-zhen. Effects of speech correction combined with oral motor training on functional articulation disorders [J]. Massage and rehabilitation medicine Science, 2015, (In Chinese)
[28] Xu lina, li feng, zhang yanyun. Phonetic characteristics and rehabilitation training of abnormal lingual root sounds in children with functional dysarticulation [J]. Huaxi kou Journal of cavity medicine, 2016. (In Chinese)
[29] Min zhiyun, li feng, xu lina. Effect of intensive oral motor training on speech articulation in children with functional dysarthria [J]. Listen Journal of mechanics and speech disorders, 2018. (In Chinese)
[30] Nosofsky R. Attention, similarity, and the identification categorization relationship. J Exp Psychol: General; 115:39–57. 1986.