

A COMBINATION OF METALINGUISTIC INSTRUCTION AND TASK REPETITION IN TEACHING CHINESE PROSODY

Jiang Liu, University of South Carolina

When learning Mandarin, where f0 differentiates word meanings and carries higher-order communicative function, prosody errors cause reduced comprehensibility (Yang, 2016). Thus, teaching prosody in Mandarin is critical to improving students' pronunciation. We compared three classes of L2 Chinese learners' prosody at the sentence level. The Control group recorded a list of target sentences in a scripted dialogue after doing a role-play activity in class at Time 1. Two trainee groups (Task-Repetition and No-Repetition groups) who did the same role-play activity received additional explicit prosody teaching. They recorded the target sentences after the explicit teaching as Time 1 recording. Then the Task-Repetition group repeated the role-play activity while the NoRepetition group did not. All three groups recorded the same target sentences at Time 2. The results showed that both training groups had significantly better comprehensibility ratings than the Control group at Time 1, and the Task-Repetition group retained the gain but the No-Repetition group did not at Time 2. These findings suggest that task repetition is necessary to maintain learned prosodic patterns.

INTRODUCTION

Previous research has shown that L2 learners' phonological development is the most active during the early stages of language learning or exposure to the L2 (Flege, 1988; Munro & Derwing, 2008). It has been suggested/recommended that teachers should focus on segmentals and suprasegmentals such as basic intonation in beginning level classes so that the learners can maximize their opportunity to develop L2 phonology in this critical time window (Darcy, Ewert & Lidster, 2012). However, in practice, it is often found that language instructors feel they lack training on pronunciation teaching; thus, they can only teach the subject with their intuitive notion of intelligibility and foreign accent (MacDonald, 2002; Henderson et al., 2012; Foote, Holtby & Derwing, 2011). Among various cues that provide grounds for instructors to judge L2 English learners' comprehensibility or foreign accent, prosodic features such as speaking rate, stress, rhythm, or intonation stand out to be the salient cues (e.g., Kang, 2010). Similarly, comprehensibility has been found to be strongly affected by prosody in Mandarin Chinese (Yang, 2016). In Chinese, there is both lexical prosody and phrase-/utterance-level prosody. Native Chinese listeners have been found to be more sensitive to prosodic deviations, including tone deviations in L2 Chinese, than non-tone language listeners (Yang, 2016). When native English speakers learn Chinese, many tend to produce sentences in a bottom-up way with more pauses than native Chinese speakers (Chen, 2014). In other words, learners ranging from beginner to advanced level have a tendency of speaking character by character or word by word, thus, causing two types of unnaturalness: (1) hyper-articulation of tones; (2) non-nativelike prosodic features (e.g., intonation, stress, pitch accent, final lengthening, boundary cues).

So far, the discussion of best practice of prosody teaching at the sentence level in L2 Chinese is scarce. Yang and Jin (2018) conducted a survey about pronunciation teaching in L2 Chinese. Among 80 respondents, the majority of instructors reported that they preferred to teach

pronunciation in a meaningful context and in an integrative and systematic approach, instead of teaching it separately. The most common approach used was to provide corrective feedback during doing communicative language speaking tasks. However, so far, few empirical studies have investigated the effectiveness of prosody teaching in L2 Chinese.

Mandarin as a tonal language uses tones to differentiate word meaning (Yip, 2002) and encodes various types of communicative functions such as focus, statement, and interrogatives (Xu, 2005). There are several non-trivial challenges for native English speakers to learn prosody in L2 Chinese. The first area is the articulatory mechanism for producing lexical tones with distinct register features (i.e., pitch range) and contour features (i.e., pitch contour) as well as how to pronounce lexical tones in various contexts known as tonal coarticulation (Duanmu, 2007; Xu, 1994, 1999, 2002). The second area is the prosodic structure of Mandarin. As an utterance can range from a monosyllabic word to multi-syllabic words, L2 Chinese learners need to know how to segment a long sentence into parts within which the transition between lexical tones occurs. However, the segmentation in an utterance does not always align with the syntactic segmentation (e.g., phrase structure). The prosody of an utterance has a hierarchical structure starting from syllables to foot to prosodic word and phrase (Tseng, Pin, Lee, Wang, & Chen, 2005). Utterance-level prosody refers to the prosodic phrasing, as well as the F0 and duration patterns, at the utterance level. Previous studies on L2 Chinese prosody mostly focus on describing the prosodic error patterns and discuss the results in relation to phonological and phonetic theories (e.g., Zhang, 2010; Yang, 2016). Although pedagogical implications are discussed, few training paradigms try to teach L2 Chinese prosody. As Munro and Derwing (1995) point out, pronunciation teaching should not aim to reduce the foreign accent, recognizing the partial independence of foreign accent, intelligibility and comprehensibility. In other words, a strong accent does not entail a failure of understanding the L2 speech. Rather, we should aim to increase the intelligibility and comprehensibility of L2 learners' speech. The current study explored the effectiveness of two training paradigms that aimed to improve L2 Chinese learners' comprehensibility by teaching them the prosody of Chinese.

As previous research has pointed out that instructors receive very little or no training on metalinguistic instruction of pronunciation teaching, in the current study, the experimenter who is the course coordinator at a public university in the U.S. first taught the instructors some basic metalinguistic knowledge about Chinese prosody and trained them how to provide one-on-one pronunciation teaching. After the teacher training, the instructors provided systematic pronunciation teaching that focuses on prosody to two sections of a beginner level Chinese class, which we treated as the trainee groups. By comparing learners' pronunciation in the two classes to the one in a Control group in which students did not receive any prosody teaching, we can test how robust the prosody teaching is. In addition to the explicit prosody teaching, we had one trainee group repeat a communicative task as part of their pronunciation learning. Previous research showed that repetition of a communicative task in the classroom improved oral proficiency (e.g., fluency and accuracy) in a second or later task cycle. However, most of the oral production gain comes from grammar accuracy and fluency (e.g., Bygate, 1996, 2013). Whether task repetition can be effective for prosody learning is still an open question. Some study has shown that a combination of task repetition and explicit pronunciation teaching benefit L2 pronunciation learning (Trofimovich & Gatbonton, 2006). The current study tried to explore whether such combination can benefit prosody teaching at sentence level in L2 Chinese.

Research Questions

1. Does explicit prosody teaching improve the pronunciation at the sentence level in L2 Chinese?
2. Does task repetition help retain and further improve pronunciation at the sentence level?

METHODS

Participants

L1-English L2-Chinese learners enrolled in three different sections of a beginning level Chinese course in a public university in the U.S. participated in the current study. Two sections taught by the same instructor (Section 1: 12 students; Section 2: 14 students) received pronunciation teaching at the beginning of the semester whereas the third section (13 students) taught by another instructor did not receive any pronunciation teaching until the end of the semester. Because of the uneven number of students in each section, we randomly selected speech samples of 12 participants from each section for the comprehensibility rating. None of the participants were heritage speakers. The mean age of onset for the group was 15.12 (SD = 2.11) years, and learners reported an average of 1.01 (SD = 1.15) years of Chinese instruction prior to university. To establish a baseline, we also included ten native speakers (NS) to record the target sentences.

Stimuli

Each participant recorded twelve sentences (6 questions and 6 statements) in a scripted dialogue taken from the textbook (Integrated Chinese, level 1, part 1; Liu & Yao, 2009), which were used for the comprehensibility rating. The target sentences are from two dialogues about making an appointment. Depending on the type of the question (e.g., a full question: 老师, 今天下午您有时间吗? ‘Professor, do you have time today’s afternoon?’ or a tag question: 明天呢? ‘How about tomorrow?’), the number of syllables for the question ranged from three to ten. For the questions, students needed to know how to segment the sentence into prosodic phrases (e.g., 今天下午 today’s afternoon + 您有时间 you have time + 吗 question mark ?). Students also needed to be aware of the pitch transition from one syllable to the toneless question marker such as 吗 and 呢. The declarative sentences used as answers to questions in the dialogue also consisted of utterances ranging from two to nine syllables. (e.g., 没有 ‘no’, 明天上午我有两节课. ‘Tomorrow morning I’ll have two classes.’). Students recorded all these target sentences at two times. The total number of students’ utterances was 864 (3 groups x 12 participants x 12 sentences x 2 times) and the total number of native speakers’ utterances was 120.

Training and recording

We first let all students do a role-play activity (i.e., making an appointment) in class. After class, all students were given a dictation assignment and dialogue completion exercise to enhance their understanding and memory of the target sentences. Then, the **Control group** was asked to record the target sentences in a scripted dialogue at **Time 1**. We provided explicit prosody teaching to the two trainee groups (**No-Repetition** and **Repetition groups**) by asking them to attend a one-on-one pronunciation clinic during the instructor’s office hours. The instructors asked students to read the scripted dialogue and gave students feedback about their pronunciation of those target sentences. More importantly, the instructors taught students how to group words and phrases in the long sentences and pronounce the sentences with a rhythm and melody such as when they

needed to stress certain words based on the message they intended to deliver, when to lengthen a syllable, or when to make a short pause. In addition to the prosody teaching, we also taught students the discourse of the dialogue such as which part of the dialogue showed a failed initial attempt to set up an appointment, which part was negotiating time and discussing schedule, which part was about checking availability, etc. After the pronunciation clinic, the two trainee groups recorded the target sentences in the scripted dialogue as their **Time 1** recording. By comparing the Control group to the two trainee groups' comprehensibility rating at Time 1, it can inform us whether explicit prosody teaching helps to improve sentence pronunciation. At the end of the chapter instruction, the Task-Repetition group was given another round of role-play activity in class. When doing the activity, the instructor provided recasts when hearing noticeable pronunciation errors students made, particularly prosodic ones. The No-Repetition trainee group, however, did not do the second round of role-play activity. After the Task-Repetition group completed the second round role-play activity, all three groups recorded the scripted dialogue for the second time. We refer it to be the recording at **Time 2**. By comparing the three groups' comprehensibility rating score at Time 2, we can test whether task repetition can help students retain and further improve their sentence pronunciation.

Comprehensibility Rating

One goal of this study was to train the Chinese language instructors' awareness of comprehensibility. We asked the two instructors, who are both native Chinese speakers, to rate the comprehensibility of the collected speech samples. Comprehensibility is defined as how easy a listener understands the meaning of words and utterances (Munro & Derwing, 1995; Kirkpatrick, Deterding, & Wong, 2008; Smith & Nelson, 1985). Each listener listened to 432 utterances (3 groups x 12 participants x 12 sentences) at two time points respectively. The rating task was carried out individually on a personal computer in a quiet office or room in the language lab. First, the native listeners listened to a few samples from all students to avoid biased ratings with the initial samples (Derwing, Rossiter, Munro, & Thomson, 2004). Prior to completing the task, they rated three practice items (not used in the experiment). The native listeners then completed the rating task by listening to the randomized speech samples one by one and rated each one by choosing a value on a 9-point Likert scale on a printed questionnaire, with '1' being the easiest to understand and '9' cannot understand at all. The listeners heard each complete sample only once before making their judgments and were instructed to use the entire scale. An experimenter controlled the playback of the stimuli to ensure that all raters stayed in step. It took about 40 minutes to complete the ratings for Time 1 and 2.

RESULTS

The inter-rater reliability coefficients (Cronbach's alpha) computed across all ratings given by two native listeners were very high at both Times 1 and 2 (Time 1: .89, Time 2: .92). This indicated a strong agreement (Larson-Hall, 2015). As expected, the raters found the Native Speaker group the most comprehensible, as seen in the mean scores reported in Table 1.

Table 1.
Mean comprehensibility ratings with standard deviation in the bracket

	Time 1	Time 2
Native speakers (n=10)	1.5 (0.04)	1.5 (0.04)
Control (n=12)	6.2 (0.15)	6.0 (0.14)
No-Repetition (n=12)	4.8 (0.12)	5.8 (0.12)
Repetition (n=12)	5.2 (0.12)	4.4 (0.13)

To further illustrate the distribution of comprehensibility ratings within each group, Figure 1 shows the boxplot¹ of the four groups of participants at Time 1 and 2. We averaged the comprehensibility ratings across the two listeners for each sentence item of each participant. We noticed that the native speakers received ratings with very small variability. Among the three learner groups, the Control group showed a large amount of variability compared to the two training groups at both Time 1 and 2 of the recording. The large individual variability found in the Control group is consistent with previous finding that learners' individual difference in terms of sensitivity to non-linguistic pitch led to different lexical tone learning outcome (Chandrasekaran et al., 2010). For the two trainee groups, with pronunciation intervention at Time 1 and 2, the variances were noticeably reduced.

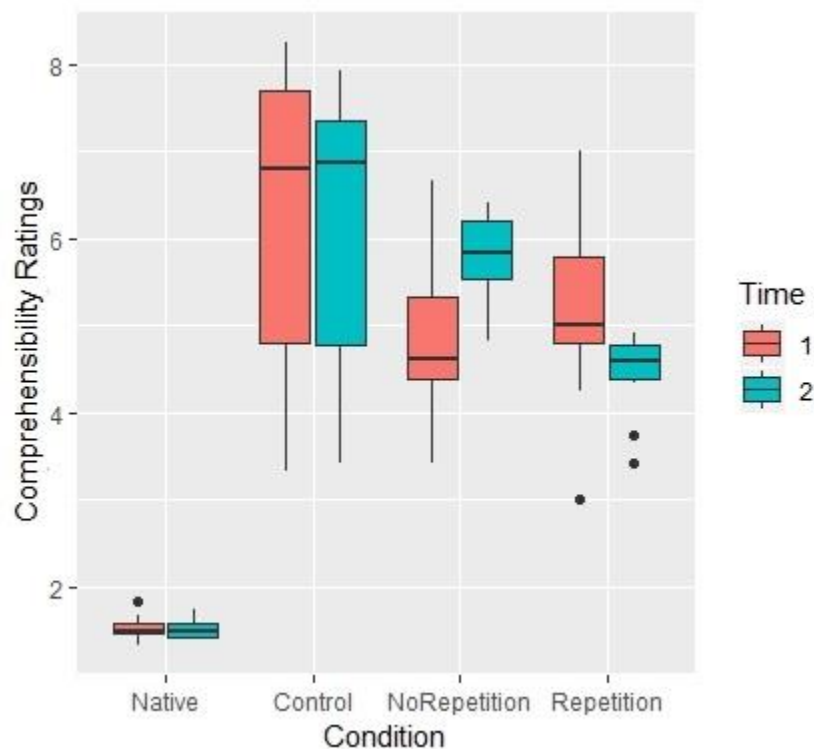


Figure 1. Boxplot of comprehensibility ratings by Condition over Time.

For statistical analysis, a linear mixed-effects model was built with the lme4 package in R (version 3.5.1; R Core Team, 2017) with Condition (Control, NoRepetition, Repetition) and Time (1 and

¹ Although the boxplot does not indicate the mean scores for each group, it better illustrates the dispersion of the data in each group. Thus, it is a supplement to the mean rating scores reported in Table 1.

2) as fixed factors and by-learner and by-sentence as random effect. The dependent variable was the comprehensibility rating score averaged across the two listeners for each item per participant. Model fit, including selection of predictors and their interactions, the random effect structure, and reported p-values were derived from the χ^2 -test of the change in deviance between the models with and without the effect of interest. Table 2 reports the optimal model output and R code. Table 2 shows that at Time 1 both NoRepetition and Repetition groups had significantly lower ratings on the comprehensibility scale than the Control group (RQ1). The result suggests that explicit prosody teaching made an improvement in comprehensibility. The result also showed a significant Condition x Time interaction. As Figure 1 shows, from Time 1 to 2, the No-Repetition group was rated as less comprehensible while the Repetition group was rated as more comprehensible. By releveling the reference level from Time1 to Time 2, we found that only the Repetition group had significantly lower rating on the comprehensibility scale than the Control group at Time2 whereas the NoRepetition group did not differ from the Control group at Time2 (NoRepetition vs. Control: $\beta = -0.28$, SE = .44, $t = -0.63$, $p = 0.53$; Repetition vs. Control: $\beta = -1.63$, SE = .44, $t = -3.63$, $p < .001$). The result suggested that task repetition helped to retain and improve comprehensibility (RQ2). Interestingly, the NoRepetition group actually had a worse comprehensibility rating at Time 2 than Time 1 (with higher rating on the scale). The result showed that without task repetition the pronunciation actually deteriorated.

Table 2.

Summary of fixed-effects coefficient estimates and significance values for linear mixed-effects model output

Fixed effects:	Estimate	Std. Error	df	t-value	Pr(> t)	
(Intercept)	6.39	0.32	37.82	16.61	< 2e-16	***
NoRepetition	-2.48	0.45	37.08	-4.58	1.77e-05	***
Repetition	-1.08	0.45	37.08	-2.41	0.021	*
Time2	-0.15	0.15	814.00	-0.96	0.335	
NoRepetition:Time2	1.10	0.21	814.00	5.13	3.64e-07	***
Repetition:Time2	-0.55	0.21	814.00	-2.57	0.010	*

[R code: `lmer(Comprehensibility~Condition*Time+(1|learner)+(1|sentence))`. Reference level for Condition is 'Control group'; Reference level for Time is Time1]

(Note: 'learner' in the model means the learner who produced the sentence. 'sentence' means the specific sentence out of 12 sentences)

DISCUSSION

To address the first research question whether explicit prosody teaching helps to improve sentence pronunciation in L2 Chinese, we found that the explicit prosody teaching indeed helped to improve students' comprehensibility as the two trainee groups that received the teaching showed better comprehensibility than the Control group. Although all three groups did in-class role-play activity and homework related to the recorded dialogue, the two trainee groups who received explicit prosody teaching outperformed the Control group in terms of comprehensibility. This suggests that doing homework may enhance the familiarity of the content but it may not directly help to improve the pronunciation.

The most interesting finding in the present study is that only the Task-Repetition group retained and improved their gain in comprehensibility. This answers RQ2, showing that the Task-Repetition group not only retained their advantage to the control group in terms of

comprehensibility but also improved the comprehensibility of the target sentences from Time 1 to Time 2. The No-Repetition group did not retain their gain from Time 1 to Time 2 when compared to the Control group. The comprehensibility of the No-Repetition group became poorer at Time 2. This suggests that prosody teaching can result in an initial gain in comprehensibility but Task-Repetition is necessary to maintain the gain. Otherwise, the gain can potentially diminished.

There has been a long debate about how important suprasegmentals are for communication in English as an international language (Jenkins, 2000; Levis, 2005). Teachers tend not to use communicative speaking tasks to teach prosody specifically. Similarly for L2 Chinese teaching, many instructors believe that lexical tones are already difficult enough for learners whose first language is not a tonal language, let alone intonation and other prosodic features. Therefore, few attempts have been made to use task-repetition to improve L2 Chinese learners' prosody. This study tries to bring CFL teachers' attention to prosody teaching in Chinese by trying out a combination of metalinguistic instruction of prosody and task-repetition.

One challenge to deliver effective prosody teaching lies in teacher training. It is not uncommon that native speakers of a language do not realize how complicated the structure of their mother tongue is until they are taught explicitly about it. This is true for the Chinese instructors (mostly Graduate Teaching Assistants) working in our Chinese program. Through the comprehensibility rating task and preparation for the pronunciation clinic, the instructors reported that it increased their awareness of prosody in Chinese and the types of prosodic errors L2 Chinese learners make when producing sentences. They became more aware of L2 Chinese learners' comprehensibility through prosody teaching.

LIMITATIONS

The current study did not include a true pre-test for the two trainee groups. Thus, the effect of metalinguistic instruction can only be tested between the trainee groups and the Control group but not within the trainee groups. In the future, we should also include a pre-test for the trainee groups to have a more robust test on the intervention methods for pronunciation teaching.

ABOUT THE AUTHOR

Jiang Liu is an Assistant Professor in the Department of Languages, Literatures and Cultures and Linguistics Program at the University of South Carolina, USA. He obtained a PhD in Linguistics from the University of Kansas. His research interests include phonetics and phonology and L2 speech learning.

Contact

Department of Languages, Literatures and Cultures
Linguistics Program
University of South Carolina
709 Welsh Humanities Building
1620 College St. | Columbia, SC 29208
Telephone: 803-576-7291
Email: jiangliu@mailbox.sc.edu

REFERENCES

- Bygate, M. (1996). Effects of task repetition: appraising learners' performances on tasks. In J. Willis and D. Willis (Eds.) *Challenge and change in language teaching* (pp. 136-146). Oxford: Heinemann.
- Bygate, M. (2013). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, M. Swain (Eds.). *Researching pedagogic tasks* (pp. 23-48). New York: Routledge.
- Chandrasekaran, B., Sampath, P. D., & Wong, P. C. (2010). Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, 128(1), 456-465.
- Chen, M. (2014). Simultaneous pauses in Chinese as second language 汉语作为第二语言的自发性言语停顿模式研究. *Experimental Linguistics 实验语言学*, 3(2), 1-14.
- Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox". In J. Levis and K. Lavelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, pp. 93-108.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679.
- Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford: Oxford University Press.
- Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America*, 84(1), 70-79.
- Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, 29(1), 1-22.
- Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Chungham, U. & Curnick, L. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, 10(1), 5-27.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38(2), 301-315.
- Kirkpatrick, A., Deterding, D., & Wong, J. (2008). The international intelligibility of Hong Kong English. *World Englishes*, 27(3-4), 359-377.
- Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R*. New York: Routledge.
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *Tesol Quarterly*, 39(3), 369-377.
- Liu, Y.H. and Yao, D.Z. (2009). *Integrated Chinese: Zhong Wen Ting Shuo Du Xie* (3rd Edition). Boston, MA: Cheng & Tsui.
- MacDonald, S. (2002). Pronunciation: Views and practices of reluctant teachers. *Prospect*, 17(3), 3-18.
- Munro, M. J. & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.

- Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58(3), 479-502.
- Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: Directions and resources. *World Englishes*, 4(3), 333-342.
- Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal*, 90(4), 519-535.
- Tseng, C. Y., Pin, S. H., Lee, Y., Wang, H. M., & Chen, Y. C. (2005). Fluent speech prosody: Framework and modeling. *Speech Communication*, 46(3-4), 284-309.
- Xu, Y. (1994). Production and perception of coarticulated tones. *The Journal of the Acoustical Society of America*, 95(4), 2240-2253.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, 27(1), 55-105.
- Xu, Y. (2002). Maximum speed of pitch change and how it may relate to speech. *The Journal of the Acoustical Society of America*, 111(3), 1399-1413.
- Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, 46(3-4), 220-251.
- Yang, C. (2016). *The acquisition of L2 Mandarin prosody: From experimental studies to pedagogical practice*. Philadelphia: John Benjamins Publishing Company.
- Yang, C. & Jin, W. (2018). Chinese as a second language pronunciation teaching survey. *Journal of National Council of Less Commonly Taught Languages*, 23, 153-188.
- Yip, M. (2002). *Tone*. Cambridge University Press.
- Zhang, H. (2010). Phonological universals and tone acquisition. *Journal of the Chinese Language*, 45(1), 39-65.