

Learning English Intonation Through Exposure to Resynthesized Self-produced Stimuli

Zhongmin Li^a

lzm1988jimmy@126.com

School of Foreign Languages,
Suranaree University of Technology, Thailand

Andrew-Peter Lian^b

mlapl1@gmail.com

School of Foreign Languages,
Suranaree University of Technology, Thailand

Butsakorn Yodkamlue

butsakorn@sut.ac.th

School of Foreign Languages,
Suranaree University of Technology, Thailand

ABSTRACT

EFL learners are prone to having problems in pronunciation, while their problems in intonation are more salient. The Chinese EFL pronunciation classroom has long been criticized for teacher-centered, “one-size-fits-all” teaching, which is inefficient and ineffective for solving individual student’s specific pronunciation problems. This study conducted an experiment to examine the effectiveness of exposure to resynthesized self-produced stimuli for intonation learning. The participants were 66 first year English majors studying at a university in China. The treatment was a form of English intonation training wherein the students in the experimental group used their resynthesized self-produced stimuli (their own voices) as the pronunciation model for learning while the control group used a model produced by a native speaker. After the training, the results of the intonation production test showed that the experimental group outperformed the control group in eight intonation patterns. The students’ problems in intonation support Mennen’s (2007) claim that intonation learning involves a first stage of acquiring the phonological representations of intonation patterns and a second stage of acquiring the phonetic realizations of those patterns. The results of this study revealed that exposure to resynthesized self-produced stimuli for intonation learning was as effective as the native speaker model for helping the students form the phonological representations of intonation patterns, while it was more effective than the native speaker model for facilitating the students to produce more accurate phonetic realizations of those patterns.

Keywords: English intonation; precision language education; modified stimuli; phonological representation; phonetic realization

INTRODUCTION

Intonation is notoriously difficult to teach (Lengeris, 2012). Romero-Trillo (2012) even claimed that “of all the elements of a target language, intonation appears to be the most difficult to acquire” (p.89). The reason might be that “the complexity of the total set of sequential and prosodic components of intonation and of paralinguistic features makes it a very difficult thing

^a (Main author)

^b (Corresponding author)

to teach” (Roach, 2009, p. 11). As a result, intonation instruction has long been neglected in EFL classrooms. Some teachers hold that intonation cannot be taught but can only be acquired through long term exposure to the language. Other teachers believe that intonation should be taught while they do not have adequate knowledge and therefore lack the confidence to teach it (Lengeris, 2012). All these made intonation the “problem child” of pronunciation teaching (Dalton & Seidlhofer, 1994).

For several decades, some researchers have appealed for a paradigm shift of pronunciation teaching: the teaching priority should be shifted from the segmental features to the suprasegmental features (Morley, 1991; Jenkins, 2004; Kang, 2010; Gilbert, 2014). Previous studies have also shown that intonation contributes to speech accentedness and intelligibility more than individual sounds do (Anderson et al., 1992; Jilka, 2000; Pickering, 2001; Kang, 2010). This means that a speaker is likely to have a stronger accent and is less likely to be understood by the listeners if s/he is weak in intonation, despite having the ability to pronounce every individual sound accurately.

EFL learners are prone to having problems in English pronunciation. At the same time, their problems in intonation are always more salient. Zhang (2015) investigated the international intelligibility of the English spoken by Chinese students and found that their speech was largely intelligible to international listeners but was frequently evaluated by the listeners as “strongly accented, fast, choppy, monotonous, truncated and hesitant” (p.51). These negative evaluations are related to the suprasegmental aspect of pronunciation, which were also supported by previous studies on Chinese EFL students’ English intonation, such as stress-timed rhythm patterns (Zhu, 2007), incorrect tone choices (Makarova & Zhou, 2006; Rui, 2007; Huo & Luo, 2017), inappropriate pauses (Yang & Mu, 2011), and incorrect assignment of intonation boundaries (Yang, 2006; Meng & Wang, 2009).

Bi and Chen’s (2013) cross-sectional study revealed that Chinese EFL university students’ problems in intonation remained unchanged throughout four years of study, while the segmental aspect was improved to some extent. This means that English intonation teaching and learning in universities of China was ineffective for solving the students’ problems in intonation. Furthermore, the classrooms in universities of China are usually large in size (with 30 students or so), and teachers always follow a teacher-centered, “one-size-fits-all”, style of teaching. The students rarely have a chance to receive feedback and there is also very limited time for them to practice pronunciation in class. Therefore, it is urgent to find an effective and efficient way to improve Chinese EFL university students’ English intonation learning.

To fill this gap, the present study attempted to validate a way of strategically using students’ own speech productions for intonation learning, i.e., providing the students with their own resynthesized self-produced stimuli as the model for learning. Informed by the concept of precision language education (Lian & Sangarun, 2017), instruction of this kind focused on individual student’s specific pronunciation problems and provided precise corrective feedback to their problems. Specifically, this approach took students’ incorrect intonation, digitally altered it to the correct one and fed it back to the students as the input for learning, so as to raise their awareness of their problems and improve their pronunciation performance.

LITERATURE REVIEW

THE PHONETICS AND PHONOLOGY OF ENGLISH INTONATION

Intonation is defined as the linguistic use of pitch variations in utterances (Tench, 2015). There are different approaches to intonation analysis, amongst which the most influential are the British school’s pitch contour approach and the American school’s autosegmental metrical approach. The British school holds that the system of English intonation consists of tonality, tonicity, and tone. Wells (2006) integrated the three “Ts” into the mechanism for making

decisions about intonation when people are speaking, i.e., how to break up the material into chunks (tonality), what is to be accented (tonicity), and what tones are to be used (tone).

Tonality involves the use of intonation boundaries to break the utterances into chunks. One chunk constitutes one intonation group. Tonality in speech plays a role like punctuation in writing (Halliday, 2015), so intonation has the function of disambiguating ambiguous syntactic structures. Tonicity concerns the focus of information which is carried by the nucleus of intonation group. The nucleus is likely to rise when the speaker wants to express a narrow focus—the particular information that the speaker wants to emphasize; a contrastive focus—the information that forms a contrast with the corresponding information that has been said; and new information (Cruttenden, 1997). Tone refers to the falls and rises of pitch contour. Tench (2015) and Wells (2006) recognized three types of primary tones that can lead to contrastive pitch movements: falling tone, rising tone, and falling-rising tone. The rising-falling tone was not included because it generally has the same function as the falling tone, while the falling-rising tone often “signals particular implications” (Wells, 2006, p.10).

The American school proposed that intonation consists of a phonological component and phonetic component (Botinis et al., 2001; Beckman & Pierrehumbert, 1986). The phonological aspect is exercised in the associations of a set of high tones (H) and low tones (L). The phonetic realizations of these tones are measured by two parameters: the scaling (pitch value) and the alignment (the temporal relations of the segments). Language learners’ phonological problems of L2 intonation may result from the intonational differences in the inventory of L1 and L2 phonological patterns. For example, English and Korean are different in distinguishing between yes/no-questions and wh-questions (MacDonald, 2011). The phonetic problems of L2 intonation result from the differences between L1 and L2 in the phonetic realization of the same phonological pattern, e.g., the English spoken by Dutch learners were featured by a smaller declination rate for falling tones and a lower starting pitch for rising tones (Willems, 1982).

Mennen (2007) claimed that L2 intonation learning involves two stages. L2 learners may first acquire the phonological patterns before they acquire the correct phonetic realizations of these patterns. His former studies (1999, 2004) found that Dutch learners of Greek were perfectly able to produce the correct phonological tonal elements but implemented these structures by using L1 phonetic regularities. As L1 interference can be an overwhelming factor for L2 intonation learning (Swerts & Zerbian, 2010), it is necessary to investigate the characteristics of the English intonation produced by Chinese learners, tonal language speakers who are learning a non-tonal language. Though previous studies have shown that Chinese EFL learners had various problems in English intonation, no studies separated their problems in phonological representation from phonetic realization. Therefore, this study aimed to examine whether Chinese EFL learners’ problems in intonation reflect the two-stage process, and whether exposure to self-resynthesized stimuli would solve those problems in each stage.

PRECISION LANGUAGE EDUCATION

Precision education was inspired by the concept of precision medicine which is an innovative approach to personalizing healthcare delivery. Precision medicine takes individual differences (such as patient’s genes, environments, and lifestyles) into consideration and enables medical professionals to tailor treatment to each patient’s unique needs. The precision medicine approach to treating diseases can be ideally applied to the field of education in dealing with learning disabilities, because learning disabilities are a variety of disorders with remarkable similarities to biomedical diseases (Hart, 2016). The current educational system has been designed for the benefits of most students or the average student, with uniform instruction, broad assessment and fixed teaching methods. However, each student is unique and each

student's learning disabilities are unique, so students should be treated as a group with considerable individual differences (Cook et al., 2018), so as to "get the right intervention in place for the right person for the right reason" (p. 5).

To specify the application of precision education in the field of language education, Lian & Sangarun (2017) proposed the concept of precision language education.

Precision language education heralds a new way of dealing with individual differences by effecting as precise a diagnosis as possible on each language learner, thus triggering specific interventions designed to target and respond to each person's specific language-learning problems. (p. 1)

Lian's computer-based answer-evaluation and markup system (see Lian & Sangarun, 2017) is a good example of instructional activity based on precision language education, which provides precise feedback for listener's answers in a listening-transcription task. It uses the student's own input to identify whether students' answers are correct or not, and then provides specific feedback to help students repair the identified problems. By so doing, students are able to modify their perceptual and comprehension systems according to their specific problems and get closer to the correct answers.

Considering the individuality and variability of students' problems in pronunciation, no student characteristic can dictate a priori what interventions will work, and nor will a given intervention be effective for all students all of the time (Fuchs et al., 2003). Therefore, the researchers of the present study hold that pronunciation instructions or interventions should be designed catering to individual student's characteristics and tailoring to their specific needs. However, in pronunciation teaching practice, teachers often select interventions in a trial-and-error fashion based on predictions or conventions, and they focus more on how to set or select good models for students to imitate rather than on how to deal with students' specific personal problems. In other words, they are trying to solve all students' problems by resorting to the correct model rather than identifying ways to correct individual students. This is the same as "shooting in the dark and hitting targets indiscriminately" (Cook et al., 2018, p. 5).

Speech input stimulus modification can be a representative for individualized pronunciation teaching, which entails efforts to modify the properties of the input speech signal to make it a better fit for individual learner's perceptual mechanisms so as to improve their production performances. Informed by the concept of precision language education, the present study attempted to solve the students' problems in English intonation by modifying the students' incorrect output and feeding it back to them as the input (resynthesized self-produced stimuli) for learning. This approach explicitly targeted each student's specific problems and sought to provide tailored interventions to solve these problems, thus enabling learning to occur in a "just in time, just enough, just for me" (Lian, 2014) fashion.

SPEECH STIMULUS MODIFICATION IN INTONATION INSTRUCTION

Speech stimulus in pronunciation instruction refers to the speech input provided to the learners to facilitate their perception or production improvement. A speech stimulus can generally be classified into three types: a natural stimulus, a stimulus naturally produced by a human; a synthetic stimulus, a stimulus produced by a machine; and a resynthesized stimulus, that was initially produced by a human but underwent a process of acoustic modification by a machine. A natural stimulus, on the one hand, contains rich redundancy of acoustic cues (Lively et al., 1993) while, on the other hand, because of its diversity and variability, it might be difficult for learners to perceive the critical acoustic cues precisely. A synthetic stimulus is purposefully generated to highlight critical cues but might contain misleading or incomplete information (Logan et al., 1991). In comparison, a resynthesized stimulus avoids the deficiencies of natural

and synthetic stimuli while making the acoustic cues more salient based on the modification of naturally produced speech.

Given its advantages, resynthesized stimuli have been frequently used for speech perception or production training in previous studies, especially for the training of vowel contrasts. In the suprasegmental aspect, related studies are still limited, while some researchers have explored the techniques for prosody modification. For example, Yoon (2009) introduced the technique of exaggerating prosody by manipulating the fundamental frequency contour, the duration, or the intensity of an utterance, which supported the viability of using exaggerated speech stimuli for pronunciation training. Lu, Wang and De Silva (2012) employed automatic stress exaggeration to enlarge the differences between stressed and unstressed syllables in order to help learners to better perceive sentence stress.

Tang, Wang and Seneff (2001) invented a voice transformation technique which can flexibly manipulate the suprasegmental features of speech signals. Felps, Bortfeld and Gutierrez-Osuna (2009) invented another voice-transformation technique which can be used to transfer native speaker's accent to student's speech. This technique can reduce foreign accentedness without significantly altering the voice quality properties of the learner. Pellegrino and Vigliano (2015) used resynthesized stimuli to help Japanese learners of Italian. The suprasegmental features of native Italian speakers' speech were transferred to Japanese learners' speech. Then the Japanese learners were asked to imitate their own modified speech. Results showed that through self-imitation, the learners significantly improved their communicative effectiveness. A similar method of using students' prosodically corrected speech as training stimuli employed by Bissiri and Pfitzinger (2009) to teach Italian speakers learning German lexical stress and by Hirose (2004) to teach non-Japanese learners the production of Japanese accents, also yielded positive results.

Modified stimuli allow instructors to make the critical acoustic cues more prominent so as to draw learners' awareness to those critical cues (Hardison, 2012). However, most of those above-mentioned studies modified the learners' speech by directly transferring the native speaker's prosody to the learners' speech. This would make the learners' speech sound unnatural, and even sometimes the learners could not recognize the modified speech as their own voices. Furthermore, it would also reduce the salience of the prosody cues. In this sense, the modified speech might lose its effectiveness. Therefore, in the present study, the researchers modified the students' intonation by manually manipulating the pitch contour of the students' speech (the detailed procedures will be presented in the methodology part). In this way, it was possible to locate precisely the students' pronunciation problems while preserving the students' original pronunciation as a reference point. By so doing, the individual student's specific problems could be solved precisely and the effectiveness of resynthesized stimuli could be given full play.

METHODOLOGY

The present study employed a pretest-treatment-posttest quasi-experimental design to examine the effectiveness of using students' own resynthesized self-produced stimuli for English intonation learning. The experiment included two groups of students—the experimental group (EG) and the control group (CG). The treatment consisted of an English intonation training wherein the students in the EG were trained by using their resynthesized self-produced stimuli (their own voices) as the model for learning, while for the CG, the training stimuli were produced by an English native speaker. Otherwise, conditions, including time on task, were identical. An English intonation test was administered to the participants before and after the training to examine the effectiveness of the two different models for intonation learning. The research questions of this study were:

- 1) What are Chinese EFL university students' problems in English intonation?
- 2) Does exposure to resynthesized self-produced stimuli facilitate the learning of English intonation? If yes, in what ways?

PARTICIPANTS

The participants were sampled from the first-year English majors studying at Hunan University of Science & Engineering (HUSE), China. In HUSE, there were altogether 204 first year English majors assigned to 6 academic classes (with about 34 students in each class). Students of two intact classes with the same teacher of English phonetics course were selected as the participants in order to control the confounding variables that might arise. One of the classes was randomly selected as the EG and the other as the CG. Students who were unwilling to participate in the study or had hearing problems were excluded. Finally, a total of 66 students (33 in each group) were recruited in the experiment.

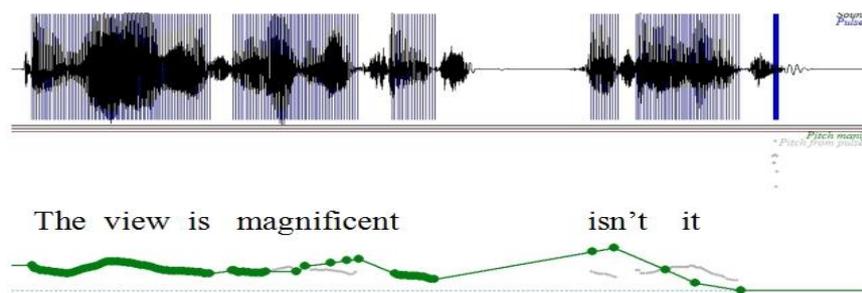
INSTRUMENTS AND PROCEDURES

ENGLISH INTONATION TRAINING COURSEWARE

The students were trained using an English intonation training courseware developed by the researchers with Lectora Inspire (version 17). The design of the intonation learning materials followed the British School's approach to intonation analysis and thus the contents for each training session were divided into three modules—tonality, tonicity and tone (See Appendix A for a sample). The learning materials altogether included 150 target utterances (25 utterances * 6 sessions), which were specially designed in particular contexts for practicing intonation patterns.

The speech stimuli (pronunciation model) for the CG were produced by a native English speaker. Before the training, the native speaker was asked to produce the target utterances and her productions were audio recorded, edited, and then uploaded to the courseware for the CG. The speech stimuli for the EG were initially produced by the students in the EG themselves. Before each training session, each student in the EG was asked to produce the target utterances and his/her productions were audio recorded, resynthesized by the researchers (modifying the incorrect intonations to the correct ones), and then uploaded to the courseware. Therefore, each student in the EG used his/her resynthesized self-produced stimuli as the pronunciation model for learning.

The modification of the students' initial speech productions was performed by using the phonetic software Praat (Boersma & Weenink, 2018). It included four major steps: First, import the recording of student's production into Praat and generate a manipulation object; second, manipulate the pitch contour to the desired one (see Figure 1 for an example); third, adjust the intensity or duration on necessary places to make it sound natural, especially for modifying sentence stress and intonation boundaries; four, generate the resynthesized speech.



Note: The student's initial production had two problems: 1) misplacing the word stress of "magnificent" on the first syllable; 2) inaccurate phonetic realizations of the falling tone of "isn't it" (This utterance was designed in the context where the tag-question "isn't it" should be produced with a falling tone). Therefore, modifications were performed to shift the stress of "magnificent" from the first syllable to the target second syllable, and to shift the starting of the falling tone from "it" to "isn't".

FIGURE 1. An example of modifying student's initial production

When modifying the student's speech, the native speaker's pitch contour was taken as a reference. The modification aimed to: 1) make the intonation phonologically and phonetically correct; 2) make the resynthesized speech sound natural and comfortably intelligible. In addition, the modification focused only on the students' incorrect intonation. If the intonation of the students' initial production was correct, no modifications would be made, and only the original version would be provided to the students.

PROCEDURES FOR CONDUCTING THE ENGLISH INTONATION TRAINING

The training included 6 sessions and each session lasted for 2 hours. Considering that modifying the students' initial speech productions was quite time-consuming, a one-week interval was placed between two consecutive sessions. Therefore, the 6 training sessions altogether spanned 12 weeks. The training was conducted in computer laboratories, and the EG and CG took the training at the same time but in different rooms. To take the training on the intonation training courseware, the students were required to do as follows:

Step 1: Run the courseware, test the microphones, and adjust the sound volume to a comfortable level.

Step 2: Log into the system with student name and ID number and click the "Start" button to start the training.

Step 3: Practice pronunciation through simple listen-compare-repeat exercises. The students can first try to produce the target utterances by themselves using the recorder on the page which can record and replay their productions. Then, they can listen to the pronunciation model and compare it with their own productions. Specifically, for students in the CG, they can click the underlined sentences to listen to the native speaker model (see Figure 2); for students in the EG, they can listen to their original productions and the modified version (see Figure 3). After a certain number of listen, compare, and repeat sequences, they were asked to upload a recording of each target utterance. Their uploaded recordings were sent to the researchers' server, and the researchers could monitor their progress by reviewing these productions.

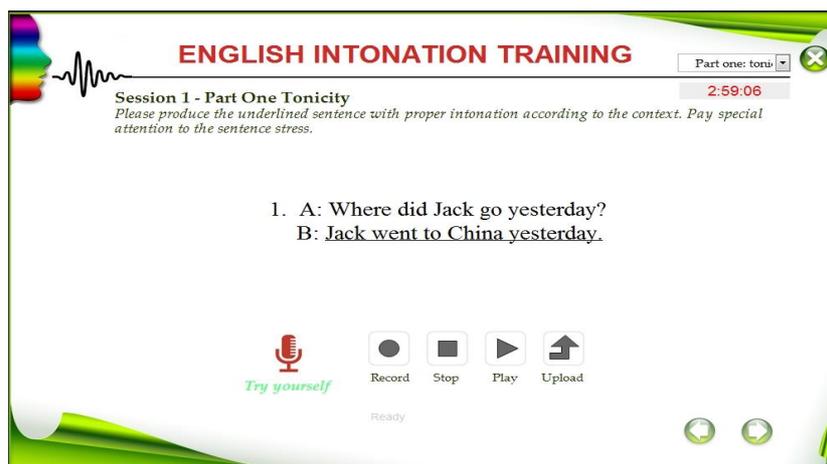


FIGURE 2. Screenshot of the courseware page for the control group

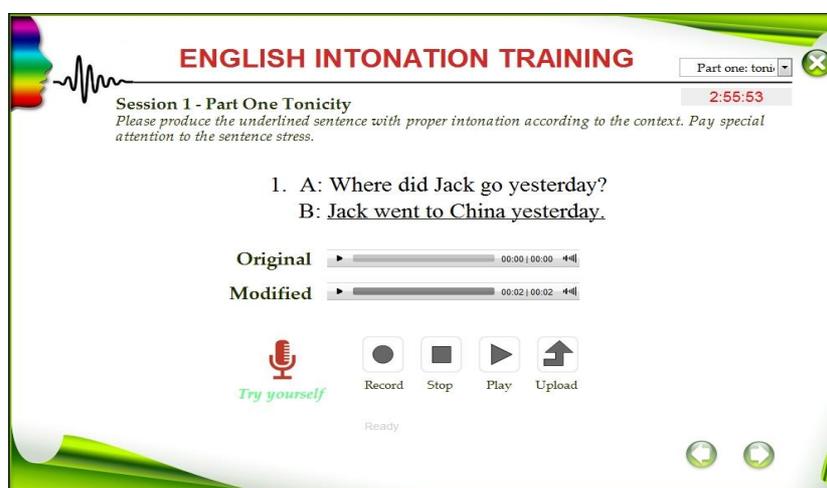


FIGURE 3. Screenshot of the courseware page for the experimental group

ENGLISH INTONATION PRODUCTION TEST

An English intonation production test was used to test the students' performance in intonation production before and after the training. The pretest and the posttest were identical, but none of the testing items were used in the training. The test (Appendix B) was composed of 38 target utterances which were specially designed in contexts. The students were asked to produce the target utterances with proper intonation according to the contexts.

The pretest and posttest were conducted in a quiet room. The students took the test one by one in order to avoid disturbing each other. They were required to produce the target utterances shown on a screen, and their voices were audio recorded and saved in WAV files for further rating.

PRONUNCIATION RATING

The students' performance in the test was assessed in two phases: 1) the rating for intonation choice, i.e., whether the student chose the correct intonation pattern for the target utterance; 2) the rating for the phonetic realizations of the intonation pattern, i.e., how well did the student realize the chosen intonation pattern in his/her pronunciation. The first phase was scored dichotomously: "0" for incorrect intonation choice, and "5" for correct intonation choice. The

rating for phonetic realization focused on the degree to which the student’s intonation deviated from the native speaker’s intonation. This concept was proposed by Kennedy and Trofimovich (2008): “how closely the pronunciation of an utterance approaches that of a native speaker” (p. 461), and by Isaacs and Thomson (2013): “how different the speaker sounds from a NS” (p. 141). The rating employed a 5-point Likert scale, where “5” represented “near native speaker’s intonation” and “1” represented “extremely different from the native speaker’s intonation”

Therefore, for the 38 items of the intonation production test, the total score was 380, with 10 marks for each item. Of the 10 marks, 5 marks were for intonation choice and the other 5 marks were for the phonetic realizations of the chosen intonation. The rating employed a double-blind procedure, i.e., the student’s productions in the pretest and posttest were mixed together and then were provided to the raters. In order to make the rating more reliable, the three raters received a training session and participated in a pilot rating. After the rating, Pearson’s Correlation Coefficient was used to calculate and check the inter-rater reliability. The results (Table 1) showed that there were strong positive correlations ($r > 0.80$) in scoring across the three raters, indicating a high level of inter-rater reliability.

TABLE 1. Inter rater reliability results (Pearson’s Correlation Coefficient)

	Rater 1	Rater 3
Rater 1	-	0.88
Rater 2	0.83	0.80

RESULTS

Descriptive statistics for the students’ scores in the pretest are shown in Table 2. As can be seen from the table, the mean scores of both groups were quite low, approximating 56% of the total score ($380 * 56\% = 212.8$). None of the students’ scores were higher than 300 (about 79% of the total score). The performance of the CG (S.D.=31.80) was more consistent than that of the EG (S.D.=36.41). However, results of the Independent-samples t-test showed that there was no significant difference between the two groups’ mean scores ($t=0.41$, $p=0.68$), i.e., before the training, the EG and the CG’s performance in producing English intonation were at about the same level.

TABLE 2. Descriptive statistics for the students’ pretest scores

Group	Number	Mean	SD	Range	Min.	Max.
EG	33	213.48	36.41	131	157	288
CG	33	210.00	31.80	136	156	292

Table 3 lists the two groups’ detailed performances for specific intonation patterns as well as the results of independent-samples t-tests for comparing the two groups’ mean scores in those patterns. As the results revealed, there were no significant differences between the two groups’ performances across all of the intonation patterns as none of the p values were lower than 0.05. In terms of tonicity and tonality, their mean scores were around half of the total score ($60 * 50\% = 30$). The pretest performances of both groups in Statements (falling tone), Commands, and Exclamations were excellent, with mean scores over 18 (90% of the total score), while their performances in Wh-questions (falling tone), Yes/no-questions (falling tone), Tag-questions (falling tone), Implications, Alternative questions, and Listings were very poor, with mean scores lower than 8 (40% of the total score).

TABLE 3. The students' pretest scores in specific intonation patterns

Intonation pattern	EG		CG		Indept-S. t-test	
	Mean	SD	Mean	SD	t	Sig.(2-tailed)
Tonicity	33.00	9.27	32.03	7.58	0.47	0.64
Tonality	26.24	8.22	24.24	8.27	0.99	0.32
Statement (F)	18.67	2.03	18.27	2.92	0.64	0.53
Statement (R)	13.67	5.76	13.42	5.79	0.17	0.87
Wh-question (F)	13.45	6.62	14.70	5.99	-0.80	0.43
Wh-question (R)	8.97	6.27	9.55	6.67	-0.36	0.72
Yes/no-question (F)	5.39	7.83	5.91	6.61	-0.29	0.77
Yes/no-question (R)	12.82	5.99	11.79	6.16	0.69	0.49
Tag-question (F)	8.79	6.96	9.21	6.96	-0.25	0.81
Tag-question (R)	13.91	6.95	13.27	6.86	0.37	0.71
Command	18.94	1.14	19.21	1.17	-0.96	0.34
Exclamation	19.00	1.23	19.06	1.14	-0.21	0.84
Implication	3.03	4.75	3.36	5.44	-0.27	0.79
Alternative question	8.79	6.55	8.42	6.05	0.23	0.82
Listing	8.82	6.20	7.55	5.33	0.90	0.37

Note:

1. "F" in brackets stands for "Falling tone" and "R" stands for "rising tone". E.g., "Statement (F)" means the case of producing a statement with falling tone.
2. "Indept-S. t-test" is the abbreviation for "Independent-samples t-test".

Results of the students' performance in the posttest are presented in Table 4. The results showed that both groups' mean scores were higher than 72% (273) of the total score. The lowest score was 223, which was still higher than the mean score of the pretest (212). Comparison between the two groups' mean scores was conducted via an Independent-samples t-test, and the results showed that there was a significant difference ($t=3.91$, $p=0.00<0.05$) between the performance of the EG ($M=300.09$) and the CG ($M=275.52$). In other words, the EG outperformed the CG in the posttest.

TABLE 4. Descriptive statistics for the students' posttest scores

Group	Number	Mean	SD	Range	Min.	Max.
EG	33	300.09	24.07	102	247	349
CG	33	275.52	26.92	106	223	329

Paired-samples t-tests were employed to compare the students' pretest and posttest scores. Results (Table 5) showed that there were significant differences between the students' pretest scores and posttest scores, both for the EG ($t=-17.51$, $p=0.00<0.05$) and the CG ($t=-14.13$, $p=0.00<0.05$). This means that the performance of both groups in the intonation test had significantly improved after the training.

TABLE 5. Comparisons between the students' pretest and posttest scores

Pair (pretest-posttest)	Mean differences	SD	95% Confidence Interval of the Difference		t	Sig. (2-tailed)
			Lower	Upper		
EG	-86.61	28.41	-96.68	-76.53	-17.51	0.00
CG	-65.52	26.64	-74.96	-56.07	-14.13	0.00

In terms of the students' detailed performances in specific intonation patterns, results of the Paired-samples t-tests (Table 6) indicated that for both groups, their performances were significantly improved across all patterns (all p values were smaller than 0.05) except for Statements (falling tone), Commands, and Exclamations. The reason is that they had already achieved high scores for these three cases in the pretest, resulting in a ceiling effect.

TABLE 6. Comparisons of the students' pretest and posttest scores in specific intonation patterns

Pair: pretest - posttest	EG			CG		
	Mean differences	t	Sig. (2-tailed)	Mean differences	t	Sig. (2-tailed)
Tonicity	-14.33	-10.97	0.00	-11.36	-8.28	0.00
Tonality	-14.52	-10.68	0.00	-11.58	-10.11	0.00
Statement (F)	-0.49	-1.15	0.26	-0.67	-1.01	0.32
Statement (R)	-4.21	-4.27	0.00	-2.55	-2.60	0.01
Wh-question (F)	-2.67	-2.56	0.02	-2.39	-2.7	0.01
Wh-question (R)	-6.21	-5.73	0.00	-4.46	-3.75	0.00
Yes/no-question (F)	-10.09	-6.98	0.00	-8.97	-7.24	0.00
Yes/no-question (R)	-4.15	-3.29	0.00	-3.18	-2.85	0.01
Tag-question (F)	-6.52	-5.57	0.00	-3.55	-3.20	0.00
Tag-question (R)	-3.58	-3.06	0.00	-3.42	-2.94	0.01
Command	-0.49	-1.97	0.06	-0.15	-0.65	0.52
Exclamation	-0.27	-1.06	0.30	-0.09	-0.39	0.70
Implication	-7.79	-8.35	0.00	-5.18	-5.90	0.00
Alternative question	-6.97	-5.82	0.00	-4.82	-3.64	0.00
Listing	-4.33	-4.02	0.00	-3.15	-4.44	0.00

As has been mentioned above, the EG outperformed the CG in the posttest in terms of the total score. However, in terms of their detailed performances in specific intonation patterns, results of the independent-samples t-tests showed that the two groups' scores in Tonality, Wh-questions (falling tone), Yes/no-questions (falling tone), and Tag-questions (rising tone) showed no significant differences, while the differences in Tonicity, Statement (rising tone), Wh-question (rising tone), Yes/no-questions (rising tone), Tag-questions (falling tone), Implications, Alternative questions, and Listings, reached statistical significance ($p < 0.05$).

As the rating for the students' pronunciation was composed of two phases, it is necessary to examine the students' scores in each phase of the rating, so as to locate which phase of the rating caused the significant differences in the two groups' scores in the above-mentioned 8 intonation patterns. The first phase was the rating for intonation choice and the second phase involved the rating for the phonetic realizations of the chosen intonation. Table 7 displays the results of the independent-samples t-tests for comparing the two groups' scores in the first phase of rating, and Table 8 shows the results of the second phase of rating.

TABLE 7. Comparisons of the two groups' scores in the first phase of rating

Intonation pattern	EG (Mean)	CG (Mean)	Independent-samples t-test		
			Mean difference	t	Sig. (2-tailed)
Tonicity	25.30	24.24	1.06	1.15	0.26
Statement (R)	9.24	8.79	0.45	0.92	0.36
Wh-question (R)	7.88	7.58	0.30	0.49	0.63
Yes/no-question (R)	8.94	8.64	0.30	0.57	0.26
Tag-question (F)	8.03	7.42	0.61	0.98	0.33
Implication	5.76	5.30	0.46	0.69	0.49
Alternative question	8.33	7.73	0.61	1.00	0.32
Listing	6.97	6.36	0.61	1.04	0.30

TABLE 8. Comparisons of the two groups' scores in the second phase of rating

Intonation pattern	EG (Mean)	CG (Mean)	Independent-samples t-test		
			Mean difference	t	Sig. (2-tailed)
Tonicity	24.24	19.15	2.88	3.32	0.00
Statement (R)	8.64	7.18	1.45	3.05	0.00
Wh-question (R)	7.30	5.94	1.36	2.24	0.03
Yes/no-question (R)	8.03	6.33	1.70	3.45	0.00
Tag-question (F)	7.27	5.33	1.94	3.35	0.00
Implication	5.06	3.24	1.82	3.74	0.00
Alternative question	7.42	5.52	1.91	3.52	0.00
Listing	6.18	4.33	1.85	3.37	0.00

As has been demonstrated above, for the 8 intonation patterns shown in table 7 and table 8, there were significant differences between the two groups' performances. However, if one takes the scores separately, in terms of intonation choice (the first phase of rating), there was no significant difference between the two groups' performance across all of the 8 patterns (Table 7); in terms of phonetic realizations (the second phase of rating), there was a significant difference in each of the 8 patterns (Table 8). In other words, the EG and the CG had equal performances in choosing intonation patterns, while in phonetic realizations of those patterns, the EG outperformed the CG. This means that it was the aspect of phonetic realization that caused the significant differences between the two group' performances. They both improved equally in ability to perceive intonative differences while the experimental group was better at production than the control group which remained unchanged.

To sum up the results, there was no significant difference between the two groups' performance in the pretest. In the posttest, both groups' performances were significantly improved across all of the intonation patterns except for Statements (falling tone), Commands, and Exclamations due to a ceiling effect. In terms of the total score, the EG outperformed the CG; in terms of their scores in specific intonation patterns, the two groups showed no significant differences in Tonality, Wh-questions (falling tone), Yes/no-questions (falling tone), and Tag-questions (rising tone), while for the other 8 patterns, the differences reached significant levels. Comparisons of their scores in the two rating phases for the 8 intonation patterns indicated that the differences did not lie in intonation choice but in the phonetic realizations of those patterns. Here, conclusions can be drawn that both the native speaker model (for the CG) and the students' resynthesized self-produced stimuli as the model (for the EG) for intonation learning were equally effective for informing the students to choose correct intonation patterns for production. However, the students' resynthesized self-produced stimuli as the model for learning were more effective than the native speaker model for facilitating the students' phonetic realizations of intonation.

DISCUSSION

CHINESE STUDENTS' PROBLEMS IN ENGLISH INTONATION

The results of the pretest revealed that Chinese students' overall performances in English intonation were far from satisfactory before the training. However, in terms of specific intonation patterns, their performance in Statements (falling tone), Commands and Exclamations were satisfactory. This was in accordance with previous research finding (Jiang, 2012; Huo & Luo, 2017; Rui, 2007) that Chinese students were good at producing falling tones, and Wells' (2006) claim that exclamations are the simplest kinds of utterance for EFL students. The reason might be that the falling tone is the most frequently used tone pattern across all languages, while the default tone for these three sentence types is a fall.

For the other intonation patterns, some were significantly improved after the training and some were more resistant to change. Especially, judging from the CG's performance before and after the training, for some intonation patterns, the students could produce them correctly as long as they made the correct intonation choice. These patterns were in: Tonality, Wh-questions (falling tone), Yes/no-questions (falling tone), and Tag-questions (rising tone). For other patterns, students' problems lay in phonetic realization, i.e., even if they made the correct intonation choices, they still could not produce them accurately. These patterns were in: Tonicity, Statements (rising tone), Wh-questions (rising tone), Yes/no-questions (rising tone), Tag-questions (falling tone), Implications, Alternative questions, and Listings.

Intonation consists of a phonological and a phonetic component, and learners' phonological problems with L2 intonation may result from the intonational differences in the

inventory of L1 and L2 phonological patterns (Mennen, 2007). This could account for the Chinese students' problems in intonation choice. Because of the differences between Chinese and English language, such as the former being a tonal language with stress-timed rhythm patterns while the latter is a non-tonal language with syllable-timed rhythm patterns, Chinese students tended to choose intonation arbitrarily or according to their first language intonation when speaking English. This was represented by overusing falling tones, over-reliance on pauses to mark intonation boundaries, and failure to highlight information focuses. As English intonation has rarely been taught in the Chinese EFL classroom, the limited instruction focused mainly on the teaching of default tones. A default tone is an "unmarked or neutral tone that is used under no special circumstances" (Wells, 2006, p.15), e.g., the default tone for statement is a fall. This misled Chinese student into linking default tones with sentence types without considering contexts. As a result, they were prone to making incorrect intonation choices in Statements (rising tone), Wh-questions (rising tone), Yes/no-questions (falling tone), and Tag-questions (falling tone).

Chinese students failed to produce correctly the above-mentioned eight intonation patterns, even though they could make the correct intonation choices. This implied that English intonation was learned by Chinese students sequentially. They might first form the phonological patterns of English intonation and then acquire the phonetic details, which is in accordance with Mennen's (2007) claim of a two-stage process for L2 intonation learning. Specifically, Chinese students' problems in phonetic realization of English intonation were: for falling tone pattern, on the one hand, some Chinese students' speech tended to sound flat or monotonous, while on the other hand, some students purposefully added pitch variations that were unnecessary, making their speech sound strange. For rising tone, they tended to place the starting point of the rise at the last word of a sentence. For falling-rising, they usually unconsciously replaced it with a rising tone. Compound tones were found to be more difficult for Chinese students to produce. For one thing, compound tones were usually incorporated in long sentences (e.g., alternative questions) which may contain more words or novel words that might result in the students' anxiety during production. For another thing, compound tones were composed of more than two tone patterns which involved more pitch variations and might increase the students' phonological memory loads.

THE EFFECTIVENESS OF EXPOSURE TO RESYNTHESIZED SELF-PRODUCED STIMULI FOR INTONATION LEARNING

Well-formed phonological representation is a sine qua non for target-like sensory motor skills and accurate L2 speech production (Lee & Lyster, 2017). Therefore, the key factor of L2 phonological acquisition lies in whether the input speech signal can arouse in learners an awareness of the perceived differences between his/her production and the target sound (Flege, 1995) that will lead to an adjustment of the representations (Leather, 1983). Adequate adjustment will lead to the establishment of new categorizations and the redundancy of input, while inadequate adjustment will result in production failure. This might explain why, in the present study, some students could correctly choose the target intonation patterns but not produce them correctly.

The native speaker model was effective in helping the students form the initial phonological representations of the intonation patterns while it failed to facilitate the students' subsequent adjustment of those representations. Brown (1999) pointed out that native speaker models of the same category are continually faced with variable realizations influenced by factors like coarticulation, sloppy articulation or interspeaker variability. Those variations do not contribute to differences in meaning while distracting learners' attention from critical cues. In comparison, the students' resynthesized self-produced stimuli as the model can filter out the irrelevant "noise" in the acoustic signal and make the critical cues more salient (Hardison,

2012). By listening to the modified speech of the students' own voices, the students could focus on the modifications (the key acoustic cues) and realize faster the differences between their incorrect intonation and the desired target without "the distractions from less relevant factors such as voice characteristics" (Tang, Wang & Seneff, 2001, p.3). Hence, the memory load put on the auditory system was greatly reduced and the processing could proceed more quickly.

Therefore, the students' resynthesized self-produced stimuli as the model for intonation learning can make the key acoustic cues more salient, facilitate the students to listen to the model critically, and improve the students' ability to produce more accurate phonetic realizations of the intonation being studied. Specifically, its effectiveness was manifested in the following aspects:

Firstly, it helped students to deaccentuate unnecessary information focuses. Influenced by the syllable-timed rhythm pattern of Chinese language, Chinese students tended to accent every word in a sentence equally, making their speech lack in prominence contrasts. It was relatively easier for the students to notice the information focus of an utterance, provided that they devoted attention to the contexts. In the present study, the students in the CG were able to perceive the information focus after listening to the native speaker model. However, in phonetic realizations of the information focus, those students would unconsciously add more prominence than necessary within one intonation group. This made the rhythm of their speech sound syllable-timed, or Chinese-accented. In comparison, for the students in the EG, after listening to their modified speech, not only did they learn to locate the information focus, but they also gained the ability to remove the unnecessary stresses in their original productions.

Secondly, it provided students with more strategies for chunking and increased their speech fluency. In the present study, by listening to the native speaker model, the students in the control group noticed the intonation boundaries in the native speaker's speech, but they failed to extract the strategies used for signaling the boundaries. For example, if there was a pitch reset between two intonation groups, the students would tend to perceive it as a pause rather than notice the difference of the pitch variations between the two intonation groups. In comparison, the students in the experimental group gained more strategies (such as pitch variations) through the training to mark intonation boundaries, which contributed to their speech fluency.

Thirdly, it enabled students to retune their phonological representations of tone patterns and facilitated the students' more accurate phonetic realizations of these patterns. For falling tones, it helped the students to eliminate unnecessary pitch variations in their speech. For rising tones, it helped the students to locate more accurately the starting point of the rise. In cases where the pitch variations were more complex, such as falling-rising tones or compound tones, it enhanced the students' phonological memory and reduced their anxiety in producing those tone patterns. This was because by listening to this kind of model, the students could focus on the problems of their pronunciation without distractions from other elements with which they had no problems. In this sense, it provided a more realistic pronunciation goal for the students to pursue and the students could gradually approach the desired pronunciation based on the corrections of their mispronunciations (a constructive way).

The CG students' failure in listening to the native-speaker model critically brought the model-oriented intervention into a "bootstraps fallacy" (Leather, 1983) — the phonetic criteria which the students needed for helpfully critical listening were precisely those which the instruction aimed to develop. The model-oriented intervention focuses on how the model is correct, rather than how the students can be correct or how students' problems can be solved. It fails to locate students' individual specific problems and attempts to solve students' problems through a repeated trial-and-error procedure. In comparison, the precision approach to pronunciation correction, informed by the concept of precision language education, focuses on what individual student's problems are, what interventions can be tailored to their needs, and

how precisely their problems can be solved. Especially, this approach fits the needs of the Chinese EFL pronunciation classroom characterized by large size, teacher-centeredness and “one-size-fits all” teaching. Instruction of this kind of precision approach can avoid educational waste, save time in solving students’ problems, meet individual student’s needs, and make learning less costly and more effective.

CONCLUSION

Intonation consists of a phonological and a phonetic component. This study found that Chinese students’ problems in English intonation reflected the two-stage process of L2 intonation learning proposed by Mennen (2007), i.e. that L2 learners may first acquire the phonological intonation patterns before they acquire the correct phonetic implementation of these patterns. To redress the students’ problems, this study conducted an English intonation training with two groups of students to examine the effectiveness of two types of model for learning: the native-speaker model and the students’ resynthesized self-produced stimuli as the model. The results showed that exposure to resynthesized self-produced stimuli for intonation learning was as effective as the native-speaker model in informing students of the phonological representations of intonation patterns, while it was much more effective than the native-speaker model in enabling students to produce more accurate phonetic realizations of those patterns. Thus, it is clear that resynthesized self-produced stimuli are capable of making critical acoustic cues more salient, can reduce students’ phonological memory load, can enhance students’ critical listening and, as a consequence, can enable students to focus on the phonetic details of intonation patterns.

However, in view of the technological limitations, the resynthesized stimuli provided to the students could be thought of as a kind of delayed rather than immediate feedback. By the time the students received the training, there is a (very unlikely) possibility that their problems might have been solved or become transformed into other problems. Furthermore, when the students were doing the exercises or the tests, they were provided with texts (scripts), so this study cannot answer the question of whether the students’ ability gained from the training can be generalized to real-time conversations or whether the correct patterns will be retained over a long period of time. The primary purpose of this study was to determine whether the kind of input modification used in the study was of significant value, and that appears to be the case. Nevertheless, in order to deal with some of the issues raised in this paragraph, the researchers in this study recommend that a real-time pitch modifier be developed to provide immediate corrective feedback to students’ intonation and that further studies be conducted with a larger sample size and longer training time to examine the generalization and retention effects of the training.

The researchers also recommend that neuroimaging experiments be carried out to examine the differences in the learner’s neuronal activities in the cortical area upon listening to resynthesized self-produced speech stimuli and native-speaker models. Hopefully, this kind of experiment can provide neuro-evidence that exposure to resynthesized self-produced stimuli can contribute to the learner’s acoustic-articulatory mapping. In addition, future research can use students’ modified speech for the instruction of segmentals. It is suggested that one begin with vowels whose modifications can be realized by manipulating the formants or only the second formant for some vowel contrasts. Considering that language learners’ problems in individual sounds are more complex and individual-specific, one can look forward to the promising outcomes achieved here yielding positive results.

REFERENCES

- Anderson-Hsieh, J., Johnson, R. & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*. 42(4), 529-555.
- Beckman, M. E. & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology*. 3, 255-309.
- Bi, R. & Chen, H. (2013). Developmental changes of tone patterns in Chinese EFL students' read speech. *Foreign Languages and Their Teaching*. 1, 50-54.
- Bissiri, M. P. & Pfitzinger, H. R. (2009). Italian speakers learn lexical stress of German morphologically complex words. *Speech Communication*. 51(10), 933-947.
- Boersma, P. & Weenink, D. (2018). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.37. Retrieved from <http://www.praat.org/>
- Botinis, A., Granström, B. & Möbius, B. (2001). Developments and paradigms in intonation research. *Speech Communication*. 33(4), 263-296.
- Brown, C. A. (1999). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 4-63). Malden, MA: Blackwell.
- Cook, C. R., Kilgus, S. P. & Burns, M. K. (2018). Advancing the science and practice of precision education to enhance student outcomes. *Journal of School Psychology*. 66, 4-10.
- Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.
- Dalton, C. & Seidlhofer, B. (1994). *Pronunciation*. Oxford: Oxford University Press.
- Felps, D., Bortfeld, H. & Gutierrez-Osuna, R. (2009). Foreign accent conversion in computer assisted pronunciation training. *Speech Communication*. 51(10), 920-932.
- Fuchs, D., Mock, D., Morgan, P. L. & Young, C. L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*. 18(3), 157-171.
- Gilbert, J. (2014). Myth 4: intonation is hard to teach. In Grant, L. & Brinton, D. (Eds.), *Pronunciation Myths: Applying Second Language Research to Classroom Teaching* (pp. 107-137). Michigan: University of Michigan Press.
- Halliday, M. A. K. (2015). *Intonation and Grammar in British English (Vol. 48)*. Walter de Gruyter GmbH & Co KG.
- Hardison, D. M. (2012). Second language speech perception: A cross-disciplinary perspective on challenges and accomplishments. In S. Gass & A. Mackey (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp. 349-363). London: Routledge.
- Hart, S. A. (2016). Precision education initiative: moving toward personalized education. *Mind, Brain, and Education*. 10(4), 209-211.
- Hirose, K. (2004). Accent type recognition of Japanese using perceived mora pitch values and its use for pronunciation training system. In *Proceedings of the International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages* (pp. 77-80), Beijing, China.
- Huo, S. Y. & Luo, Q. (2017). Misuses of English Intonation for Chinese Students in Cross-Cultural Communication. *Cross-Cultural Communication*. 13(1), 47-52.
- Isaacs, T. & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*. 10(2), 135-159.
- Jenkins, J. (2004). Research in teaching pronunciation and intonation. *Annual Review of Applied Linguistics*. 24, 109-125.
- Jiang, H. L. (2012). An empirical study of the learning effect of college English majors in

- intonation. *Foreign Languages in China*. 9(2), 65-80.
- Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent*. Doctoral dissertation, University of Stuttgart, Germany.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*. 38(2), 301-315.
- Kennedy, S. & Trofimovich, P. (2008). Intelligibility, comprehensibility and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*. 64, 459-489.
- Leather, J. (1983). Second-language pronunciation learning and teaching. *Language Teaching*. 16(3), 198-219.
- Lee, A. H. & Lyster, R. (2017). Can corrective feedback on second language speech perception errors affect production accuracy?. *Applied Psycholinguistics*. 38, 1-23.
- Lengeris, A. (2012). Prosody and second language teaching: Lessons from L2 speech perception and production research. *Pragmatics and Prosody in English Language Teaching*. 15, 25-40.
- Lian, A. P. (2014). On-Demand Generation of Individualized Language Learning Lessons. *Journal of Science*. 9(1), 25-38.
- Lian, A. P. & Sangarun, P. (2017). Precision Language Education: A Glimpse Into a Possible Future. *GEMA Online® Journal of Language Studies*. 17(4), 1-15.
- Lively, S. E., Logan, J. S. & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*. 94(3 Pt 1), 1242-1255.
- Logan, J. S., Lively, S. E. & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*. 89(2), 874-886.
- Lu, J., Wang, R. & De Silva, L. C. (2012). Automatic stress exaggeration by prosody modification to assist language learners perceive sentence stress. *International Journal of Speech Technology*. 15(2), 87-98.
- MacDonald, D. (2011). Second language acquisition of English question intonation by Koreans. *In Proceedings of the 2011 annual conference of the Canadian Linguistic association*.
- Makarova, V. & Zhou, X. (2006). Prosodic characteristics in the Speech of Chinese EFL learners. *Proceedings of Speech Prosody*. Dresden, Germany.
- Mennen, I. (1999). The realisation of nucleus placement in second language intonation. In *Proceedings of the fourteenth international congress of phonetic sciences* (pp. 555-558).
- Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*. 32(4), 543-563.
- Mennen, I. (2007). Phonological and phonetic influences in non-native intonation. *Trends in Linguistics Studies and Monographs*. 186, 53-76.
- Meng, X. J. & Wang, H. M. (2009). Boundary tone patterns in Chinese English learners' read speech. *Foreign Language Teaching and Research (Bimonthly)*. 6, 447-451.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *Tesol Quarterly*. 25(3), 481-520.
- Pellegrino, E., Vigliano, D. (2015). Self-imitation in prosody training: A study on Japanese learners of Italian. In Steidl, S., Batliner, A. & Jokisch, O. (Eds.), *Workshop on Speech and Language Technology in Education* (pp. 53-57), Leipzig, Germany.
- Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *Tesol Quarterly*. 35(2), 233-255.
- Roach, P. (2009). *English phonetics and Phonology: A Practical Course*. Cambridge: Cambridge University Press.
- Romero-Trillo, J. (2012). *Pragmatics, Prosody and English Language Teaching*. Dordrecht:

Springer.

- Rui, T. (2007). The English intonation of Chinese EFL learners: A comparative study. *CELEA Journal (Bimonthly)*. 12(6), 34-45.
- Swerts, M. & Zerbian, S. (2010). Intonational differences between L1 and L2 English in South Africa. *Phonetica*. 67(3), 127-146.
- Tang, M., C. Wang & S. Seneff, (2001). Voice transformations: from speech synthesis to mammalian vocalizations. In *Proceedings of the 7th European Conference on Speech Communication and Technology* (pp. 357-360), Aalborg, Denmark.
- Tench, P. (2015). *The Intonation Systems of English*. Bloomsbury Publishing.
- Thomson, R. (2017). Measurement of accentedness, intelligibility, and comprehensibility. In Kang, O. & Ginther, A. (Eds.), *Assessment In Second Language Pronunciation* (pp. 11-29). London: Routledge.
- Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge: Cambridge University Press.
- Willems, N. (2010). *English Intonation from a Dutch Point of View*. Walter de Gruyter.
- Yang, J. (2006). Inappropriate divisions of intonation groups in Chinese university students' read speech. *Modern Foreign Languages (Quarterly)*. 29(4), 409-417.
- Yang, M., & Mu, F. Y. (2011). Study on fluency-related IP-internal pauses in Chinese EFL learners' spontaneous speech. *Foreign Languages and Their Teaching*. 6, 16-21.
- Yoon, K. C. (2009). Synthesis and evaluation of prosodically exaggerated utterances. *Phonetics and Speech Sciences*. 1(3), 73-85.
- Zhang, L. (2015). An empirical study on the intelligibility of English Spoken by Chinese university students. *Chinese Journal of Applied Linguistics*. 38(1), 36-54.
- Zhu, L. (2007). *The rhythm patterns of English spoken by Chinese and its pedagogical implications*. Doctoral dissertation, Minzu University of China, China.

APPENDIX A

A SAMPLE OF THE ENGLISH INTONATION TRAINING MATERIALS

TRAINING SESSION ONE

Module I Tonicity: the placement of information focus

Instruction: Please read the following dialogues and try to produce the underlined sentences with proper intonation according to the context. Then listen to the models and compare with your own productions. Pay special attention to the placement of information focus.

Part 1: Narrow focus

(1) A: Where did Jack go yesterday?

B: Jack went to China yesterday.

(2) A: When did Jack go to China?

B: Jack went to China yesterday.

Part 2: Contrastive focus

(1) A: This donation is for him?

B: From him, not for him.

(2) A: You bought it before Christmas?

B: After Christmas, not before Christmas.

Part 3: Old and new information

(1) A: Can you give me a cigarette?

B: I thought you have quit smoking.

(2) A: What soup do you want?

B: I prefer beef soup.

Module II Tonality: chunking

Instruction: Please read the following dialogues and try to produce the underlined sentences (the punctuation has been removed) with proper intonation according to the context. Then listen to the models and compare the models with your own productions. Pay special attention to the intonation boundaries.

Part 1: Attributive clauses

(1) A: Jane is my sister who lives in Canada

B: Where's your other sister Ella?

(2) A: Who is Jane? Is she your only sister?

B: Yes. Jane is my sister who lives in Canada

Part 2: Adverbials

(1) A: I will talk to the students in the garden

B: OK. I'm going to take them to the garden.

(2) A: I will talk to the students in the garden

B: OK. I'm going to take them to your office.

Part 3: Parallel structures

(1) A: Has she washed the dishes?

B: She washed and ironed her blouse

(2) A: What did she do to her blouse?

B: She washed and ironed her blouse

Module III Tone: the falls and rises in pitch

Instruction: Please produce the underlined sentence with proper intonation according to the context. Pay special attention to the falls and rises in pitch.

(1) Statement: falling tone

A: What are they doing?

B: They are waiting outside.

- (2) Statement: rising tone
A: They are waiting outside?
B: No, they aren't. I think inside.
- (3) Wh-question: falling tone
A: I bought a new phone.
B: How much?
A: 1000 dollars.
- (4) Wh-question: rising tone
A: This new phone cost me 1000 dollars.
B: How much?
- (5) Yes/no-question: falling tone
A: It would be nice to have a new kitchen.
B: Can we afford one? You know we can't even afford the food.
- (6) Yes/no-question: rising tone
A: Will you be at the meeting?
B: I'm not sure now.
- (7) Tag-question: falling tone
Well it's not very good, is it?
(Note: the speaker is sure that the hearer will agree)
- (8) Tag-question: rising tone
It's snowing, isn't it?
(Note: the speaker is not sure.)
- (9) Command: falling tone
A: What should I do next?
B: Add the seasoning.
(Note: the speaker intends to give an normal instruction)
- (10) Exclamation: falling tone
A: I just got a promotion.
B: What good news!
- (11) Implication: falling-rising tone
A: Can we set up an appointment?
B: I could see you on Tuesday. (but that might not suit you)
- (12) Alternative question: falling tone + rising tone
A: Is Mary ready or does she need some more time?
B: She is ready now.
- (13) Listing: falling tone+ falling tone +...+ rising tone
A: What fruits do you like?
B: I like apples, oranges and pears.

APPENDIX B

ENGLISH INTONATION PRODUCTION TEST

Part I Tonicity: the placement of information focus

Instruction: Please produce the underlined sentences according to the context. Pay special attention to the placement of information focus.

- (1) A: Which team is going to win?
B: The red team is going to win.
- (2) A: That mobile looks familiar.
B: It's your phone.
- (3) A: Do you remember what he said?
B: I only care how he said it.
- (4) A: Does she write books?
B: No, but she used to write books.
- (5) A: Shall we walk there?
B: Yes. I like going on foot.
- (6) A: Do you like winter?
B: No, I can't stand cold weather.

Part II Tonality: chunking

Instruction: Please produce the underlined sentences according to the context. The punctuation has been removed. Pay special attention to the boundaries between intonation groups.

- (1) A: The villagers who like running live longer
B: Yes, I can tell. All the people in this village, old or young, like running very much.
- (2) A: The defendant said the accuser should be punished
B: I agree. Obviously, it's the defendant's fault.
- (3) A: She was talking to the man she met on the bus
B: She told me already. They first met at a party.
- (4) A: Those who spoke quickly got an angry response
B: He always requires students to keep quiet in class.
- (5) A: Imported apples and oranges are expensive
B: The price of apples is reasonable. But the oranges are domestic.
- (6) A: Who will clean the table?
B: I'm going to clean and repair the bathroom
A: You don't need to repair the bathroom. Just clean the table.

Part III Tone: the falls and rises in pitch

Instruction: Please produce the underlined sentences with proper intonation according to the context or instructions.

- (1) A: Most left handed people are creative.
B: I agree. I'm left handed and obviously I'm creative.
- (2) A: The sun rises from the east and sets in the west.
B: But to me, I feel it rises from the west and sets in the east.
- (3) A: We're going to have to let you go.
B: You are firing me?
B: Yes. You disappointed all of us.
- (4) A: Anybody home?
B: Oh, Tony. It's you! Come on, in.
- (5) A: Where are the kids staying?
B: They are staying with their grandmother.
- (6) A: How long did it take to get there?
B: Fifty minutes drive.

- (7) A: Ten of my friends will join our dinner.
B: How many? Our table is not that big.
- (8) A: Jenny won a big prize.
B: What did she do?
C: I said she won a big prize. A: Answer me!
- (9) Will you marry me?
B: I will, but please give me some time.
- (10) A: I'm sorry. I really didn't want to hurt her. It was not on purpose.
B: But did you hurt her?
- (11) A: Was she pleased to see you?
B: Yes, sure she was.
- (12) A: Can you speak Chinese?
B: No, I can't. It's so difficult.
- (13) A: Why did I only get a C?
B: Because you made a lot of mistakes, didn't you.
- (14) A: She is pretty smart, isn't she? (Note: The speaker is asking for agreement)
B: Yes, she's always smart.
- (15) A: You are Japanese, aren't you? (Note: The speaker is quite sure)
B: Yes, I am. How did you know?
- (16) A: We have met before, haven't we? (Note: The speaker is not sure)
B: No, I think we haven't.
- (17) A: Move out of my way!
B: Why are you shouting at me!
- (18) A: Stop! I told you many times. Don't feed the dog from the table!
B: Alright. Don't be angry. I'll never do that again.
- (19) A: I just won the lottery!
B: Why are you yelling at me?
A: I'm sorry. I'm just too excited.
- (20) A: He has donated all of his properties.
B: He's such a kind soul!
- (21) A: What time should I come in tomorrow?
B: Can you come in at 3?
A: I can...
B: So what?
A: But should I? The meeting starts at 1.
- (22) A: Will their parents be coming to the dinner?
B: They're invited.
A: But?
B: They refused.
- (23) A: Is something up?
B: Was that a knock at the door, or am I imaging things?
- (24) A: He was very rude, wasn't he?
B: Is he always like that, or had something upset him?
- (25) A: Do you have something to recommend?
B: We have fried chicken, hamburger, French fries...
A: OK. A hamburger please.
- (26) A: We can paint it in red, white, blue...
B: Red and blue.
A: I haven't finished yet. We can also choose brown, purple, and green.

ABOUT THE AUTHORS

Zhongmin Li is a Ph.D candidate in the School of Foreign Languages at Suranaree University of Technology, Thailand, and a lecturer in the School of Foreign Languages at Hunan University of Science and Engineering, China.

Andrew-Peter Lian is Professor of Foreign Language Studies at Suranaree University of Technology, Thailand, Professor of Postgraduate Studies in English Language Education at Ho Chi Minh City Open University, Vietnam, Professor Emeritus of Languages and Second Language Education at the University of Canberra, Australia. He is the current President of AsiaCALL (Asia Association of Computer-Assisted Language Learning).

Butsakorn Yodkamlue is a lecturer in the School of Foreign Languages, Suranaree University of Technology, Thailand. She earned her Ph.D degree in linguistics at the University of South Carolina, USA in 2008.