TRAINING TAIWANESE EFL LEARNERS TO PERCEIVE ENGLISH LEXICAL STRESS CONTRAST: A PILOT STUDY

Shu-chen Ou

National Sun Yat-sen University, Taiwan
sherryou@faculty.nsysu.edu.tw

ABSTRACT

This study investigates the effect of perceptual training on the learning of non-native lexical stress contrast. Nineteen participants who speak Mandarin or Taiwanese as their native language and have acquired English as a foreign language for an average of 7.2 years attended a perceptual training course to learn to identify English disyllabic nouns and verbs which differ only in the location of stress (e.g., *pérmit* (n.) vs. *permít* (v.)). A comparison of the L2 learners' accuracy rates in the pretest and post-test shows (i) an increase of 33% when stress is realized in the rising intonation and (ii) a decrease of 13% when stress is realized in the falling intonation. This suggests that perceptual training is useful not only for the perception of non-native segmental contrasts and lexical tones, but also has effects, either positive or negative, on the perception of the lexical stress contrast which involves multiple phonetic correlates.

Keywords: perceptual training, word stress

1. INTRODUCTION

It is well known that adults experience difficulties in perceiving the non-native phonological contrasts that are very similar to those in their native language [1, 2]. Nevertheless, empirical studies have shown that sufficient exposure to the target language can improve learners' perception of the non-native sound contrasts to some extent, revealing that adults' phonological perception can be adjusted to be more target-like with experience. In addition, several studies investigate whether non-native speakers' perception of non-native phonological contrasts can be modified via laboratory-controlled training, in which a particular phonological feature is auditorily presented in a contrastive way to non-native speakers. Overall, these studies indicate that the application of such a training technique has a positive effect on American English speakers' perception of the VOT in a three-way dimension [7], French speakers' perception of the English /θ-ð/ contrast [4] and Japanese speakers' perception of the English /l-r/ contrast [6]. The technique has also been proved to be effective for the perception of non-native lexical prosody (e.g., American English speakers' perception of Mandarin lexical tones) [12].

Non-native speakers are also found to perceive the lexical stress contrast in a different way from native speakers. For instance, French speakers are reported to be partially deaf to Spanish non-word pairs that use stress to differentiate meanings (e.g., *p ki* vs. *pik*) [10]. In addition, Dutch native speakers are more sensitive to the pitch difference than English native speakers when perceiving English primary and secondary stress (e.g., *mú*sic vs. *mu*sé*um*) [2]. Moreover, native speakers of lexical tone languages like Mandarin tend to over-rely on the cue of high pitch in identifying primary stress, as evidenced by their high accuracy rates in identifying non-word pairs which differ only in the location of stress (e.g., *fórcept* vs. *forcépt*) when stress is associated with the nuclear pitch accent of high tone (H*) and low accuracy rates when stress is associated with low tone (L*) [8, 9].

No studies have been carried out so far to train non-native speakers to perceive lexical stress (e.g., *pérmit* vs. *permít*), which involves multiple phonetic correlates, including F0, duration and intensity. The motivation of this study, therefore, is to examine whether perception training of Taiwanese EFL learners has an effect on their perception of English lexical stress contrast.

2. DESIGN

Lexical stress, in this study, refers to the phenomenon that some English disyllabic nouns and verbs are mainly differentiated by the location of stress (e.g., *pérmit* vs. *permít*), but not by the categorical change of vowel qualities. The experiment includes three phases: pretest, perceptual training and post-test. The objective of the tests is to determine whether the perceptual training produces some sort of effect on EFL learners' perception of lexical stress. The experiment was done in 14 days: The pretest was run on the first day and the post-test on the last day. Between the two tests, the participants needed to choose any six days to finish six training sessions in the laboratory, depending on their availability.

2.1. Materials

Six pairs of English disyllabic nouns and verbs which differ only in the location of primary stress (i.e, *pérmit* vs. *permít, súrvey vs. survéy, résearch* vs. *reséarch, tránsplant* vs. *transplánt, ínsert* vs. *insért, ímport* vs. *impórt*.) were selected. These words were embedded in two contexts: (i) yes/no questions (e.g., *Did you say pérmit?* vs. *Did you say permít?*), and (ii) affirmative-answer sentences (e.g., Yes, I said *pérmit.* vs. *Yes, I said permít.*). The affirmative-answer carrier sentence is meant to realize stress in higher pitch since it receives high tone (H*), which was previously reported that Taiwanese EFL learners to mainly rely on in perceiving stress [8, 9]. On the other hand, in the yes/no carrier questions, the stressed syllable receives low tone (L*) and the perception of stress may have to rely on other correlates (e.g., duration change and contour shape) [5, 9]. The test words were recorded three times each by a female trained phonetician, a native speaker of English with North American accent on a SONY Hi-MD recorder, digitalized at 44 kHz (16 bits). The words were then segmented from the carrier sentences, and only the words were used in the experiment. Two recordings of each item, similar in the phonetic measurements of pitch, intensity and duration, were then chosen for the experiment. All six pairs were used in the pretest, training and post-test. The design allows us to see whether the training has effects on the participants' perception of stress by comparing their performance in the pretest with that in the post-test. Table 1 presents the phonetic measures of stressed and unstressed vowels of the experimental items in the falling intonation (recorded in the affirmative sentences).

Table 1: The phonetic measures of stressed and unstressed vowels in the falling intonation.

	V́	V	V	V́
Mean F0 (Hz)	273	145	186	227
Duration (ms)	166	165	121	213
Intensity (db)	75	72	70	74

As seen in Table 1, the stressed vowel is more prominent in the cue of mean F0 for all the items. The average intensity of the stressed vowels is slightly higher, but variability is found, i.e., six out of 24 have slightly higher intensity in the unstressed ones. In addition, the duration of the stressed vowels is not significantly longer than that of the unstressed ones (i.e., 166 ms vs. 165 ms). This can be accounted for by the intrinsic properties of the vowels in the selected words. Specifically, the second vowel of the words, *survey* and *import*, is intrinsically longer in both of the stress patterns. In sum, the robust cue in the falling intonation is likely to be higher F0 and perhaps intensity. Table 2 presents the phonetic measures of stressed and unstressed vowels in the rising intonation (recorded in the yes/no questions).

Table 2: The phonetic measures of stressed and unstressed vowels in the rising intonation

	V́	V	V	V́
Mean F0 (Hz)	172	276	174	178
Duration (ms)	280	199	111	252
Intensity (db)	70	71	69	70

The F0 of the first vowel has no significant difference, since, when it is stressed, it receives L*, and when it is unstressed, its pitch stays low. The pitch difference of the word pairs is mainly on the second syllable in terms of the mean pitch (i.e., lower when it is stressed and higher when it is unstressed) or the contour shape (i.e., high rising when it is unstressed and low rising when it is stressed) since the stressed vowel receives L* while the unstressed one does not in this type of intonation. In addition, the duration of the stressed vowel is always longer but the intensity of the stressed vowel is not significantly louder.

The training materials for the control group were six English songs. In each session, the participants were required to fill in the blanks with proper words by listening to the song provided and then record a version sung by the trainee. The training has nothing to do with the perception of English word stress pairs. The average length of practicing each song was about 15 minutes, similar to the length of the experimental group's training.

2.2. Participants

In the beginning, forty-one undergraduate, non-English major students, who had acquired English as a foreign language in Taiwan for 7.2 years (SD = 0.8) on average were recruited. They were randomly assigned to two groups, yielding 21 in the experimental group and 20 in the control group. The two groups took the same pretest and posttest but had different training materials. During the training, two participants in the experimental group and one in the control group withdrew so that the results from only 38 participants are reported in this study. All of them reported no hearing or speech problems. Each was paid 450 NT dollars (about 15 US dollars).

2.3. Procedure

2.3.1. Pretest

The procedure of the pretest and post-test was the same: the participants were visually presented with two words, one with stress on the first syllable and

the other with stress on the second syllable (e.g., *pérmit* vs. *permít*), followed by an auditory stimulus (e.g., *pérmit*), and then they were asked to make a choice of which word matched what they heard by pressing the appropriate button. The six pairs of words were constructed into 6 blocks, each of which had 8 stimuli (2 stress patterns x 2 intonations x 2 tokens). So there were 48 trials in total. Once the decision was made, the next trial proceeded. The total length of the test was about 10 minutes.

2.3.2. Perceptual training

The materials used in the training were the same test words as in the pretest. Each pair of words was presented in a contrastive way. Specifically, a pair of words (e.g., *pérmit* vs. *permít*) were presented simultaneously on the screen and below the visual stimuli were two corresponding sound stimuli. Participants clicked on the sounds to listen to the subtle differences in pronunciation as many times as they wanted. Subsequently, they moved onto the next slide which required them to perform six steps: (i) listen to a sound stimulus, (ii) say whether it is a noun or a verb, (iii) immediately check their answer by listening to a pre-recorded answer (i.e., *It's a noun.* or *It's a verb.*), (iv) listen to the same sound file again and repeat it orally at least once, (v) listen to a pre-recorded question (i.e., *The noun's* _____ or *The verb's* _____) and say the answer, (vi) listen to a sound file that is the correct answer to the previous question to self-check if their pronunciation is accurate. In each training session, there were 2 pairs of words. So there were 3 sessions in total. After they finished all three, they had to repeat the three sessions once more. The six sessions had to be finished in any six days after the pretest and before the post-test in a program-installed PC in the laboratory. On average each participant spent 15 mins in each training session, adding up to 90 mins for the whole course.

2.3.3. Post-test

The post-test was a repetition of the pretest.

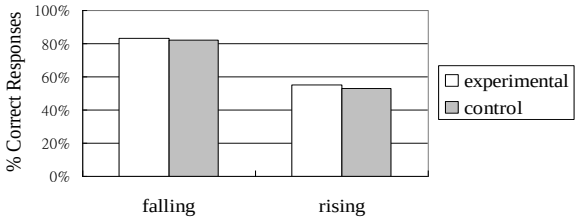
3. RESULTS

3.1. Pretest

Only the results of the 38 participants who attended the tasks in all three phases are reported. Figure 1 shows the accuracy rates for the stress contrasts in the two intonation patterns by the two groups before the training.

The accuracy rates in the falling intonation are 83% for the experimental group and 82% for control group while the accuracy rates in the rising intonation are 55% for the experimental group and 53% for the control group. An analysis of variance using *Group* (experimental and control) as the between-subjects factor, and *Intonation* (rising and falling) as the within-subjects factor was conducted. The effect of *Intonation* is significant [$F(1 , 46) = 70.32$; $p < 0.001$] while the effect of *Group* is not [$F(1, 46) = 0.01$, *n.s.*]. This indicates that both groups had few difficulties in perceiving the stress contrast in the falling intonation, but both of them were apparently poor at perceiving such a contrast in the rising intonation.

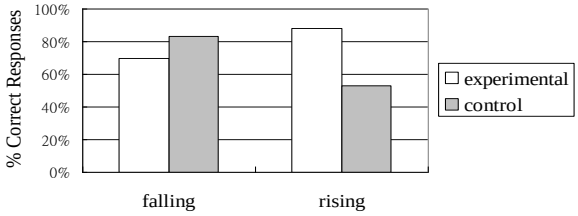
Figure 1: Overall correct identification on the lexical stress contrast in the pretest by group



3.2. Post-test

Figure 2 presents the performance of the two groups in the post-test after the perceptual training.

Figure 2: Overall correct identification on the lexical stress contrast in the posttest by group



The accuracy rates in the falling intonation are 70% for the experimental group and 83% for control group while the accuracy rates in the rising intonation are 88% for the experimental group and 53% for the control group. Overall, the control group did not improve or deteriorate in their identification of the stress contrast [$t(11) = -0.84$ for falling and $t(11) = 0.11$ for rising, both n.s.]. On the other hand, the experimental group's perception was significantly affected by the training sessions. That is, their perception of the stress contrast in the rising intonation has improved by 33% [$t(11) = -4.60$, $p < 0.01$]. This finding is compatible with the previous studies that trained non-native speakers to perceive segmental contrasts and lexical tones. However, the experimental group's perception of the stress contrast in the falling intonation dropped after the training (i.e., -13% from pretest 83% to post-test 70%) [$t(11) = 3.02$, $p < 0.01$]. This is somehow surprising since the training produced a negative effect on the learners' perception

of the stress contrast in the falling intonation. A detailed look into the data reveals that the drop occurs particularly in the word pairs of which the second syllable was longer in length than the first syllable no matter whether stress is on the first or second syllable (i.e., *survey* and *import*). A double check with the participants about how they made their choices confirmed the author's speculation that they mainly relied on the high pitch to determine the stress in the pretest, while during the training, they became aware that the cue of duration might also play a role. So they determined stress sometimes by higher pitch, but sometimes by longer duration, in the post-test. This may explain why their perception of stress in the falling intonation, in which the stressed syllable receives high pitch (H*), deteriorated after the training.

4. DISCUSSION AND CONCLUSION

The laboratory training has been applied in numerous studies to help non-native speakers perceive phonological contrasts in the target language, including segments such as consonants and vowels and suprasegmental features such as lexical tones. The general findings suggest that the method is effective in improving non-native speakers' phonological sensitivity to the target language. This study used a similar paradigm to train EFL learners in Taiwan to perceive the lexical stress contrast of English because previous studies have reported that these learners failed to identify word pairs differing in the location of stress using phonetic correlates such as rising contour shape or duration change. The study produced two main findings. First, their perception of stress in the rising intonation improved significantly (i.e., a significant increase of 33% from the pretest to post-test). This suggests that the perceptual training is effective in improving not only learners' perception of non-native phonological contrasts such as segments and lexical tones, but also their perception of English lexical stress in the rising intonation condition. Second, it was also found that the training reduced the learners' accuracy rates in perceiving stress contrast in the falling intonation (i.e., a decrease of 13% from the pretest to post-test). The second finding is interesting since it indicates learners' development of awareness that stress is signified not only by the cue of higher pitch, but also by other cues such as duration according to their self-reports and an analysis of the nature of the test items. The accuracy dropping serves as an indication that learners were adjusting the way they perceived English lexical stress from reliance on high F0 to other phonetic cues (e.g., longer duration), even though their perception was probably still non-native-like.

This study can be expanded in several directions. First, whether the trainees' ability of perceiving stress can be generalized to new items is worth investigating. In addition, how long their acquired ability can be retained is also an issue to look at. Moreover, whether their ability can be extended to their production can be explored too. Finally, the psychological and pedagogical implications of the accuracy drop as an effect of the perceptual training can also be studied further.

5. ACKNOWLEDGEMENTS

This project was supported by the National Science Council, Taiwan, from 2010 to 2011 (Project number: NSC 99-2410-H-110-055). Thanks go to Professor Karen Chung, Professor Janice Fon and Miss Sally Chen's assistance in sound recording. Thanks also go to the assistants in my work team, Cathy Cheng and Ana Hsieh, Ya-wen Hsiao and Anlo Liu for their assistance in preparing and conducting the experiments.

6. REFERENCES

[1] Best, C. 1995. A direct realist view of cross-language speech perception. In Strange, W. (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press.

[2] Cutler, A., Wales, R., Cooper, N., Janssen, J. 2007. Dutch listeners' use of suprasegmental cues to English stress. *Proc.16th ICPhS* Saarbrücken, 1913-1916.

[3] Flege, J. 1995. Second language speech learning: Theory, findings, and problems. In Strange, W. (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press.

[4] Jamieson, J., Morosan, D. 1986. Training non-native speech contrasts in adults: Acquisition of the English /θ/ and /ð/ contrast by francophones. *Perception and Psychophysics* 40, 205-215.

[5] Ladd, D. 1996. *Intonational Phonology*. Cambridge: CUP.

[6] Logan, J., Lively, S., Pisoni, D. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *J. Acoust. Soc. Am*. 96, 2076-2078.

[7] McClaskey, C., Pisoni, D., Carrell, T. 1983. Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics* 34, 323-330.

[8] Ou, S. 2010. Identification and discrimination of word stress by Taiwanese EFL learners. *Speech Prosody 2010* 100962, 1-4.

[9] Ou, S. 2010. Taiwanese EFL learners' perception of English word stress. *Concentric: Studies in Linguistics* 36, 1-23.

[10] Pepperkamp, S., Dupoux, E, Sebastian-Galles, N. 1999. Perception of stress by French, Spanish and bilingual subjects. *Proceedings of Eurospeech* 6, 2683-2686.

[11] Pisoni, D., Aslin, R., Perey, A., Hennessy, B. 1982. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology* 8, 297-314.

[12] Wang, Y., Spence, M. 1999. Training American listeners to perceive Mandarin tones. *JASA* 106(6), 3649-3658.