Identification and Discrimination of Word Stress by Taiwanese EFL Learners

Shu-chen Ou

Department of Foreign Languages and Literature, National Sun Yat-sen University, Taiwan

sherryou@faculty.nsysu.edu.tw

Abstract

This paper investigates Taiwanese EFL learners' identification and discrimination of English word stress when the cue of pitch is manipulated. Forty Taiwanese EFL learners and twenty English controls participated in two perceptual experiments. In Experiment 1, the participants were asked to identify a perceived non-word when its stressed syllable was realized phonetically either (a) in higher pitch, or (b) in a low rising pitch contour. In Experiment 2, they were asked to discriminate a given paired stimuli with either identical or different stress pattern(s). The results show that while all of the subjects had little difficulty in identifying word stress in the condition (a), they all had great difficulty in doing so in the condition (b). However, these subjects showed substantial sensitivity to stress differences in the condition (b) and their perception was psycho-acoustically based whereas the English native controls' was more categorical. The findings suggest that Taiwanese speakers are still sensitive to the stress variations of English in spite of the lack of lexical stress contrasts in Taiwanese Mandarin.

Keywords: prosody perception, L2 lexical stress, tone-stress interlanguage

1. Introduction

Languages differ from each other in how they use the cue of pitch in their phonological systems [2]. Some languages rely primarily on pitch in lexically marking certain syllables in a word to differentiate meanings. Such languages are typologically referred to as tone languages (e.g., Chinese) or pitch accent languages (e.g., Japanese). Other languages use stress (sometimes in combination with lexical tone or pitch accent) to mark prominent syllables (e.g., English). Stress has a number of phonetic correlates, including duration, spectral tilt, and segmental quality, and also often affects -- in a rather complicated way -- the pitch contour of the utterance.

The language-specific phonological use of pitch may influence speakers' perception of the lexical prosody of another language. For instance, it is indicated that native speakers of Mandarin Chinese tend to interpret English word stress as tonal differences in the code-switching context [2]. Specifically, it is reported that an English unstressed syllable can trigger the 3rd tone sandhi (i.e., a low tone becomes a rising tone when it is followed by another low tone) when English words are inserted into Chinese sentences (e.g., *hao* LL *professor* → *hao* MH *professor* 'good professor'). In contrast, English primary stress and secondary stress do not trigger the tone sandhi rule because they carry the feature [+high]. It is not clear, however, whether the tendency of Mandarin Chinese speakers to interpret word stress as tonal differences translates to a tendency to identify the location of English stress in the setting of second language (L2) acquisition. While it might be possible that the tonal interpretation occurs only in code-switching context, it might also be possible that the effect is still persistent in the course of L2 acquisition and thus impedes the development of target-like perception of English lexical stress.

Investigations of the perception of lexical prosody variations by non-native speakers reveal that French speakers are sensitive to the phonetic variations of lexical tone and stress in discrimination tasks even though their phonological system uses neither tone nor stress to distinguish word meanings [3][4][5][6]. These findings suggest that non-native speakers without lexical prosody are still sensitive to the phonetic variations of lexical prosody in another language even though they are unable to use relevant cues phonologically. Then a question arises is whether Taiwanese speakers are sensitive to the phonetic variations of English lexical stress even though their L1 phonological system does not use stress to differentiate word meanings.

To summarize, this study is to investigate how Taiwanese EFL learners identify the location of stress in English non-word pairs and when the cue of pitch is manipulated and whether they can discriminate English word stress at the phonetic level.

2. Method

Do native speakers of lexical tone languages over-rely on the cue of F0 in identifying stressed and unstressed syllables when learning English as a second language? To be specific, do they equate primary stress with [+high] tones and weak stress with [-high] tones. This question is tested by seeing how Taiwanese EFL learners identify a stressed syllable when the F0 is manipulated in the different intonation patterns of English. Specifically, in affirmative statements of North American English, the stressed syllable of the focused word is signified by the high nuclear pitch accent (H*) and followed by a low boundary tone (L%), resulting in a falling tone contour. On the other hand, in yes/no questions, the stressed syllable of the focused word is signified by the low nuclear pitch accent (L*) and followed by a high boundary tone (H%), resulting in a rising tone contour. Though in some varieties of English, the H*L% intonation is used in yes/no questions (e.g., Brighton, England), this does not influence our study because the rising intonation is used for the purpose of manipulating the cue of pitch in stressed and unstressed syllables only. The design allowed us to see whether EFL learners were able to identify the position of word stress when high pitch is removed since in the rising intonation, the stressed syllable cannot be cued by high pitch. If learners over-rely on the cue of high pitch in determining primary stress, they will have difficulties in identifying stress when it is carried by the rising intonation. To avoid the effect of lexical memorization and word frequency, the study used non-words in probing into the L2 learners' perceptual knowledge of English word stress they have acquired.

2.1. Experiment 1

2.1.1. Materials

A pair of nonsense words differing only in the location of lexical stress were constructed (i.e., *fércept* vs. *fercépt*) and were put into two contexts: (i) yes/no-question carrier sentences (i.e., *Are you a fércept?* vs. *Are you a fercépt?*), and (ii) affirmative-answer carrier sentences (e.g., *Yes, I am a fércept. vs. I am a fercépt.*). In the affirmative answers, one of the phonetic cues signifying the stressed syllable is higher pitch because it receives the nuclear pitch accent of high tone (H*). On the other hand, in the yes/no questions, the stressed syllable is targeted by L* so stress is cued by something else. In addition to the two sentential contexts, the non-words were also recorded in their isolated forms (*fércept* [fɝ́sɛpt] vs. *fercépt* [fɚsɛ́pt]), with a falling intonation for use in the learning phrase before the experiment. Another pair of nonsense words with a segmental contrast (i.e., *tóoper* vs. *tóoker*) was designed and meant to establish baseline performance because both /p/ and /k/ occur in the phonological systems of English and Mandarin Chinese. The items were recorded 3 times each by a trained female phonetician, a native English speaker from North America, on a SONY HI-MD recorder. All recorded items were digitized at 44 kHz (16 bits). Two recordings which were more similar in the measurement of three phonetic features (i.e., pitch, duration and intensity) were selected for each item. Figure 1 shows the pitch contour of the non-word pair. Table 1 shows the means for F0, duration and intensity for the average of the two tokens of each non-word in falling intonation. It is clear that the stressed syllables have higher pitch than the unstressed neighbors.

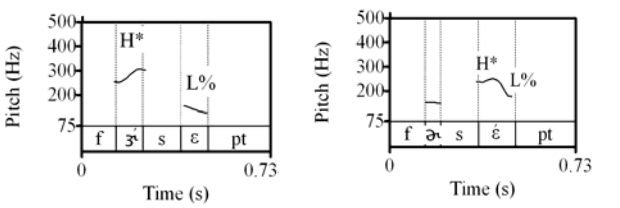

Figure 1: *Pitch contours of fércept (left) and fercépt (right) in the falling intonation*

Table 1. *Phonetic measures of stressed and unstressed syllables of non-words in the falling intonation*

	fércept		*fercépt*	
	[ɝ]	[ɛ]	[ɚ]	[ɛ]
F0 average (Hz)	281	141	147	233
duration (ms)	118	98	55	133
intensity (db)	79	69	70	74

In addition to higher pitch, the stressed syllable has longer duration and greater intensity in the falling intonation condition.

Figure 2 presents the pitch contours of the non-words in the rising intonation. Table 2 shows the phonetic measures of the word pair in the rising intonation.

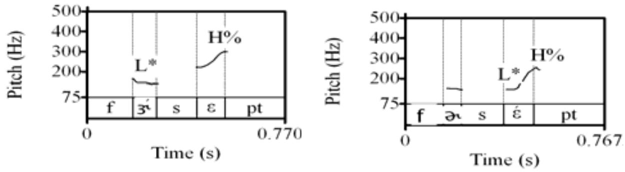

Figure 2: *Pitch contours of fércept (left) and fercépt (right) in the rising intonation*

Table 2. *Phonetic measures of stressed and unstressed syllables of non-words in the rising intonation*

	fércept		*fercépt*	
	[ɝ]	[ɛ]	[ɚ]	[ɛ]
F0 average (Hz)	151	256	153	177
duration (ms)	116	119	63	129
intensity (db)	65	73	63	69

In the rising intonation pattern, the F0 of the first syllable has no difference (151 Hz vs. 153 Hz) because when it is stressed, it receives L*; when it is unstressed, it receives a low tone automatically. The pitch difference of the two words lies on the second syllable, i.e., higher pitch when it is unstressed (256 Hz) and lower pitch when it is stressed (177 Hz) because the stressed syllable receives L*. Put simply, the second syllable has a high rising contour when it is unstressed while it has a low rising contour when it is stressed. The duration of the stressed syllables is not necessarily longer than the unstressed neighbors. Again, intensity seems not to be very different in the stressed and unstressed syllables. In other words, in this intonation pattern, the stressed syllable is not signified by high pitch, but instead by a low or high rising pitch contour in the second syllable.

2.1.2. Participants

Two experimental groups of Taiwanese EFL learners were recruited to participate in the experiment: 20 graduate students who had learned English as a foreign language for at least 10 years (Group H hereafter) and 20 high school students (Group L hereafter) who had learned English as a foreign language for about 3 years. All of them reported no hearing and speech problems. In addition, 20 English native speakers were also included as controls. Each was paid 120 NT dollars (about 3.5 US dollars).

2.1.3. Procedure

The whole procedure consisted of two phases: a learning phase and a test phase. In the learning phase, participants were trained to match sound stimuli with corresponding pictures (i.e., *fércept* matches the picture 'teacher', *fercépt* matches 'student', *túpu* 'boy' and *túku* 'girl'). The sound stimuli which were recorded in isolation with falling pitch-accent were used in this learning phase. Participants were allowed to listen to the sound stimuli as many times as possible in relation to the pictures until they felt that they had memorized all the four words. There was a short quiz at the end of the learning phrase to ensure that they had learned the pairings. In the test phase, they saw two pictures on the screen, heard a nonsense word in a sentence, and then made a choice of which picture matched what they had just heard by pressing the associated button. Each minimal pair was constructed in two experimental blocks so that there were 4 blocks in total (2 minimal pairs (segment and stress) x 2 intonation patterns (rising and falling)). Each block contained 8 trials (2 non-words x 2 tokens x 2 picture orders (i.e., student–teacher and teacher–student)). All the visual and auditory items were presented randomly in each block, controlled by E-Prime. The experiment lasted 25 minutes on average.

2.1.4. Results

Figure 3 shows the overall accuracy rates for three groups (English controls and Learner Groups H and L) for the segmental contrast embedded in the two intonation patterns (i.e., rising and falling). The result suggests that these learners had little difficulty in identifying the segmental contrast no matter what type of intonation pattern was used to carry the non-words.

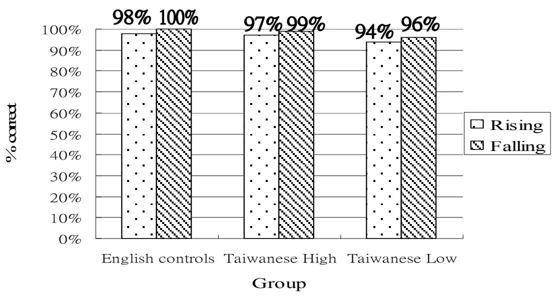

Figure 3. *Overall identification performance on the segmental contrast (tóoper vs. tóoker) by group*

Figure 4 shows the overall accuracy rates for the stress contrast in the rising intonation for three groups. The error rates were clearly high in the two experimental groups.

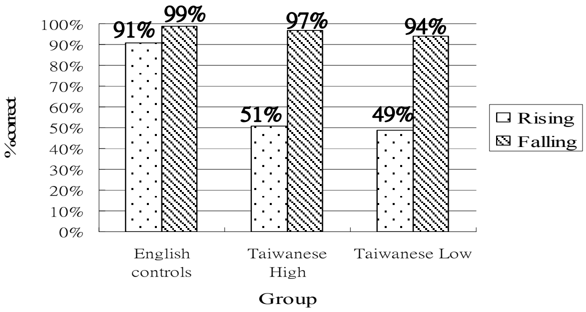

Figure 4. *Overall identification performance on the stress contrast (fércept vs. fercépt) by group*

An analysis of variance with the between-subjects factor Group (English controls, Group H and Group L), and 2 within-subjects factors, Contrast (segment and stress) and Intonation (rising and falling) was conducted. There were significant effects for all three factors, Group [$F(2,57) = 26.85$, $p < 0.01$], Contrast [$F(1,57) = 866.14$; $p < 0.01$], and Intonation [$F(1,57) = 754.42$; $p < 0.01$]. There was a significant interaction between Group and Contrast [$F(2,57) = 125.82$; $p < 0.01$]. Bonferroni's post hoc test shows that the English group was different from the two experimental groups at the 1% level, and the two experimental groups were not different from each other. This indicates that these L2 learners had few difficulties in perceiving the two nonsense words in the falling intonation, but they were apparently poor at perceiving the contrast in the rising contour. The low error rate of the English controls indicates that there were phonetic cues other than high F0 available to identify the stress position (most notably the different contour F0 shapes and duration ratio change according to Table 2), which native speakers of English must have used to identify the words in the rising intonation, but the Taiwanese EFL learners were impervious to. This suggests that Taiwanese EFL learners have over-relied on the cue of high pitch in identifying stressed syllables at the expense of other correlates of stress such as the contour F0 shapes and duration ratio change phonologically. An extensive question to ask, then, is whether the Taiwanese EFL learners can perceive English stress differences at the phonetic level.

2.2. Experiment 2

An AX discrimination task was employed to evaluate the Taiwanese subjects' ability to acoustically discriminate the English stress contrasts. In this task, subjects hear a pair of sounds, one of which contains, for example, a *fércept*, and the other a *fercépt* and have to indicate whether the two sounds are the same or different.

2.2.1. Materials

The items used in this experiment were the same as those in the first experiment except that they were not in the sentence contexts but segmented to be individual words (i.e., *fércept* and *fercépt* from *Are you ____?* and *I am a ______.*) In total, there were 8 stimuli (2 stress patterns x 2 tokens x 2 intonation patterns). One token of each stress pattern (e.g., Token A of *fércept*) was, then, paired with another sound stimulus that was either a different or identical token of the same stress pattern (e.g., Token A or Token B of *fércept*) or a token of the other stress pattern word (e.g., one of two tokens of *fercépt*). The pairing created 32 trials which were divided into two blocks: (i) 16 pairs in falling intonation and (ii) 16 pairs in rising intonation.

2.2.2. Participants

The participants in this task were the same as those in the first experiment, namely, the 20 Taiwanese speakers who had learned English as a foreign the language for at least 10 years (Group H) and the 20 high school students (Group L) who had learned English for about 3 years. They took part in the task within one week after their participation in the first experiment. Twenty native speakers of English were also recruited to participate for comparison.

2.2.3. Procedure

In each trial, subjects heard a pair of sounds displayed one after the other with an interval of 500 ms between the two sounds, and then they had to determine whether the two sounds were the same or different. Immediately after their response, there was a neutral sound, "OK", played. They were told that there were 32 trials in the beginning of the task and there was a break for one minute. Each trial was played only one time. The task lasted for 4 minutes or so.

2.2.4. Result

A response that any paired sounds of the same stress pattern were "same" was counted as correct (i.e., 1 point) and "different" as incorrect (i.e., zero point). Similarly, a response that the sounds of different stress patterns were "different" was counted as correct and "same" as incorrect. In other words, sixteen points were given if the sounds of different stress patterns were responded as being "different" while the sounds of the same stress patterns were responded as being "the same".

Figure 5 shows the overall accuracy rates for three groups (English controls and Learner Groups H and L) to discriminate the stress differences (i.e., initial stress vs. final stress) in two intonation conditions (i.e., rising and falling).

The result suggests that these learners were just as good as the English native speakers at discriminating stress differences in the falling intonation condition whereas the same was not found in the rising intonation condition.

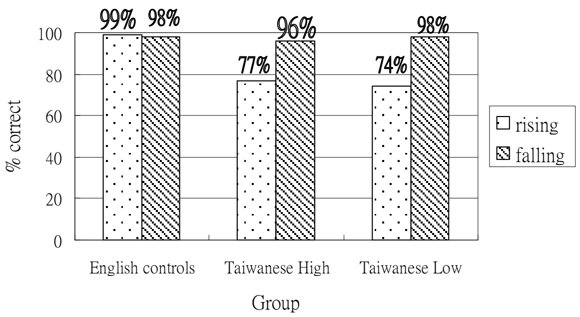

Figure 5. *Overall discrimination performance on the stress contrasts (fércept vs. fercépt) by group*

An analysis of variance with the between-subjects factor, Group (English controls, Group H and Group L), and the within-subjects factor, Intonation (rising and falling), was conducted. There were significant effects for both of the factors, Group [$F(2,57) = 85.33$, $p < 0.01$] and Intonation [$F(1,57) = 243.00$; $p < 0.01$]. There was also a significant interaction between Group and Intonation [$F(2,57) = 65.44$; $p < 0.01$]. Bonferroni's post hoc test shows that the English group was different from the two experimental groups at the 1% level, and the two experimental groups were not different from each other. This indicates that these L2 learners had fewer difficulties in discriminating the stress contrast in the falling intonation, but they were apparently poor at discriminating the stress contrast in the rising intonation.

In addition, a closer look into the errors committed by the two groups yields some more interesting findings, as shown in Table 3.

Table 3. *Error patterns of the experimental groups*

	"same" perceived as *"different"*	*"different perceived as "same"*
Group H	95% (73/77)	5% (4/77)
Group L	93% (76/82)	7% (6/82)

Table 3 shows that both of the Taiwanese groups tended to perceive the sound pairs that were indeed two different tokens of the same stress pattern as being different (e.g., Token A of *fércept* and Token B of *fércept*) in the rising intonation, but this tendency was not found in the English controls. This indicates that these EFL learners were sensitive to the phonetic differences of English word stress evel even though they could not use such a stress contrast to distinguish word meaning, as shown in the first experiment.

3. Discussion and Conclusion

In the two experiments, it was found that Taiwanese EFL learners were unable to identify word stress when the stressed syllable was signified by the low or rising pitch contours. The reliance on high pitch in perceiving stress may show a persistent effect of learners' L1 lexical tone. However, it is also possible that the reliance on pitch is a general tendency for L2 beginning learners' perceiving stress regardless of their L1 background. The theoretical account for Taiwanese EFL learners' reliance on the tonal feature [+high] in perceiving stress needs further research to be more thoroughly evaluated. In addition, even though Taiwanese EFL learners exhibited great difficulties in the phonological use of stress to differentiate word meanings, they were found to be sensitive to the stress differences in the AX discrimination task. Their sensitivity was not categorical and is likely to be psycho-acoustic because they judged the two different tokens of a certain stress pattern in the rising intonation condition to be different, suggesting that they were partially deaf to stress at the phonological level but not completely deaf to stress at the phonetic level. This is compatible with the finding that French speakers' perception of Mandarin tones is non-categorical but psycho-acoustic even though lexical tone is not used to differentiate word meaning in French [2].

This study can be expanded in a few directions: (a) Is the reliance on pitch in identifying stress really an effect of the tonal system of Chinese? Can it not be a general feature of L2 stress perception? (b) Will native speakers of a pitch accent system (e.g., Japanese, Basque) who also use pitch a primary cue to lexical contrast also show this effect in L2 English stress perception? Both questions are worth investigating in the future.

4. Acknowledgements

The author would like to thank Dr. Mits Ota and Prof. Robert Ladd at the University of Edinburgh for their comments and suggestions on this paper, Prof. Karen Chung, Prof. Janice Fon and Miss Sally Chen at National Taiwan University for assistance with sound recording, and participants in the experiments from National Sun Yat-sen University and National Lu-chu High School in Taiwan. This paper is written based on a research project, NSC 97-2410-H-110-055, granted by National Science Council, Taiwan. Part of this study (i.e., Experiment 1) has also appeared in *Concentric 36*(1).

5. References

[1] Beckman, M. E., Stress and Non-stress Accent. Dordrecht: Foris, 1986.

[2] Cheng, C. C., "English stresses and Chinese Tones in Chinese sentences", Phonetica, 18: 77-88, 1968.

[3] Dupoux, E., Peperkamp, S. & Sebastian-Galles, N. "A robust method to study stress "deafness". *Journal of the Acoustical Society of America, 110* (3), 1606-1618, 2001.

[4] Dupoux, E., Sebastian, N., Navarrete, E., Peperkamp, S., "Persistent stress "deafness": The case of French learners of Spanish", Cognition, 106: 682-706, 2008.

[5] Hallé, P. A., Chang, Y. C., Best. C. T. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. Journal of Phonetics, *32*, 395-421, 2004.

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