Concentric: Studies in Linguistics
36.1 (January 2010):1-23

Taiwanese EFL Learners' Perception of English Word Stress*

Shu-chen Ou
National Sun Yat-sen University

This paper investigates how Taiwanese EFL learners perceive non-word pairs which differ only in the location of stress (e.g., *fércept* vs. *fercépt*) when the phonetic cue of pitch is manipulated. Fifty-eight Taiwanese EFL learners participated in two forced choice perceptual experiments, in which they were asked to identify a perceived non-word when its stressed syllable was signified either (i) by higher pitch or (ii) by a low rising tonal contour. The results show that, while these L2 learners had little difficulty in perceiving stress when the stress was signified by higher pitch, they all had great difficulty in doing so when the stress was signified by the low rising tonal contour. In addition, analyses of their errors show that less experienced learners relied mainly on higher pitch or rising pitch contour in guessing the position of stress, which may indicate a persistent effect of their L1 tonal system or L2 learners' universal tendency of perceiving stress, while more experienced learners referred to the information of morpho-syntactic categories as a strategy in guessing the position of stress, suggesting their phonological awareness of the difference between lexical tone and lexical stress at their developmental stage.

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

1. Introduction

Languages differ from each other according to three basic lexical prosody phenomena: tone, pitch-accent, and stress (Beckman 1986). Some languages rely primarily on pitch to lexically mark certain syllables in a word to differentiate meanings. Such languages are typologically referred to as tone languages (e.g., Chinese and Hausa) or pitch accent languages (e.g., Japanese and Basque). In lexical tone languages, pitch height and contour shape are used to distinguish one word from another. Mandarin Chinese, for example, contains four lexically contrastive (or phonemic) tones (e.g., Chao 1968, Cheng 1973). In this language, for example, the syllable *ma* means 'mother' when its pitch height is high level, 'hemp' when it is rising, 'horse', when it is low, and 'scold' when it is falling. The second type of language has lexical pitch-accent: one syllable per word is made prominent by means of a specific pitch height. Japanese is a typical pitch-accent language. In Japanese, the

* 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, Mr. Allen Handel for his editorial assistance, 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. Thanks also go to two anonymous reviewers of *Concentric: Studies in Linguistics* for their helpful and constructive suggestions. This paper is written based on a research project, NSC 97-2410-H-110-055, granted by the National Science Council, Taiwan. Portions of this paper were previously presented at the 17th Manchester Phonology Conference on May 28-30, 2009.

syllable that carries the lexical pitch accent is marked by a fall in pitch height. For example, when the word *ame* means 'rain', its pitch accent occurs on the first syllable, followed by a fall on the second syllable. One the other hand, when *ame* means 'candy', its pitch accent is on the second syllable, followed by a fall on the following syllable, and the first syllable gets a default low automatically. In addition to lexical tone and lexical pitch-accent languages, other languages use stress, sometimes in combination with lexical tone or pitch accent, to mark prominent syllables.

Stress has a number of phonetic correlates, including duration, intensity, and segmental quality, and also often affects, in a rather complicated way, the pitch contour of the utterance. In comparison with lexical pitch-accent languages, a stressed syllable may also have higher pitch, but this is not always the case. English is of this type. A stressed syllable in English generally tends to be longer in duration, greater in intensity, and less centralized in vowel quality. However, the pitch of a stressed syllable varies with various intonation patterns. Specifically, a stressed syllable has high pitch when it receives a type of intonation, or it has low pitch when it receives another. According to Ladefoged (2006), the intonation pattern that indicates simple statements in English is H*L-L%, where H* is a high nuclear pitch accent, L- a low phrase pitch accent, and L% a low boundary tone, whereas the intonation pattern that indicates yes/no questions is L*H-H%, where L* refers to a low nuclear pitch accent, H- a high phrase pitch accent, and H% a high boundary tone. In both cases, the nuclear accents are associated with the stressed syllables. Therefore, the pitch of a stressed syllable may be high or low in pitch. The following shows different pitch realizations of stressed syllables in English, taken from Ladefoged (2006: 126).

(1) a. Amélia. (simple statement)

	H*			L-L%
ə	m	i	l i ə	

b. Amélia? (yes/no questions)

	L*			H-H%
ə	m	i	l i ə	

For example, the word, *Amélia,* is stressed on the second syllable *-me-*, as marked by the apostrophe. When the word is used to respond to the question, *What is her name?*, its stressed syllable is high in pitch (H*), as shown in (1a). However, the same syllable is low in pitch when it is carried in a question like *Did you say Amélia*?, as shown in (1b). In other words, stress languages like English do not use pitch in the same way as lexical tone and lexical pitch accent languages.

The language-specific phonological properties may influence speakers' speech perception of the phonological contrast in another language and further affect L2 learners' acquisition of a new sound (Best 1994, Brown 1998, 2000, Flege 1995). For instance, Goto (1971) has reported that Japanese listeners tend to map American /l/

and /r/ into a single sound category; as a result, they failed to discriminate the two sounds. Advanced by Brown (2000), the Japanese speakers' failure in discriminating the /r/ and /l/ of English is due to the fact that these speakers have long been trained to ignore the phonetic features that are not used contrastively in their first language (L1) phonological system (i.e., [coronal] is not contrastive in Japanese). In addition, the influence of L1 suprasegmental properties onto second language (L2) word stress has been attested, but the findings are somehow conflicting. Altmann's (2006) investigation of stress perception in English as an L2 suggests that a predictable stress position in the L1 (e.g., French) is problematic while native speakers of L1 lexical tone languages (e.g., Chinese) do not have problems. On the other hand, Peperkamp et al. (1999) found that French subjects exhibit great difficulties in identifying non-words that differ only in the location of stress (e.g., *mípa* vs. *mipá*), and this inability is attributed to the fact that French does not use stress to mark lexical differences at the phonological level. Furthermore, while Kijak (2007) reports that both French and Chinese learners of Polish display problems with perceiving stress position in nonsense words, other studies do not (e.g., Pater (1997) with French-English bilinguals and Ou (2007) with Chinese and Vietnamese EFL learners).

Despite the controversies of L2 stress perception, it is worth mentioning that the previous studies mentioned above investigate this issue when the stressed syllables are signified by multiple phonetic cues such as higher pitch, longer duration, and greater intensity. It is, therefore, not clear whether the perception of word stress by these L2 learners involves multiple phonetic cues or only certain phonetic cues in a non-native-like way. For instance, it is highly possible that native speakers of lexical tone languages may rely on the cue of higher pitch in identifying the location of word stress at the expense of other phonetic cues of stress due to the fact that pitch is the most salient cue of lexical prosody in their native languages. The speculation is reasonable because, in the code switching context, it is indicated that native speakers of Mandarin Chinese tend to interpret English word stress as tonal differences (Cheng 1968). Specifically, it has been 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 using tonal contours in the setting of L2 acquisition. While it might be possible that the tonal interpretation occurs only in a code-switching context because speakers must adapt a non-native prosody into their

native phonological system, 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. Due to the controversies from the studies of L2 stress perception and evidence of tonal adaptation of lexical stress in the code-switching context, this study is motivated to investigate how Taiwanese EFL learners perceive the location of stress in English word pairs when the cue of pitch is manipulated.

2. Method

Do native speakers of lexical tone languages over-rely on the cue of pitch in identifying stressed and unstressed syllables when learning English as a second language? That is, does the tendency of native speakers of Chinese in Taiwan to interpret English word stress as tonal differences (i.e., primary stress carries the [+high] feature and weak stress carries the [-high] feature) translate to a tendency to identify the location of English stress in the setting of L2 acquisition? This question is investigated by testing how Taiwanese EFL learners determine English word stress when the phonetic cue of F0 is manipulated. Specifically, the author used two intonation patterns of North American English to manipulate the cue of pitch height of an English stressed syllable (i.e., F0 or fundamental frequency). As reviewed in the first section, in simple or affirmative statements of this dialect, the stressed syllable of the focused word is signified by the high nuclear pitch accent (H*), whereas in yes/no questions, the stressed syllable of the focused word is signified by the low nuclear pitch accent (L*). Though in some varieties of English, the H*L-L% intonation is used in yes/no questions (e.g., Brighton, England), this does not influence our study because the intonation pattern is used to manipulate the cue of pitch in stressed and unstressed syllables only.[1] The design allowed us to see whether EFL learners were able to identify a stressed syllable when the cue of the high nuclear pitch accent (H*) is replaced by the low nulear pitch accent (L*). If learners over-rely on the cue of high pitch in perceiving primary stress, they will have difficulties when the primary stress is signified by a low tone, that is, L*.

[1] Thanks go to Dr. Mits Ota for mentioning this dialectal difference in the intonation patterns of yes/no questions.

2.1 Experiment 1

2.1.1 Materials

Nonsense words were designed for this study based on several considerations. First of all, because our focus was to investigate whether the pitch changes affect the learners' perception of stress, real words with stress contrasts accompanied by vowel reductions (e.g., *récord* [rékɚd] vs. *recórd* [rɪkɔ́rd]) were not considered in this study. Even though there is a small number of English word pairs that differ only in stress positions (e.g., *pérvert* vs. *pervért* and *pérmit* vs. *permít*), they were not used due to another consideration, the lexical memorization effect, as first proposed by Pater (1997) and followed by many other studies of L2 stress acquisition (e.g., Davis and Kelly 1997, Peperkemp et al. 1999, Dupoux et al. 2008, Guion 2005, Ou 2006, 2007). Specifically, EFL learners who learn English in foreign language settings may have been exposed to various kinds of non-target-like input from non-native English instructors and other L2 learners. It may be the case that the learner has been exposed to an environment where both *résearch* and *reséarch* are produced without any distinction (e.g., initial stress). In other words, using real words in the experiment requires some additional work such as checking an individual learner's knowledge of English stress patterns. Moreover, English stress minimal pairs contain dialectal variations. For instance, the same stress pattern, *tránsfer*, is used for both nouns and verbs in American English but different stress patterns are used in British English (i.e., *transfer* (n.) vs. *transfér* (v.)) Again, this leads to the uncertainty of the input that learners have been exposed to. In order to avoid various confounding effects of using real words in the experiments, this study designed non-words to test Taiwanese EFL learners' perception of word stress.

A pair of nonsense words differing only in the location of stress were constructed (i.e., *fércept* vs. *fercépt*) and were carried in two contexts: (i) affirmative-answer sentences (i.e., *Yes, I am a <u>fércept</u>.* vs. *Yes, I am a <u>fercépt</u>.*), and (ii) yes/no-question sentences (i.e., *Are you a <u>fércept</u>?* vs. *Are you a <u>fercépt</u>?*), as shown in (2).

(2) a. The word pair, *fér.cept* and *fer.cépt*, carried in the H*L-L% intonation

H* L-L% H*L-L%

Yes, I am a *fér.cept*. vs. Yes, I am a *fer.cépt*.

b. The word pair, *fér.cept* and *fer.cépt*, carried in the L*H-H% intonation

L*H-H% L*H-H%

Are you a *fér.cept*? vs. Are you a *fer.cépt*?

When the word pair is carried in the affirmative-answer sentence with a falling intonation pattern (i.e., H*L-L%), as in (2a), the stressed syllable receives the nuclear pitch accent of high (H*). Under this condition, the pitch of the first syllable is higher when compared with its unstressed neighbor (i.e., *fér-* is higher than *-cept* and *-cépt* is higher than *fer-*). In contrast, when the non-words are carried in the yes/no questions with a rising intonation pattern (i.e., L*H-H%), as in (2b), the stressed syllable receives the low pitch accent (L*). If the description of the English rising intonation in (1b) is adequate, the non-word with initial stress (i.e., *fér.cept*) should have a rising contour which starts earlier than the non-word with final stress (i.e., *fer.cépt*) even though both words would have higher pitch in the second syllable because of the high phase accent pitch (H-) and the high boundary tone (H%). That is, the word with initial stress should have a high rising tonal contour on the second syllable whereas the word with final stress should have a low rising contour on the second syllable. If Taiwanese EFL learners over-rely on the high pitch in identifying a stressed syllable, they will not have difficulties in distinguishing the stress minimal pair when the words are carried in the falling intonation like (2a), but they will have significant problems in doing so when the words are carried in the rising intonation like (2b).

In addition to the stress minimal pair, another pair of nonsense words with a segmental contrast (i.e., *tóoper* vs. *tóoker*) was designed to allow a comparison of L2 learners' perceptual ability in segmental phonology. The phonemic contrast /p/-/k/ was meant to be equally easy for Taiwanese EFL learners and native speakers of English because the two segments occur in both languages, but the lexical stress contrast would be difficult for the L2 learners since the two languages are different in terms of their lexical prosody typologically. The phonemic contrast was meant to establish baseline performance, as suggested by Dupoux et al. (2001) and Peperkemp et al. (1999). The following lists the non-word items designed in the experiment.

(3) stress minimal pair (experimental items): *fércept* [fɝ́sɛpt] vs. *fercépt* [fɚsépt]
phonemic minimal pair (control items): *tóoper* [túpɚ] vs. *tóoker* [túkɚ]

The items (i.e., 4 non-words in two carrier sentences as in (2)) 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. The phonetic measures of the word pair (i.e., *fércept* [fɝ́sɛpt] vs. *fercépt* [fɚsépt]) were made on the vowels. Because it was hard to draw a boundary between the vowel and the retroflex /r/ in the first syllable, the measure of the first syllable included the

segment /r/ for 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 the falling intonation.

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

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

When the word is carried in the falling intonation, the three phonetic cues of the stressed syllable are more prominent than those of the unstressed syllable, that is, the F0 is higher in the first syllable of *fércept* while it is higher in the second syllable of *fercépt*. In addition, the duration is also longer in the first syllable of *fércept* while it is longer in the second syllable of *fercépt*. Finally, the intensity is slightly greater in the first syllable of *fércept* whereas it is slightly greater in the second syllable of *fercépt*. In other words, the stressed syllable in the falling intonation is signified by higher pitch, and longer duration and probably greater intensity as well. Figure 1 presents the contour shapes of an instance of the non-word pair carried in the falling intonation.

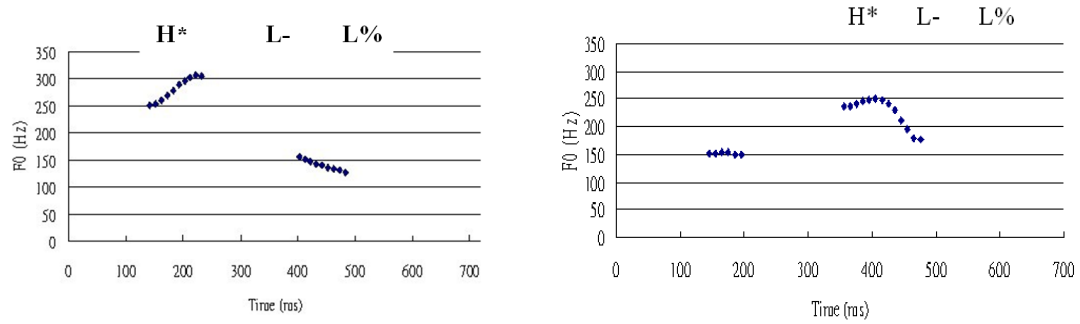


Figure 1. Pitch contours of fércept and fercépt in the falling intonation

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

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

	fércept		*fercépt*	
	[ɝˊ] of *fér*	[ɛ] of *cept*	[ɚˌ] of *fer*	[ɛ] of *cépt*
F0 average (Hz)	151	256	153	177
duration (ms)	116	119	63	129
intensity (db)	65	73	63	69

In the rising intonation, as expected, the F0 of the second syllable is higher in both *fércept* and *fercépt*, but in *fércept*, the difference is about 100 Hz, while in *fercépt* it is about 25 Hz. The bigger pitch differences in *fércept* is due to the fact that the nuclear pitch of high (L*) falls on the first syllable and the second syllable is approximately the locus of the two high tones (i.e., H- and H%). The smaller pitch differences in *fercépt* can be explained by the fact that the first syllable receives a low tone automatically since it is unstressed, and the second (stressed) syllable receives the nuclear low pitch accent (L*) followed by H- and H%. In other words, when the second syllable is stressed, it has a low rising pitch contour; when the second syllable is unstressed, it has a high rising pitch contour. As for duration and intensity, the stressed syllable is not always longer and greater than the unstressed neighbor according to the measurement. In sum, under the rising intonation condition, the stress minimal pair is signified by a low rising tonal contour (if stressed) and a high rising tonal contour (if unstressed). Figure 2 presents an instance of the non-word pair embedded in the rising intonation.

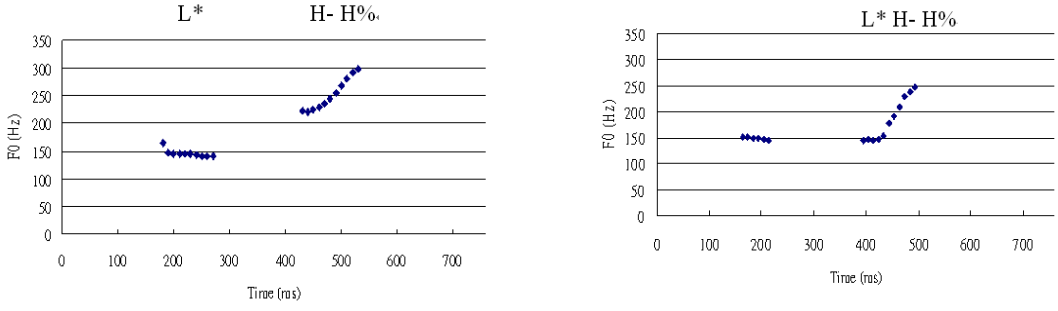


Figure 2. Pitch contours of fércept and fercépt in the rising intonation

2.1.2 Participants

In order to see whether the tonal reliance, if any, changes over the course of L2 development, two experimental groups of Taiwanese EFL learners were recruited to

participate: 20 graduate students who had learned English as a foreign language for at least 10 years (Taiwanese High hereafter) and 20 high school students who had learned English as a foreign language for less than 3 years (Taiwanese Low hereafter). All of the participants reported no hearing or speech problems. In addition, 20 English native speakers were also included as controls. Each participant was paid 120 NT dollars (about 3.60 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., *<u>fér</u>cept* matches Picture A, *fer<u>cépt</u>* matches Picture B, *tóo<u>p</u>er* Picture C and *tóo<u>k</u>er* Picture D). The sound-picture correspondence is detailed in the appendix. The non-words used in the learning phases were the same as those used in the tests (i.e., *fércept* vs. *fercépt* and *tóoper* vs. *tóoker*). However, the non-words used in the learning phases were isolated forms while those used in the experiments were carried in sentences (i.e., yes/no questions and affirmative statements). The isolated forms were segmented from the recorded sentences used in the tests (e.g., *fércept* from *Are you a fércept?* and *Yes, I am a fércept*). 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 four words. There was a short quiz containing the non-words read in the falling intonation at the end of the learning phrase to ensure that the participants had learned the pairings to an accuracy of more than 80%. In the test phase, participants were given a Forced Choice Picture Selection task (first used by Brown and Matthews 1993, 1997 in testing L2 segmental identification), in which a participant is presented with two pictures and a sound stimulus that corresponds to one of the pictures. In each trial, the participant will see a picture which refers to a non-word on the left side of the screen and a picture which indicates the other non-word on the right side. Meanwhile, the participant will hear a sound. The participant's task is to decide which of the pictures matches the sound stimulus s/he hears, then to indicate which by pressing the corresponding button. Each minimal pair was constructed in two experimental blocks so that there were 4 blocks in total (= 2 word 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). All the visual and auditory items were presented randomly in each block, controlled by E-Prime. In total, there were 32 trials, divided in the 4 blocks. The whole experiment, including the learning phase and the test phase, lasted 25 minutes on average.

In order to have successfully completed this task, the learner must have referred to

his or her internal lexical prosody representations of the words. If the participant's phonological representations of the two pictures were not distinguished from each other, then s/he would be unable to determine to which picture the verbal cue corresponded and should perform the task with chance accuracy. Successful completion of this task indicated that the subject had acquired the contrasts in English.

2.1.4 Results and discussion

If participants selected the target picture (i.e., the correct picture-sound match), a correct response was counted; otherwise, an error was counted. Figure 3 shows the overall accuracy rates for the three groups (English controls, Taiwanese High, and Taiwanese Low) for the segmental contrast (i.e., /p/ vs. /k/) embedded in the two intonation conditions (i.e., rising and falling). The results show near perfect performance was attained by the Taiwanese group on the control items in this task, suggesting 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.

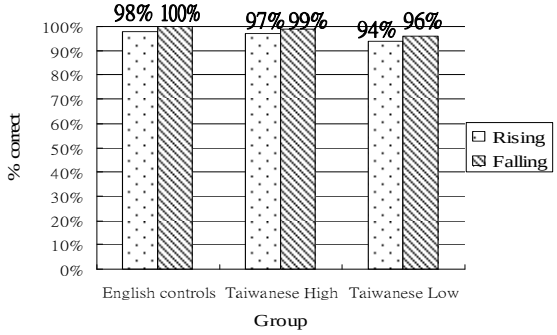


Figure 3. Overall picture performance on the segmental contrast (tóoper vs. tóoker) by group

Figure 4 shows the overall accuracy rates for three groups for the stress contrast carried by the two intonation conditions (i.e., rising and falling). Comparatively, the Taiwanese speakers were significantly poorer than the English controls at differentiating the initial stress and final stress in the rising intonation.

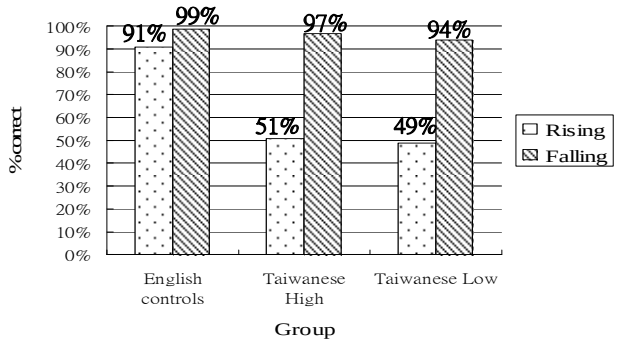


Figure 4. Overall picture performance on the stress contrast (fércept vs. fercépt) by group

A mixed ANOVA of accuracy rates using Group (English controls, Taiwanese High, and Taiwanese Low) as the between-subjects factor, and two within-subjects factors, Contrast (segment and stress) and Intonation (rising and falling) revealed 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 [$F(1,57) = 754.42$; $p < 0.01$]. In addition, there was a significant interaction between Group and Contrast [$F(2,57) = 125.82$; $p < 0.01$]. The interaction was due to the fact that there was an effect of Contrast for the two experimental groups, with stress yielding more errors than segment, but not for the English controls. Bonferroni's post hoc test further shows that the English group was different from the two Taiwanese groups at the 1% level, and the two Taiwanese groups were not different from each other. As expected, these L2 learners had few difficulties in perceiving the stress contrast in the falling intonation condition in which stressed syllables receive a high tone (H*), but they were apparently poor at perceiving the stress contrast in the rising contour in which stressed syllables receive a low tone (L*). However, the low error rate of the English controls indicates that there were phonetic cues other than the F0 available for native speakers of English to identify the stressed syllables in the rising intonation (e.g., different F0 contour shapes of the second syllable), but the Taiwanese EFL learners were impervious to these. This will be discussed in the following section.

Although the statistical results above show significant differences in L2 learners performance on the stress contrast carried in the rising intonation condition, one question readers may have is that their accuracy rate (about 50%) would seem to indicate that these learners have some knowledge of the contrast cued by the duration (although it is not native-like). However, in order to properly interpret these results, it is necessary to consider the baseline performance, that is, what the chance performance would be. If a learner has no phonological knowledge of the stress contrast cued by a low rising tonal contour and is, therefore, not able to determine which picture corresponds to the sound, he or she simply will be unable to decide

which picture corresponds to the sound (i.e., since the verbal representations of the non-word pair are not different, both pictures correspond to the given verbal cue). With a choice of two pictures, this participant has a 50 percent chance of selecting the correct one, just guessing. The observed performance, then, at 50%, is not significantly different from chance. We can infer with reasonable confidence from the Taiwanese speakers' performance on this task, that the stress contrast, in particular when it is phonetically cued by L* from the rising intonation, is not differentiated in their grammar. Performance on the other two contrasts, on the other hand, indicates that both the segmental contrast (i.e., /p/ vs. /k/) and the stress contrast cued by higher pitch are differentiated. In other words, the phonological structure that represents the segmental contrast and stress contrast cued by higher pitch has successfully been acquired by these learners while the stress contrast cued by a low tone (L*) has not.

In sum, the results show that, while the L2 learners had little difficulty in perceiving stress in the two nonsense words in the falling intonation, they were apparently poor at perceiving it in the rising intonation. A closer look into the errors committed by the two groups generates some more interesting findings, as shown in Table 3.

Table 3. Error patterns of the experimental groups

	fércept perceived as *fercépt*	*fercépt* perceived as *fércept*
Taiwanese High	26% (20/78)	74% (58/78)
Taiwanese Low	74% (61/82)	26% (21/82)

Table 3 shows that Group Taiwanese High (with at least 10 years of English learning experience) tended to identify non-words with initial stress in the rising intonation, but Group Taiwanese Low (with less than 3 years of English learning experience) tended to identify the non-words as having final stress. The chi-square test indicates that there is a significant difference between the two groups in this task (χ^2 (1) = 36.08, $p < 0.01$). This suggests that our less experienced learners made guesses of stressed syllables based on the final rising contours. In the rising intonation, both *fércept* and *fercépt* were rising on the second syllable although in the latter case the second syllable receives a low tone (L*), and the less experienced learners failed to distinguish the difference. These L2 learners seemed to take a tonal explanation towards stress when performing the experimental task and to equate higher or rising pitch with stress. On the other hand, our more experienced learners did not rely on higher pitch in determining the location of stress, suggesting that the effect of the tonal explanation may have weakened. So, what were the more experienced learners' responses based on? It is noticed that the non-words used in this experiment were carried as nouns. According to Davis and Kelly (1997), L2 learners from 15 kinds of

L1 backgrounds, including native speakers of Chinese Mandarin, were able to learn the stress typicality of disyllabic nouns and verbs in English without any difficulties (i.e., most of disyllabic nouns have initial stress and disyllabic verbs have final stress). A plausible explanation may, therefore, be that these learners were aware of the difference between lexical tone and lexical stress, but they had not yet developed the ability to detect other phonetic correlates (e.g., low rising contour vs. high rising contour) in identifying word stress, resulting in the use of an alternative strategy in the experimental task; i.e., guessing the first syllable to be stressed because it is a noun. In order to test this hypothesis, a follow-up experiment was conducted to investigate the L2 learners' responses.

2.2 Experiment 2

This follow-up experiment was to verify the strategic perception of the more experienced L2 learners reported in the first experiment, i.e., they used morpho-syntactic categories in guessing the location of stress—because non-words were embedded as nouns in the carrier sentences, they should have initial stress. In order to test this hypothesis, another pair of non-words with a stress contrast was designed and this time the non-words were put into a carrier sentence as a verb (i.e., *Do you* _____? and *Yes, I* _____.)

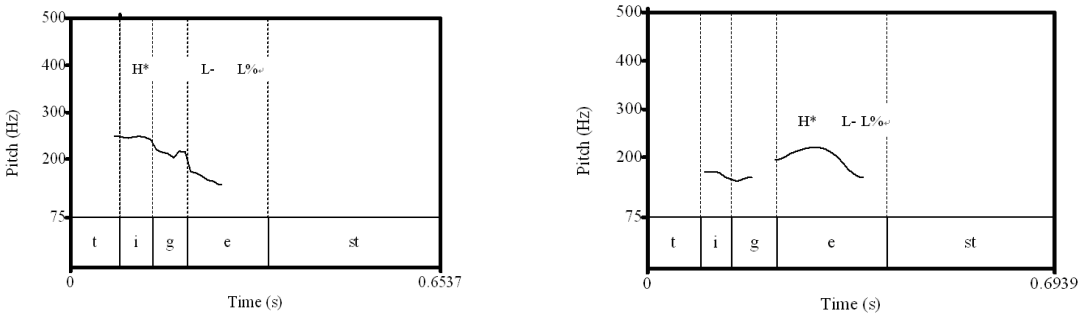
2.2.1 Materials

Another set of stress minimal pairs (*tígest* vs. *tigést*) were designed to match two pictures (i.e., Picture E and Picture F) respectively. The non-words were put into yes/no question and affirmative statement carrier sentences as verbs (*Do you* _____? and *Yes, I* _____.). The same English native speaker as in the first experiment recorded the sound stimuli 3 times. Two tokens were selected and were double-checked by 5 native speakers to ensure that the sound stimuli contrasted in their intonation patterns. The sound-picture correspondence is detailed in the appendix. Table 4 shows the means for F0, duration and intensity for the average of the two tokens of each non-word in the falling intonation.

Table 4. Phonetic measures of stressed and unstressed syllables of non-words in the falling intonation

	tígest		*tigést*	
	[ɪ] of *tí-*	[ɛ] of *-gest*	[ɪ] of *ti-*	[ɛ] of *-gést*
F0 average (Hz)	252	143	162	202
duration (ms)	60	142	49	191
intensity (db)	77	71	70	73

When the word pair is carried in the falling intonation, the pitch is higher in the first syllable of *tígest* while it is higher in the second syllable of *tigést*. This is similar to the pitch contrasts of the non-words in the first experiment. As for duration, the first syllable is shorter no matter whether it is stressed or not. This can be explained by the fact that segments have their inherent properties (Ladefoged 2005), e.g., some segments are longer than the others. As for intensity, the stressed syllables are slightly greater than their unstressed neighbors. In sum, the stressed syllable in the falling intonation is mainly signified by higher pitch and probably by greater intensity, but not by absolute duration in this word pair. Figure 5 presents the F0 contour shapes of an instance of the non-word pair carried in the falling intonation.



t ɪ́ g ɛ s t t ɪ g ɛ́ s t

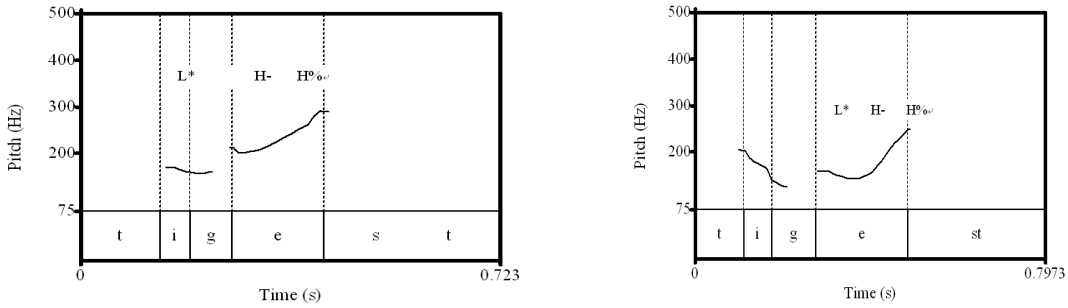
Figure 5. Pitch contours of tígest and tigést in the falling intonation

Table 5 shows the means for F0, duration and intensity for the average of the two tokens of each non-word in the rising intonation.

Table 5. Phonetic measures of stressed and unstressed syllables of non-words in the rising intonation

	tígest		*tigést*	
	[ɪ] of *tí-*	[ɛ] of *-gest*	[ɪ] of *ti-*	[ɛ] of *-gést*
F0 average (Hz)	170	262	171	180
duration (ms)	60	159	55	209
intensity (db)	74	74	72	71

In the rising intonation condition, the first syllables of the non-words do not differ in pitch because both receive low tones (i.e., an automatically associated low tone for the initial unstressed syllable vs. the nuclear low pitch accent (L*) for the initial stressed syllable). The stress contrasts can be attributed to the contour shapes of the second syllable. That is, when the second syllable is stressed, it has a lower F0; in contrast, when the second syllable is unstressed, it has a higher pitch. As for duration, the first syllable is shorter than the second one no matter whether it is stressed or not so that stress is unlikely to be cued by duration under this condition. Finally, intensity seems not significantly different in the two stress patterns, so it is unlikely to cue stress either. Figure 6 presents the contour shapes of an instance of the non-word pair carried in the rising intonation.



t í g ɛ s t t ɪ g é s t

Figure 6. Pitch contours of tígest and tigést in the rising intonation

2.2.2 Participants

We were fortunate to be able to call back 22 subjects who participated in Experiment 1 (12 of Group Taiwanese High and 10 of Group Taiwanese Low) to attend the follow-up experiment. In order to reach the requirements for statistical significance, another 18 participants were recruited (i.e., 8 undergraduates for Group Taiwanese High and 10 high school students for Group Taiwanese Low) to take part in both Experiment 1 and Experiment 2. Those who had done the first experiment only took part in the second experiment and their results from Experiment 1 were retrieved.

2.2.3 Procedure

The procedure was identical to that of Experiment 1: participants learned the mapping between sound stimuli and corresponding pictures, and then they did a forced choice test, in which they saw two pictures on the screen, then heard a sound

stimulus (either *tígest* [tíɡɛst] or *tigést* [tɪɡést]), and had to make a choice of the picture which matched what they heard.

2.2.4 Results and discussion

The prediction is that if L2 learners of Group Taiwanese High made guesses of stress in the rising intonation condition by referring to the morpho-syntactic categories, they would prefer to guess final stress in this task because the non-words were carried as verbs. If L2 learners of Group Taiwanese Low made guesses of stress according to the location of high tones, they would prefer to guess stress on the final syllable in both tasks. Figure 7 shows the overall accuracy rates for the two learner groups for the stress contrast carried by the two intonation conditions (i.e., rising and falling). Comparatively, the L2 learners were significantly poorer at differentiating stress when it was carried in the rising intonation than when it was carried in the falling intonation.

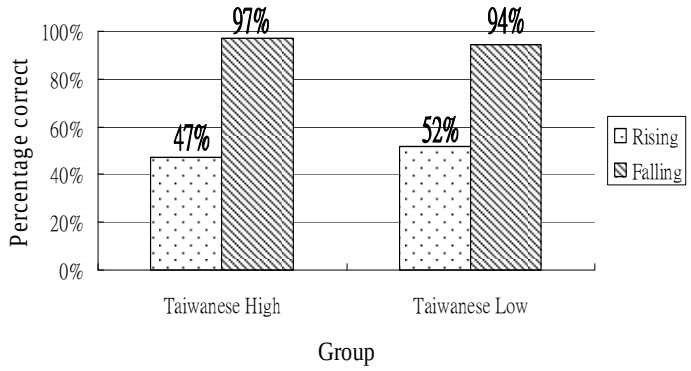


Figure 7. Overall picture performance on the stress contrast (tígest (v.) vs.tigést (v.)) by group

In comparison of the results of the two experiments (i.e., Experiment 1: *fércept* vs. *fercépt* and Experiment 2: *tígest* vs. *tigést*), an analysis of variance with the between-subjects factor Group (i.e., Taiwanese High and Taiwanese Low), and 2 within-subjects factors, Category (i.e., noun and verb) and Intonation (i.e., rising and falling), was conducted. There was a significant effect for the factor Intonation [$F(1,38) = 2281.71$; $p < 0.01$] while the other two factors were not significant (Category [$F(1,38) = 0.46$, ns.] and Group [$F(1,38) = 0.17$; ns.]). This suggests that the two experimental groups had problems in identifying word stress cued by different rising contours no matter whether the non-words were embedded as nouns or verbs. The error patterns for the two groups are shown in Table 6.

Table 6. Error patterns of the experimental groups

	tígest perceived as *tigést*	*tigést* perceived as *tígest*
Taiwanese High	84% (71/85)	16% (14/85)
Taiwanese Low	79% (61/77)	21% (16/77)

Both groups tended to choose final stress when the non-words were carried as verbs. In comparison with Table 3, it is found that Group Taiwanese High tended to choose initial stress when a non-word was embedded in the rising intonation as a noun, but they tended to choose final stress when a non-word was used as a verb. The chi-square test indicates that there is no significant difference between the two groups in this task (χ^2 (1) = 0.25, n.s.). This indicates that the more experienced participants used morpho-syntactic categories (i.e., noun and verb) as their perceptual strategy in identifying the stress position. On the other hand, Group Taiwanese Low still tended to choose the non-word with stress on the final stress in the rising intonation no matter whether the non-words were carried as nouns or verbs. This indicates that they still relied primarily on the cue of high or rising pitch to identify word stress.

3. General discussion

In the two experiments it was found that Taiwanese EFL learners were unable to perceive word stress which was cued by a difference in the high rising and low rising contours. The less experienced learners primarily relied on the cue of high or rising pitch in guessing word stress while the more experienced learners primarily relied on the information of morpho-syntactic categories in guessing the location of stress of English disyllabic non-words. The reliance on pitch height, at the expense of F0 contour shape differences, in perceiving stress may have two possible theoretical explanations. Firstly, it may show a persistent effect of the learners' L1 lexical tonal system. As mentioned earlier, Cheng (1968) reports that Mandarin Chinese speakers tend to equate weak stress in English with a low tone, carrying the feature [-high], and to treat primary and secondary stress in English as carrying the feature [+high] in the code-switching context. The tendency of Mandarin Chinese speakers to interpret word stress as tonal differences not only occurs in the code-switching context, but also translates to a tendency to identify the location of English stress using higher pitch or rising pitch contours in the setting of second language (L2) acquisition, as shown in our study. The less experienced learners' attention being drawn to F0 at the expense of the cue of F0 contour shapes impeded the development of native-like perception of lexical stress. Secondly, it is also possible that the reliance on high pitch is a general tendency for L2 learners' perceiving stress regardless of their L1 background.

Because speech prosody is a product of the phonetic realizations from the word, phrase and utterance levels, the stressed syllable is normally realized in some sort of pitch height. Due to the fact that the stressed syllable is usually realized in higher pitch when it receives the pitch-accent of high tone (H*) in English, L2 learners' attention may be drawn to this cue at the expense of the other cues. The theoretical account for Taiwanese EFL learners' reliance on pitch in perceiving stress needs further research to be more thoroughly evaluated. Therefore, this study can be expanded in a number of 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 as a primary cue to lexical contrast also show this effect in L2 English stress perception? Both questions are worth investigating in the future.

In addition, the strategic perception of the more experienced learners yields some interesting implications for L2 phonological development. That is, when learning a second language, learners can develop implicitly or explicitly the awareness that the same phonetic signal (e.g., pitch) may have a different phonological function in the target language. However, whether the native-like perception of lexical prosody in the target language can be developed is another issue. Apparently, even the participants of Group Taiwanese High in this study, with at least ten years of English learning experience, have not acquired the native-like perceptual ability.

This study provides a different conclusion from the previous view that L2 learners whose native language does not have lexical stress do not suffer from "stress deafness" at the phonetic level. Specifically, French speakers are reported to be able to discriminate between nonsense words of Spanish which differ in stress location (e.g., *mípa* vs. *mipá*), though they failed to identify them phonologically (Peperkamp et al. 1999, Dupoux et al. 2008). It is not clear what phonetic cues of stress were being used by the French speakers. This is also true of other previous studies on L2 stress perception when isolated forms or the falling intonation were used. Do those L2 learners behave like the Taiwanese EFL learners in the present study using only certain information in perceiving stress, or do they behave like the English native speakers using multiple cues of stress in perceiving stress? This question is worth investigating further as well.

4. Conclusion

This paper investigates how Taiwanese EFL learners use phonetic cues of stress in perceiving English non-word pairs which differ only in the location of stress (e.g., *férсept* vs. *fercépt*). Fifty-eight Taiwanese EFL learners were recruited to participate in

two forced choice perceptual experiments, in which they were asked to identify a perceived non-word when its stressed syllable was signified by either high pitch or different rising contour shapes. The results show that while these L2 learners had little difficulty in perceiving stress when it was phonetically realized in higher pitch, they all had great difficulty in doing so when the stressed syllable was signified by different F0 contour shapes. In addition, it is found that less experienced learners mainly relied on high pitch or rising contour in guessing the location of the stressed syllable, which may indicate a persistent effect of their L1 tonal system or a universal tendency of L2 learners' perception of stress, while more experienced learners used the information of morpho-syntactic categories as a perceptual strategy in guessing word stress, suggesting their awareness of the difference between lexical tone and lexical stress at their developmental stage.

References

Altmann, Heidi. 2006. The Perception and Production of Second Language Stress: A Cross-linguistic Experimental Study. Doctoral dissertation, University of Delaware.

Beckman, M. E. 1986. *Stress and Non-stress Accent.* Dordrecht: Foris.

Best, Catherine T. 1994. The emergence of native-language phonological influence in infants: A perceptual assimilation model. *The Transition from Speech Sounds to Spoken Words: The Development of Speech Perception*, ed. by Judith C. Goodman and Howard C. Nusbaum, 167-224. Cambridge, MA: MIT Press.

Brown, Cynthia A. 1998. The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research* 14:136-193.

Brown, Cynthia A. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. *Second Language Acquisition and Linguistic Theory*, ed. by John Archibald, 4-63. Oxford: Blackwell.

Brown, Cynthia A., and John Matthews. 1993. The acquisition of segmental structure. *McGill Working Papers in Linguistics* 9:46-76.

Brown, Cynthia A., and John Matthews. 1997. The role of feature geometry in the development of phonemic contrasts. *Generative Studies in the Acquisition of Phonology*, ed. by S. J. Hannahs and Martha Young-Scholten, 67-112. Amsterdam: John Benjamins.

Chao, Yuan Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.

Cheng, Chin-chuan. 1968. English stresses and Chinese tones in Chinese sentences. *Phonetica* 18:77-88.

Cheng, Chin-chuan. 1973. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

Davis, S. M., and M. H. Kelly. 1997. Knowledge of the English noun-verb stress difference by native and nonnative speakers. *Journal of Memory and Language* 36:445-460.

Dupoux, Emmanuel, Sharon Peperkamp, and Núria Sebastián-Gallés, 2001. A robust method to study stress "deafness". *Journal of the Acoustical Society of America* 110:1606-1618.

Dupoux, Emmanuel, Núria Sebastian-Gallés, Eduardo Navarrete, and Sharon Peperkamp. 2008. Persistent stress "deafness": The case of French learners of Spanish. *Cognition* 106:682-706.

Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. by W. Strange, 233-273. Baltimore: York Press.

Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9:317-323.

Guion, Susan. 2005. Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition* 27:503-533.

Kijak, Anna. 2007. Does regularity of L1 stress matter in L2 perception? A typological study. Poster presented at the 16th European Second Language Association Conference, Antalya, Turkey.

Ladefoged, Peter. 2005. *Vowels and Consonants*, 2nd edn. Malden, MA : Blackwell.

Ladefoged, Peter. 2006. *A Course in Phonetics*, 5th edn. Boston, MA: Thomson Wadsworth.

Ou, Shu-chen. 2006. Factors and Mechanisms in L2 Word Stress Acquisition: Evidence from Chinese-English Interlanguage. Doctoral dissertation, University of Edinburgh.

Ou, Shu-chen. 2007. Linguistic factors in L2 word stress acquisition: A comparison of Chinese and Vietnamese EFL learners' development. *Proceedings of the 16th ICPhS*, 1681-1684. Retrieved September 30, 2007, from www.icphs2007.de.

Pater, J. 1997. Metrical parameter missetting in second language acquisition. *Focus on Phonological Acquisition,* ed. by S. J. Hannahs, and Martha Young-Scholten, 235-261. Amsterdam: John Benjamins.

Peperkamp, Sharon, Emmanuel Dupoux, and Núria Sebastian-Galles. 1999. Perception of stress by French, Spanish, and bilingual subjects. *Proceedings of Sixth European Conference on Speech Communication and Technology*:

2683-2686. Retrieved December 30, 2008, from www.isca-speech.org/archive/eurospeech_1999.

[Received 28 September 2009; revised 17 December 2009; accepted 6 January 2010]

Department of Foreign Languages and Literature
National Sun Yat-sen University
Kaohsiung, TAIWAN
Shu-chen Ou: sherryou@faculty.nsysu.edu.tw

Appendix

The sound-picture correspondence in Experiment 1 and its learning phase:

(1) Stress contrast

Picture A
fércept [fɜ́ˠsɛpt]

Picture B
fercépt [fəˠsépt]

(2) Segmental contrast

Picture C
tóoper [túpəˠ]

Picture D
tóoker [túkəˠ]

The sound-picture correspondence in Experiment 2 and its learning phase:

Picture E
tígest [tígɛst]

Picture F
tigést [tɪgést]

Participant doing experimental task:

台灣英語學習者之字重音感知

歐淑珍
國立中山大學

本篇論文探討在音高語音訊號被操弄的狀況下，台灣英語學習者如何感知英語字重音對比（如：*fércept* vs. *fercépt*）。五十八位英語學習者參與兩個強制性選擇的感知實驗，其任務是指出所聽取的音檔所對應的假字圖片爲何。其中，假字的重音節以兩種語調方式呈現：(i) 較高的音高或（ii）低升的音高曲線。結果顯示，台灣英語學習者對以「較高的音高」的語音線索來呈現的重音沒有感知的困難，但對用「低升的音高曲線」來呈現的重音卻有極大的感知困難。此外，錯誤分析顯示初階學習者主要仰賴在「上揚音高」的語音訊號來猜測重音節所在，這有可能是來自母語聲調系統的影響或是第二語言學習者重音感知的共通傾向。相對地，高階學習者則使用詞類做爲猜測重音所在的策略，顯示他們在該中介語發展階段已具有「聲調不同於重音」的音韻知覺。

關鍵詞：跨語言韻律感知；第二語言字重音；聲調-重音中介語