THE TRANSFER OF L1 ACOUSTIC CUES IN THE PERCEPTION OF L2 LEXICAL STRESS

Qian Wang, Tae-Jin Yoon
Department of Linguistics, University of Victoria, Canada; wangqian@uvic.ca; tyoon@uvic.ca

1. INTRODUCTION

In second language (L2) acquisition, first language (L1) transfer is observed at different levels, e.g. syntax, morphology, and also phonology. For example, L2 learners' problems with English lexical stress were attributed to the difference between L1 and L2 phonological systems [1]. The present research aimed to explore the nature of such problems from a different perspective, the nonnative use of acoustic cues in stress acquisition. It is proposed that L1 transfer also operates at the level of acoustics, i.e. the transfer of acoustic cue reliance.

It is known that at least three acoustic correlates contribute to English lexical stress perception, i.e. F0, duration and intensity. The acquisition of stress involves the correct use of *all* three cues. Nonnative speakers, even with the correct phonological representation of the L2 stress system, may not be able to employ the correlated cues in a way similar to L1 speakers. The present study focused on the use of acoustic cues by Chinese learners of English in stress perception. Chinese is a tone language, and F0 plays a key role in shaping the L1 phonological system, while the contribution by duration and intensity is marginal. As the acoustic cues of F0, intensity, and duration are used differently in L1 and L2, we expect to find the heavy reliance on F0 by CE will be transferred to the perception of English lexical stress. Two perceptual experiments were conducted to explore the hypothesis. The first experiment involved the stress judgment of synthesized disyllabic nonsense tokens and the second one was an oddity test using real English words extracted from carrier sentences.

2. EXPERIMENT ONE

In the first experiment, F0 difference on disyllabic nonsense words was manipulated in 5 levels, while duration and intensity cues were kept consistent on the both syllables. If there is a positive transfer of F0 from the L1, then Chinese learners would have no problems in perceiving stress on words with only F0 cue.

2.1 Research design

Three disyllabic nonsense words were created, *latmab, nizdit,* and *tetsep*. The nonsense words were produced by a trained phonetician as the basis for manipulation. The two syllables of each nonsense word were first manipulated to have the same duration and intensity. Then, F0 difference between the two syllables was changed in five levels (see Table 1).

Table 1: 5 levels of F0 manipulation

5 levels	1	2	3	4	5
F0 difference: s1 –s2	-50	-25	0	25	50

For example, in the first level of manipulation, the first syllable is 50Hz lower than the second syllable and in the third level, the two syllables have the same F0. This resulted in a total of 15 tokens were created (5 levels * 3 word forms).

Two groups of participants listened to the tokens, an experimental group of 68 CE and a controlled group of 38 NE. They indicated the stress position of each token by clicking **1st Syllable** or **2nd Syllable** on a computer screen.

2.4 Data analysis and results

The number of Initial Stress (IS) judgments (i.e. stress on the first syllable) was collected for each token. The Initial Stress Percentage (ISP) was defined as the percentage of IS judgments over the total number of participants for each token. ISP was calculated for the two groups separately. For example, an ISP of 50% by NE means that half of the NE participants agreed on the IS judgment.

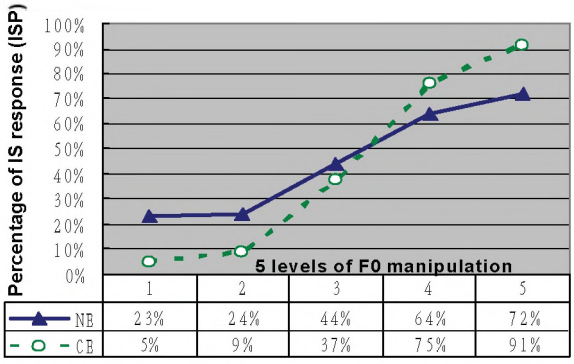

Figure 1: ISP of CE and NE as a function of F0 manipulation

The ISP values of NE and CE groups, respectively, are summarized for each level of F0 manipulation (see Figure 1). The dotted line in Figure 1 shows that CE demonstrated a systematic ISP change as a function of F0 manipulation. This change is in consistence with the change induced by NE (the solid line). A 2 (language group) × 5 (F0 manipulation level) mixed-model ANOVA on ISP showed that there is a main effect for F0 manipulation ($F(4,16) = 39.99$, $p < .001$, $\eta^2 = .909$). No significant effect of group was found ($F(1,4)=0.010$, $p>0.9$). The interaction between F0 and language group was also insignificant ($F(4,16)=2.858$, $p>0.1$). It can be concluded that F0 cue is sufficient for CE to perceive stress. However, we can't conclude that CE behaved in the same way as NE did. If we compare the slope of two lines above, we can see that for NE, F0 change from level one to five led to an increase of 49% in IS responses (23% to 72%). For the CE group, the same F0 change led to a much greater increase of 86% in IS responses in CE group (5% to 91%). This steeper slope of CE may suggest that their's reliance on F0 exceeds the native usage of F0.

3. EXPERIMENT TWO

If CE relies excessively on F0, then we can expect the negative effect when F0 information is absent. That is, CE would perform significantly worse than NE. The second experiment is designed to test this hypothesis.

3.1 Research design

In the second experiment, we chose one word form with two possible stress patterns: PERmit as a noun and perMIT as a verb (where capital letters indicate the stressed syllable). Two versions of either forms were constructed, one with F0 difference between the stressed and unstressed syllable cue and one without such difference. This was realized by extracting target words from two focus contexts: one where the target word is in the focus of the sentence, hence accented, and the other where the target word is not focused and thus unaccented, i.e., where another content word in the sentence is under focus. Previous literature [2] have shown that, the stressed syllable on an accented word in a sentence is reinforced by F0 cue, but stress on an unaccented word is usually cued by duration and intensity, but not F0. As a result, four types of stimuli (2 stress patterns * 2 accent conditions) were constructed and used in an oddity test. Participant listened to 3 tokens in a triad and decided whether the tokens are all the same, i.e. all nouns or all verbs, and if not, which token was the odd one. A total of 64 triad were constructed with two different accent conditions, half of the triads were realized with all three **A**ccented tokens (AAA), and half with all **U**naccented tokens (UUU). Two groups of participants (11 CE and 13 NE) were involved in the second experiment,.

3.2 Data analysis and results

Error rates were collected from each participant and used in a 2 (language group) × 2 (accent condition) mixed-model ANOVA. The results revealed the main effect for language group ($F(1, 21^{a}) = 10.747$, $p < .01$, $\eta^2 = .339$). There was an overall difference in the error rate of CE (M=19.89%) compared to NE (M=6.51%). A significant main effect for the accent condition ($F(1, 21)=27.61$, $p<.001$, $\eta^2 = .57$), and a significant interaction of language group by accent were also observed ($F(1,21)=14.32$, $p<0.01$, $\eta^2 = .405$). Two-tailed independent *t*-tests between NE and CE for the two accent conditions, respectively, indicated that the significant effect of group actually lies in UUU condition ($t(21)=3.54$, $p<.01$) but not in AAA ($t(21)=.61$, $p>.5$) (see Figure 2 below). In other words, when F0 information is present, CE resembles NE in stress perception, which is consistent with the results of the first experiment. However, when F0 information is absent, CE performed significantly worse.

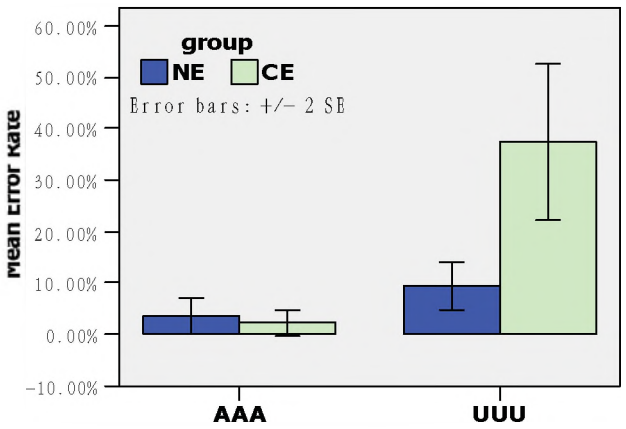

Figure 2. Mean error rates as function of accent effects by two language groups. Two types of accent combination are presented on the x-axis and error rates are represented on the y-axis. The performance of NE is indicated by blue bar, and that of CE is represented with green bar. Error bars are shown.

[a] One NE subject was identified as an outlier.

4. DISCUSSION

The results of the two experiments suggest that the difference between CE and NE in stress perception lies in their reliance on different of acoustic cues. CE rely heavily on F0 information and they don't have problems in stress perception if F0 information is provided (Experiment One). Unlike NE, they could make little use of other acoustic correlates that contribute to stress perception (Experiment Two). Further analysis through paired-sample *t*-test in the second experiment showed that while NE's performance in UUU condition is worse than in the AAA condition, the difference is not significant ($t(11)=2.421$, $p> 0.01$). CE, on the contrary, performed significantly worse in UUU condition then in AAA ($t(10) = 4.553$, $p < 0.01$). In UUU condition, NE can rely on the acoustic correlates of duration, intensity and possibly other correlates in stress perception when they are deprived of the F0 information. This is not the case for CE, who have a tonal language background. For CE, F0 is decisive and other cues may not be incorporated into the L2 phonological system and actively exploited in the construction of prosodic elements. Studies with Vietnamese learners and Japanese learners of English have found similar results [3,4]. As Adams and Munro [5] commented, although non-native speaker may not lack the acoustic cues for the appropriate prosodic elements, they nevertheless often fail to use these cues in a native-like way to achieve satisfactory prosody in a second language.

REFERENCES

[1] Archibald, J. (1997b). The acquisition of english stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics, 35*, 167-181.

[2] Sluijter, A.M.C. & van Heuven, V.J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100(4), 2471-2485.

[3] Nguyen, T.A.T. (2003). *Prosodic Transfer: The Tonal Constraints on Vietnamese Acquisition of English Stress and Rhythm.* Doctoral dissertation, University of Queensland, Queensland, Australia.

[4] Ueyama, M., *Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese.* 2000, University of California, Los Angeles.

[5] Adams, C. & Munro, R.R. (1978). In search of the acoustic correlates of stress: Fundamental frequency, amplitude, and duration in the connected utterances of some native and non-native speakers of English.