

Lexical Stress Realisation: Native vs. ESL Speech

Hua-Li Jian

Department of Foreign Languages and Literature
National Cheng Kung University, Tainan, Taiwan
hljian@mail.ncku.edu.tw

Abstract

English stress placement in phrase-medial and phrase-final position is investigated. Current results indicate that Taiwanese ESL learners realise polysyllabic words that carry various degrees of stress in two prosodic positions with considerable differences relative to the native American English speakers, and the differences are demonstrated from acoustical and phonetic perspectives.

1. Introduction

Different languages exhibit noticeable variation of typical placement of stressed syllables within a word. Some languages are characterised by fixed stress position, such as Welsh, in which stress almost always falls on the penultimate syllable. Other languages, including English, manifest movable or free stress. Several studies have also addressed stress placement of English words by learners of English as a second or foreign language. For instance, Davis and Kelly [5] found that Japanese and Mandarin ESL learners are both aware of the stress placement contrast between English nouns and verbs. In contrast, Archibald [1] reported that Chinese ESL learners in general assign stress unsystematically. This study examines how Taiwanese (TW) ESL learners realise stress placement in speech production of polysyllabic words where the primary stress occurs on the first syllable and where the same test words are placed in two prosodic conditions, namely, nuclear, intonational phrase-final position and in pre-nuclear non-final (or medial) position.

2. Background

Previous research has examined possible ways of signalling the intonational boundary in English. One way of indicating an intonational boundary in British English is to lengthen the final syllable in a phrase and simultaneously change the pitch contour on lexically unstressed syllables, i.e. alter the pitch level, or change from falling to rising or vice versa ([3], [4], and [8]). Cruttenden [4] further indicated that the lengthening of the final syllable is a feature that can be found in many or most languages, but that the amount of lengthening is language-specific. As shown in a study of Arabic words [6], a typologically similar language to English in terms of speech rhythm, speakers illustrate extensive word-

final lengthening (on the order of 60%) and utterance-final lengthening effects (on the order of 100%), which are smaller than those found for stressed, phrase-final syllables in English [2]. Low and Grabe [10], however, did not find similar final syllable lengthening effect in polysyllabic words by British English speakers where the primary stress is located at the initial syllable in either phrase-final or phrase-medial position. What Low and Grabe did identify was a final syllable lengthening phenomenon that occurred at the phrase-final position articulated by the Singaporean English speakers. It is therefore natural to expect that English spoken by TW subjects may show a different lengthening effect to American (US) subjects on the final syllable in a phrase, in addition to possible differences of the first two syllables of tri-syllabic words between the two speech samples.

Finally, Low and Grabe [10] pointed out that a pitch displacement often can be found on the stressed or accented syllable in British English speech compared to the following non-stressed syllables in the polysyllabic word, and that this displacement is less visible in the speech produced by Singaporean subjects. Consequently, this can strengthen the perception of word-final stress in intonation phrase-final position. Further, if pitch is not used to signal stress among Singapore English speakers, then listeners will have to more heavily rely on syllable durations to identify stress. One objective of this work is to test whether the same observations hold for the English spoken by TW ESL learners versus US English. In this study, the aim is to contrast the placement of English lexical stress articulated by native US English speakers with that by TW ESL learners. One can hypothesise that there exists little or no difference in lexical stress placement of the two groups; in other words, there is no shift of prominence. Further, what makes the two varieties different is the way the primary stress and the following unstressed syllables are realised acoustically. Since TW subjects are native tonal-language speakers, the way they employ suprasegmental features to realise stress or nuclear accents in polysyllabic words in a stress-timed language such as English may also differ from native non-tonal English speakers. Furthermore, in addition to reflect phonemic vowel category, vowel quality may reflect effects imposed by prosodic factors such as stress or length [9],

and hence vowel quality is examined along with other suprasegmental properties, in particular f0 and duration.

3. Method

3.1. Subjects and materials

Ten subjects participated in the experiment, namely three male and two female native TW students in their mid twenties and three male and two female US English teachers ranging in age from 20 to 40. The materials, which are inspired by the ones employed by Low and Grabe [10], consisted of 20 sentences for each subject, i.e. a total of 200 sentences. Each subject read 10 sentence pairs. Each sentence pair comprised first one sentence where a polysyllabic word (three syllables) in question appears in phrase-final position and one sentence where the same word appears in phrase-medial position. The rationale behind this approach is that polysyllabic words in phrase-final position are subject to intonation phrase boundaries. The test words in phrase medial position should be less affected by the effect. The polysyllabic words were selected to ensure that US English subjects would place the stress on the initial syllable and stress-neutral suffixes were used to avoid stress shift. The syllables immediately preceding the test items in phrase-final position are mostly function words as they are usually unstressed. Similarly, syllables immediately following the test items in phrase-medial position are usually unstressed to prevent possible shift of stress.

3.2. Procedure

The subjects were digitally using a Shure KSM32 microphone and a Sony minidisk recorder. The subjects were allowed to familiarise themselves with the text, printed on a sheet of paper and they were asked to practice once. The subjects were asked to repeat the pair of sentences twice during the recording and the second version of each sentence was used.

3.3. Measurements and analysis

The duration of each of the three syllables in the target word was measured for both the sentence-final position and sentence-medial position. These durations were found by establishing the boundary points between the three syllables. Next, the mean duration for each of the three syllables for sentence final and sentence medial tri-syllabic words were calculated for the TW and US subjects respectively. In addition, the f0 endpoint values approximately representing the f0 contour of each syllable were extracted. Only the voiced part each syllable was considered to achieve reliable readings. Further, as the f0 measurements are often inaccurate at the start and end of the voiced parts, a mean with a small window size was used. The mean f0 was computed for

the endpoints of each syllable for each subject group. Finally, F1 and F2 formant values were extracted based on the midpoints of the vowels of the syllables as vowels are characterised by their formant values. An un-paired two-tailed t-test was used to test for significance in the data using the analysis tools in Microsoft Excel 2003 with a significance level of 0.05.

4. Results and discussion

4.1. Duration

Figure 1 shows the mean duration for polysyllables in phrase-final position. The bars depict means for syllables in first, second and third position for both subject groups. Figure 2 shows matching results for polysyllables in phrase-medial position.

In phrase-final position, the TW speakers have significantly longer initial syllable durations than the US natives ($t=4.338$; $p<.001$; $df=57$). As for the following two syllables, both sets of measurements exhibit similar durational patterns. The duration decreases from the penultimate to the final syllable in both groups, with TW subjects showing longer mean duration.

When the tri-syllabic test words occur in phrase-medial position, both speech groups show similar patterning as in phrase-final position. There is greater difference between the two varieties in the final syllable in pre-nuclear medial position than in nuclear phrase-final position. The final syllable is the shortest and the first syllable longest in both speech groups. A t-test reveals a significant difference in the final syllable in both groups ($t=4.441$; $p<.001$; $df=80$). The US set has the final syllables with much shorter durations than the TW group. This is also true in phrase-final position.

A similar trend was also found in the study by Low and Grabe in that British speakers show much shorter final syllables in phrase-medial position. In the current data, the TW speakers tend to lengthen the first syllable, as is also the case in the phrase-final position. This lengthening is more visible than for the US group. The data shows that TW learners tend to articulate the first syllable of tri-syllabic word with significantly greater duration than the US natives ($t=2.762$; $p<.01$; $df=46$), which is quite the opposite of the findings by Low and Grabe, where they found that the Singaporean English speakers did not have as long first syllables as the British English speakers. Instead, Singaporean English speakers tend to lengthen the final syllable at phrase-final position. It may be that TW subjects signal the primary stress of multi-syllabic words by means of lengthening that syllable, with some degree of difference at phrase-final versus phrase-medial position, employing more lengthening (of the first syllable) at intonational nuclear position than at the pre-nuclear position.

4.2. F0

Figures 3, 4, 5 and 6 show the mean f0 traces for Taiwan and US subject's polysyllables in phrase-final and phrase-medial position respectively. The measurements of the f0 in phrase-final position show an overall decrease over time in both speech groups.

There is a significant step-up of f0 between the offset of the initial syllable and the onset of the penultimate syllable in TW group but not for the US group ($t=.564$; $p>.05$; $df=26$). This finding is different from Low and Grabe's study where the British English group shows greater step-down between the offset of the initial syllable and the onset of the penultimate, while the f0 contours for Singaporean English speakers show a small step-up of f0 between the same ranges.

For the other transitional area between the f0 contours, that is, between the offset of the penultimate and the onset of the final syllable, both speaker groups show a step-up in pitch, with the one for the Taiwan group being larger in magnitude though both are statistically insignificant (TW: $t=.77$; $p<.05$; $df=86$; US: $t=.34$; $p<.05$; $df=40$). For the penultimate syllable, TW and US groups both show a drop in f0 (i.e., falling pitch).

4.3. Formants

Stress has a tendency to increase the differences in formants between vowels [6] and that greater stress increases f1 [7]. Figures 7, 8 and 9 illustrate the formant patterning of both speaker groups in the various syllables and the two prosodic positions. The vertical axis represents the second formant (F2) based on the mean values extracted, and the horizontal axis displays the mean first formant (F1) values.

The two variety groups (TW and US) demonstrate different formant patterns for different syllables. For the first syllable, the US group shows higher f2 but lower f1, while Taiwan subjects have higher f1 and lower f2. Both sets show similar direction of change from medial to final position (or from final to medial). The contrastive formant region, but with similar transitional patterning in the first syllable of both groups, illustrates that TW speakers realise the primary stressed syllable in the two prosodic conditions in a similar way as employed by the US natives, but with different vowel quality.

US speakers articulate the penultimate syllable that contains the reduced vowel with a higher f2 but a lower f1, compared to TW speakers' higher f1 but lower f2. However, the US subjects have their f2 reduced from the medial to final position, while the TW learners have both f1 and f2 reduced towards the final position. The two groups both occupy similar formant ranges as in the first syllable. For the US subjects, there is a reduction in

f1 for both prosodic conditions from the first primary stressed syllable to the penultimate unstressed syllable, as found in de Jong's study [7], but the same phenomenon is less obvious for the TW subjects. As for the final syllable, the f1 decrease in both prosodic conditions and for both subject groups seems all the more obvious, suggesting the lowest degree of stress in the multi-syllabic word-final syllable in the current data. Additionally, both speaker groups appear far from each other at the medial position, with high f1 and low f2 in the Taiwan set and high f2 and low f1 of the US set, they then draw back to a similar region at the final position. This unusual formant patterning in the final syllable possibly illustrates the different way or quality the Taiwan subjects, as compared to the native US English speakers, end the multi-syllabic test words in a pre-nuclear intonational phrase, which is different from how one signals a word ending in the intonational phrase-final position.

5. Acknowledgements

This work is supported by the National Science Council of Taiwan grant NSC93-2411-H006-021.

6. References

- [1] Archibald, J. (1997). The acquisition of English stress by speakers of nonaccidental languages: Lexical storage versus computation of stress. *Linguistics*, 35, 167-181.
- [2] Beckman, M. E., & Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In *Papers in laboratory phonology I* (J. Kingston & M. E. Beckman, editors), pp. 152-178. Cambridge, UK: Cambridge University Press.
- [3] Cooper, W. E., & Paccia-Cooper, I. (1980). *Syntax and Speech*, Cambridge, MIT Press.
- [4] Cruttenden, A. (1997). *Intonation* (second ed.), Cambridge, U.K., Cambridge University Press.
- [5] Davis, S. M., & Kelly, M. H. (1997). Knowledge of the English noun-verb stress difference by native and nonnative speakers. *Journal of Memory and Language*, 36, 445-460.
- [6] de Jong, K., & Zawaydeh, B.A. (1999). Stress, Duration, and Intonation in Arabic Word-level Prosody, *Journal of Phonetics*, 27, pp. 3-22.
- [7] de Jong, K. (2004). Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration, *Journal of Phonetics*, 32, pp. 493-516.
- [8] Hofhuis, E., Gussenhoven, C., & Rietveld, A. (1995). Final lengthening at prosodic boundaries in Dutch. *ICPhS 13*. Stockholm: KTH and Dept Linguistics, University of Stockholm. 154-157.
- [9] Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.
- [10] Low, E. L., & Grabe, E. (1999). A Contrastive Study of Prosody and Lexical Stress Placement in Singapore English and British English, *Language and Speech*. 42(1), 39-56.

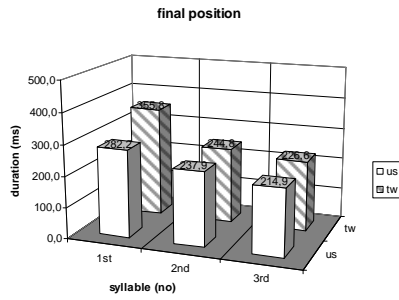


Figure 1. Syllable duration characteristics for polysyllables in phrase-final position. Error-bars show standard deviation.

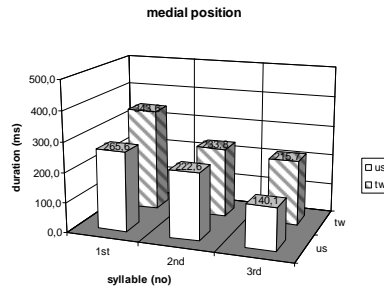


Figure 2. Syllable duration characteristics for polysyllables in phrase-medial position. Error-bars show standard deviation.

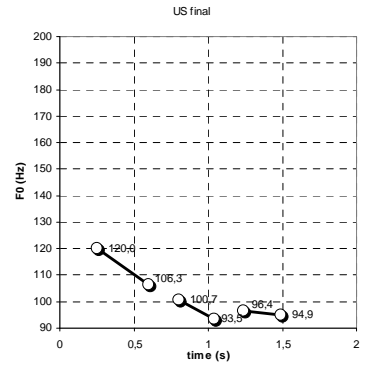


Figure 3. Mean f0 traces for the first, second and third syllables of US-English polysyllables in phrase-final position.

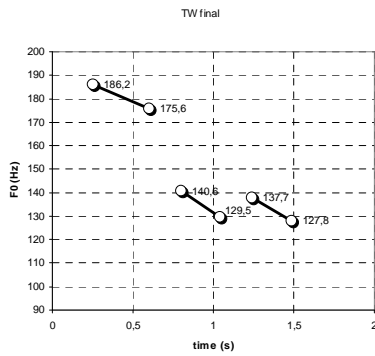


Figure 4. Mean f0 traces for the first, second and third syllables of Taiwan-English polysyllables in phrase-final position.

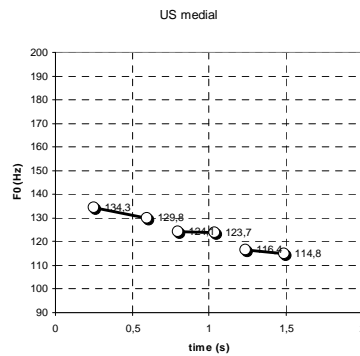


Figure 5. Mean f0 traces for the first, second and third syllables of US-English polysyllables in phrase-medial position.

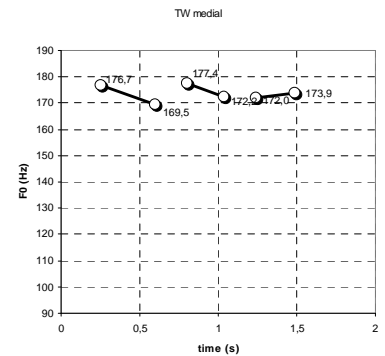


Figure 6. Mean f0 traces for the first, second and third syllables of Taiwan-English polysyllables in phrase-medial position.

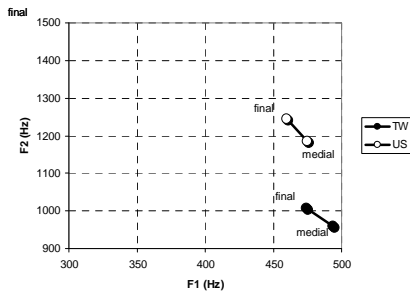


Figure 7. Mean f1/f2 formant space for the first syllable of tri-syllabic words in both phrase-final and phrase-medial positions.

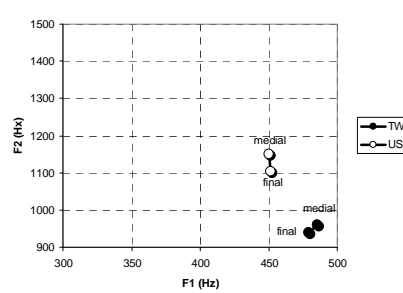


Figure 8. Mean f1/f2 formant space for the penultimate syllable of tri-syllabic words in both phrase-final and phrase-medial positions.

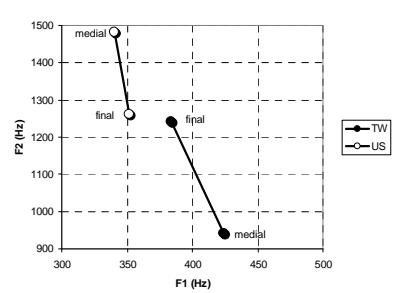


Figure 9. Mean f1/f2 formant space for the final syllable of tri-syllabic words in both phrase-final and phrase-medial positions.