Are Nuclear Accents Easier to Acquire than Prenuclear Accents? — Using Peak Alignment in Advanced Mandarin EFL Learners as an Example

Hui-Yu Chien and Janice Fon

National Taiwan University
r06142011@ntu.edu.tw; jfon@ntu.edu.tw

ABSTRACT

This study investigated how advanced EFL learners of Taiwan Mandarin align accent peaks in their L1 and L2, as compared to their native English counterparts. Twenty-two advanced Mandarin EFL learners and 22 native American English speakers read 21 English monosyllabic stimuli embedded in a declarative carrier sentence. The EFL learners also read an additional list of 21 sentences in which phonotactically-matched Mandarin stimuli were embedded. Results showed that EFL learners aligned pitch peaks in Mandarin earlier than those in English in the prenuclear position but later in the nuclear position. Moreover, EFL learners' relative peak alignment of English H* pitch peaks were comparable to that of their native English counterparts in both prenuclear and nuclear accents. However, there was a tendency for EFL learners to realize the prenuclear accent as L+H* instead of the default H*. Such difference is not likely due to L1 transfer, but insufficient L2 proficiency.

Keywords: speech prosody, L2 acquisition, tonal alignment

1. INTRODUCTION

Prosody is an important element in successful speech communication. However, its acquisition has been especially challenging for second language (L2) learners. Previous studies have shown that L2 intonational deviations from the target norm range from pitch accent alignment [3, 5, 9], to pitch range realization [2, 17, 18, 20], to phrasing [6], among others. This study focuses on the alignment of pitch accents. Peak alignment is one of the essential issues discussed in suprasegmental research [10, 11]; Previous studies have shown that languages may have their own preferred characterizing peak alignment locations. For example, English distinguishes between early and late peak alignments [14, 15]; the peak alignment of prenuclear accents in Dutch is determined by the structure of the accented syllable [7]; the peak alignment of the Mandarin rising tone is near the end of syllable boundary [19]. Different tone types may also be manifested mainly by the timing of peak alignments. For example, according to the prosodic labeling system of English Tone and Break Indices (ToBI) [1], one of the defining characteristics between an L+H* and an L*+H tone is in fact timing, with the H tone in the former being aligned earlier than that of the latter. In other words, pitch accent alignment is influenced by both linguistically intrinsic factors such as pitch accent types, and extrinsic factors such as language preferences. In L1 acquisition, the extrinsic factor may not pose as a difficulty, as children should naturally acquire the preferred alignment location of a given tone in their native language. However, in an L2 setting, when learners happen to come from a language that has a different preference for tonal alignment, one would expect it to be more challenging. The difficulty might be somewhat analogous to what was found for VOT in segmental studies, in which different voiceless-vs.-voiced divides are preferred by different languages [4].

In this study, Mandarin learners of English were chosen as the target of study. This language pair is especially interesting for examining pitch accent alignment, as tonal alignment pattern in Mandarin remains stable in spite of contextual differences [19] while English seems to have more variability in alignments according to other factors such as speech rate [16] and dialectal variation [8]. Whether this will affect learners' English is a topic worth investigating, as previous studies have not yet been unanimous on this. For example, Lu and Kim [9] found that the alignments produced by Mandarin speakers of English was much later than those by their English counterparts. As Mandarin indeed tended to prefer a later peak alignment [19], they attributed this to L1 transfer. However, since Lu and Kim [9] have not included Mandarin in their study, it is difficult to ascertain whether it was indeed due to L1 transfer or it was caused by insufficient glottal-supraglottal coordination in one's less proficient language. In the work of Chen and Fon [3] on early Mandarin-English bilinguals, they also found later peak alignment among Mandarin learners, but only in prenuclear accent positions. For nuclear accents, no difference was found. As nuclear accents by definition occur later in time than prenuclear ones, it is possible that a differential degree of articulatory preparedness might be involved. In addition, since the Mandarin learners in

Chen and Fon [3] were early near-native ESL learners arising from an immersion environment, while those in Lu and Kim [9] were late nonnative EFL learners living in an immersion environment, there may be some substantial differences between the two groups. Therefore, in this study, we focused on advanced yet nonnative EFL learners in Taiwan to see more clearly the potential learning trajectory of pitch alignment.

2. AIMS OF THE STUDY

There are two specific aims in this study. First, we would like to investigate whether advanced Mandarin EFL learners in Taiwan would align their English tones later than their native counterparts in general, as was suggested by Lu and Kim [9], or whether they would show a learning edge towards nuclear accents and master them earlier, as was implied in Chen and Fon [3]. Secondly, we would also like to examine the claim about Mandarin tonal alignment being later than English, so as to see whether L1 transfer is in fact at work. If learners align their Mandarin and English in a similar fashion, then the L1 transfer hypothesis can be more substantiated. On the other hand, if there is little correlation between learners' alignment in Mandarin and English, then other factors might be at work in the learning process.

3. METHOD

3.1. Participants

Two groups of speakers were recruited in this study. The first group included 22 native Mandarin EFL learners in Taiwan (11 males and 11 females), aged 18-30. All of them were considered advanced learners, as they were either English majors in college and/or had obtained an advanced level of CEFR B2 or C1 in an English proficiency test of CEFR B2 or C1. The second group included 22 native American English speakers from the U.S. (11 males and 11 females), aged 18-30, who served as native controls.

3.2. Material

The reading materials were 21 English declarative sentences of a simple SVO structure in the form of *I know the word X*. Words of a CV or CVC syllable were embedded as the last word in the sentence. In addition, 21 Mandarin declarative sentences were included as reading materials for the EFL learners. These were comparable sentences of *Zhege zi nian X* 'This character reads X'. Phonotactically matched Mandarin Tone 4 (i.e., high-falling) syllables were embedded as the last word in the sentence. The sentences were printed on individual index cards for recording.

3.3 Equipment and recording procedure

A KORG DAT MR-1000 digital recorder along with a SHURE SM10A head-mounted microphone were used for the recording. All participants were recorded individually in a soundproof recording room. The participants were asked to read aloud the sentences at a normal speed and in a natural fashion. They were asked to repeat if there was any pause, cough, or slip of tongue occurring in the middle of their production. The native speakers read only the English sentences while the EFL learners read the additional Mandarin sentences after they read the English sentences. In each session, all 21 sentences were randomized and each speaker had a different randomization order.

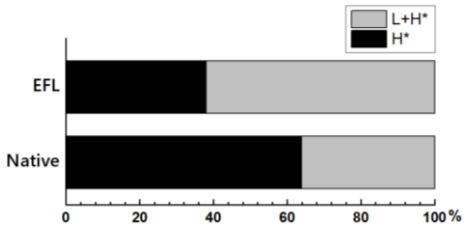
3.4 Measurement

The English sentences were labeled using the English ToBI system [1]. Both nuclear and prenuclear accents were identified, and their pitch peaks were extracted using Praat [12]. The Mandarin sentences were labeled according to the Pan-Mandarin-ToBI system [13], and pitch peaks of stressed syllables (mostly on *zi* 'character' and the target word) were also extracted. Syllable boundaries were determined based on visual inspection of the spectrogram.

4. RESULTS

Figure 1 shows the choice of pitch accent tone types among EFL learners and native speakers in English prenuclear accents.

Figure 1: Distribution of pitch accent tone types for English in the prenuclear position among EFL learners and native English speakers



Interestingly, tone type preferences were quite different between EFL learners and native speakers. Native speakers preferred to use H* more while EFL learners preferred to use L+H* more [$\chi^2(1) = 56.67$, $p < .0001$]. This indicates that for prenuclear

accents, H* tone seemed to be more of a default choice for native speakers: of all the 396 tokens, 253 (64%) of them were realized as H* and 143 (36%) as L+H* [$\chi^2(1) = 15.58$, $p < .0001$]. As for EFL learners, they tended to realize the prenuclear accent as L+H* instead of the default H*: of all the 455 tokens, 173 (38%) were realized as H* and 282 (62%) as L+H* [$\chi^2(1) = 13.26$, $p < .001$].

4.1. Prenuclear accents

Figures 2 and 3 show the alignment of prenuclear accents realized as H* and L+H*, respectively. The timing of H* occurrence in total syllable duration are also shown in the bar chart of Figure 4. Figure 5 shows the speech rate (syllable per second) of both prenuclear and nuclear accents among the two groups of speakers. In general, EFL learners showed different alignment preferences for their Mandarin and English, with the former being much earlier than the latter (the high tone and H* occurred at about 51.78% and 85.86% of the total syllable duration, respectively; [$t(29) = -6.84$, $p < .0001$]). EFL speakers were also found to utter their Mandarin significantly faster than English (5.24 syll/s and 3.89 syll/s for Mandarin and English, respectively; [$t(21) = 5.29$, $p < .0001$]. It is also interesting to find that, in both H* and L+H* cases, EFL learners indeed placed their English accent peaks later than their native counterparts, yet the difference was mainly due to their slower speech rate (3.89 syll/s and 5.42 syll/s for EFL and native speakers, respectively [$t(42) = -4.12$, $p < .0001$]). The relative location for pitch alignment was rather comparable, as shown in Figures 2 and 3. In the case of H*, the peaks occurred at about 85.86% and 79.64% of the total syllable duration in EFL and native speakers, respectively [$t(21) = 0.69$, ns]. As for L+H*, the peaks were placed at about 98.09% and 99.46% of the total syllable duration in EFL and native speakers, respectively [$t(20) = -0.21$, ns]).

Figure 2: Alignment of prenuclear accent realized as H* among EFL learners and native English speakers. MN: Mandarin; EN: English.

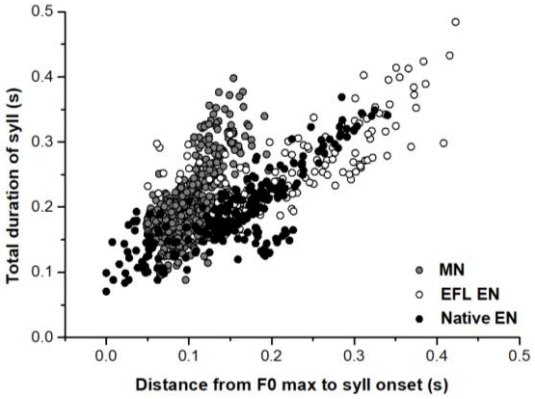


Figure 3: Alignment of prenuclear accent realized as L+H* among EFL learners and native English speakers.

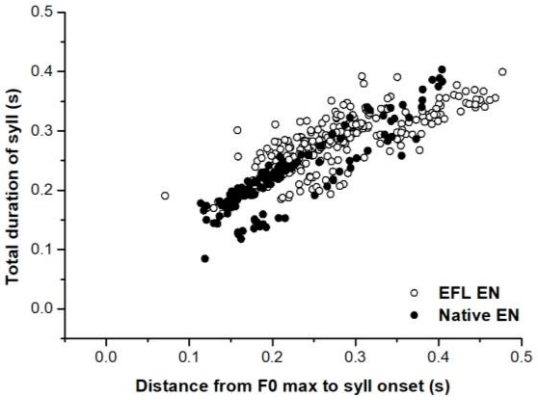


Figure 4: Timing of H* in total syllable duration for prenuclear accent, in cases of H* (left) and L+H* (right).

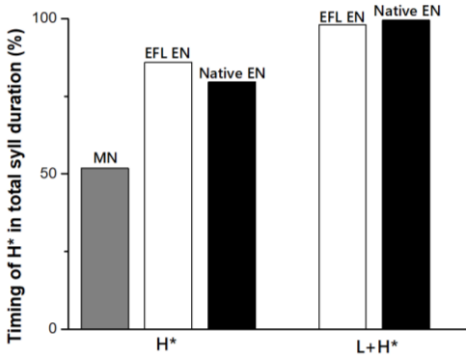
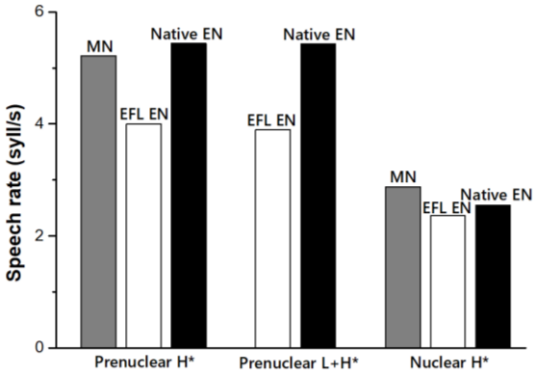


Figure 5: Speech rate of prenuclear (H* and L+H*) and nuclear accents among EFL learners and native speakers.



4.2. Nuclear accents

Figure 6 shows the alignment of nuclear accents realized as H* among EFL learners and native English speakers. The timing of H* occurrence in total syllable duration is also shown in the bar chart

of Figure 7. Contrary to the case of prenuclear accent, EFL learners aligned their pitch peaks earlier in their English than in their Mandarin for nuclear accents (the H* occurred at about 30.20% and 22.54% of the total syllable duration in Mandarin and English, respectively [$t(21) = 5.19$, $p < .0001$]). The two figures suggest that, for nuclear accents, EFL learners showed a similar alignment pattern with that of native English speakers (both had H* occurring at about 23% of the total syllable duration [$t(42) = -0.28$, ns]). Their speaking rates were also comparable, as shown in Figure 5 (2.36 syll/s and 2.55 syll/s for EFL and native speakers, respectively [$t(42) = -1.67$, ns]).

Figure 6: Alignment of nuclear accent realized as H* among EFL learners and native English speakers

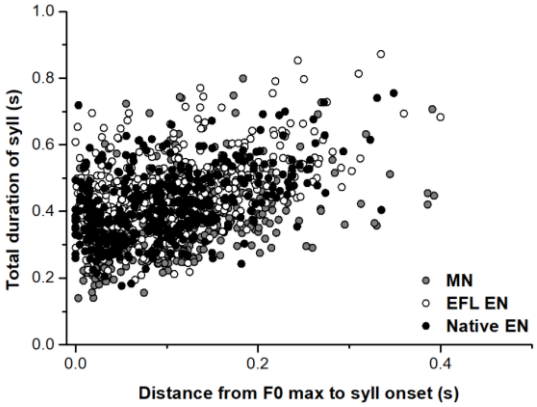
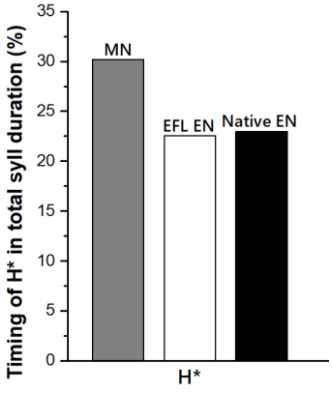


Figure 7: Timing of H* in total syllable duration for nuclear accent.



5. DISCUSSION AND CONCLUSION

The results in this study were somewhat surprising, and did not exactly correspond to previous studies. First of all, compared to native English speakers, EFL learners indeed showed some differences between nuclear and prenuclear accents, but the difference mainly lay in speaking rate, not relative peak alignment. This is in contrast with what was found in the previous studies on ESL learners [3]. One suspects that the incongruence might have something to do with the fact that previous studies had not taken into consideration the effect of speech rate (cf. [3]). Once speech rate is taken into consideration, L2 speakers actually align their pitch peaks at fairly comparable positions.

Secondly, we found that Mandarin peak alignments of falling tones were unanimously earlier than English in the prenuclear position. This implies that any late peak alignment found for EFL learners could not have been due to L1 transfer from Mandarin, but should have been due to some other factors, such as difficulties in articulatory coordination between glottal and supraglottal gestures. The fact that there were disproportionately more L+H* in the prenuclear but not the nuclear position implies that such a possibility exists.

Finally, the results suggest that in an L2 setting, choice of pitch accent tone types may be somewhat challenging for EFL learners. Similar results have also been found in Graham & Post [5], where both Spanish and Japanese learners of English showed pitch accent choices different from those of native English speakers. In this study, although EFL learners had different choices of tone types from native speakers in prenuclear accents, their realization of pitch alignment was shown to be very similar to that of native speakers once they master the particular tone type in question (i.e. H*). Therefore, the challenge that EFL learners face in the acquisition of L2 prosody may lie in choosing a target-like pitch accent contour shape rather than in the actual realization of tonal alignment itself. This may also imply that H* in English is a relatively challenging tone type for L2 learners to acquire, so instead of trying to realize H*, they chose to use another tone type, L+H*, which they find easier to master. Whether such strategic difference can also be found in other aspects of prosody awaits further investigation.

In summary, there was indeed some differences between prenuclear and nuclear accents in L2 prosody. Advanced EFL learners showed peak alignment and syllable duration comparable to their native counterparts in the nuclear position, but only showed comparable alignment but not syllable duration in the prenuclear position. Late peak alignment for the prenuclear position found in previous studies was probably not a genuine alignment difference, but a natural consequence of a slower speech rate. Nonnative-like prosody in advanced EFL learners could possibly be due to difficulty in articulatory coordination, rather than simple L1 transfer.

6. REFERENCES

[1] Beckman, M. E., & Ayers, G. 1997. *Guidelines for ToBI labelling*. The OSU Research Foundation, 3.

[2] Busà, M. G. & Urbani, M. 2011. A cross linguistic analysis of pitch range in English L1 and L2. Proc. 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong, 380-383.

[3] Chen, S. & Fon, J. 2008. *The peak alignment of prenuclear and nuclear accents among advanced L2 English learners*. Oral presentation at the meeting of Speech Prosody 2008, Fourth International Conference, Campinas, Brazil.

[4] Cho, T and P. Ladefoged. 1999. Variation and Universals in VOT: Evidence from 18 Languages, *Journal of Phonetics*, 27, pp. 207-229.

[5] Graham, Calbert & Brechtje Post. 2018. Second language acquisition of intonation: Peak alignment in American English. *Journal of Phonetics* 66. 1-14.

[6] Jun, Sun-Ah & Oh, Mira 1996. A prosodic analysis of three types of Wh-phrases in Korean, *Language and Speech* 39(1):37-61.

[7] Ladd, D. R., Mennen, I., & Schepman, A. 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America*, *107*, 2685–2696.

[8] Ladd, D., Schepman, A., White, L., Quarmby, L., & Stackhouse, R. 2009. Structural and dialectal effects on pitch peak alignment in two varieties of British English. *Journal of Phonetics*, 37, 145–161.

[9] Lu, Yu-an & Kim, Miran. 2016. Prosody Transfer in Second Language Acquisition: Tonal Alignment in the Production of English Pitch Accent by Mandarin Native Speakers. *Tsing Hua Journal of Chinese Studies*, *46*(4), 785–816. https://doi.org/10.6503/THJCS.2016.46(4).04

[10] Möbius, B.; Jilka, M., 2007. Effects of syllable structure and nuclear pitch accent on peak alignment: A corpus-based analysis. *Proceedings of the XVI ICPhS*, Germany: Saarbrücken, 1173-1176.

[11] Mücke, D.; Hermes, A., 2007. Phrase boundaries and peak alignment: An acoustic and articulatory study. *Proceedings of the XVI ICPhS*, Germany: Saarbrücken, 997-1000.

[12] Paul Boersma & David Weenink. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/

[13] Peng, S., Chan, M. K. M., Tseng, C., Huang, T., Lee, O. J., & Beckman, M. E. 2007. Toward a Pan-Mandarin system for prosodic transcription. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

[14] Pierrehumbert, Janet B. 1980. *The Phonology and Phonetics of English Intonation.* Cambridge: Massachusetts Institute of Technology, 1980.

[15] Pierrehumbert, J. B., & Steele, S. A. 1989. Categories of Tonal Alignment in English. *Phonetica*, 46(4), 181–196. doi:10.1159/000261842

[16] Silverman, K., & Pierrehumbert, J. 1990. The timing of prenuclear high accents in English. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 72–106). Cambridge: Cambridge University Press.

[17] Ullakonoja, R. 2007. Comparison of pitch range in Finnish (L1) and Russian (L2). Proc. 16th International Congress of Phonetic Sciences (ICPhS XVI), Saarbrücken, 1701-1704

[18] Urbani, Martina. 2003. The pitch range of Italians and Americans. a comparative study. Ph.D. dissertation, University of Padova, 2003.

[19] Xu Yi. 1998. Consistency of Tone-syllable Alignment Across Different Syllable Structures and Speaking Rates, *Phonetica*, 55, pp. 179-203. doi: 10.1159/000028432

[20] Zimmerer, F., Jügler, J., Andreeva, B., Mobius, B., & Trouvain, J. 2014. Too cautious to vary more? A comparison of pitch variation in native and non-native productions of French and German speakers. In Proc. *7th international conference of speech prosody, Dublin, Ireland* (pp. 1037e1041).