



On Rhythm, Prosodic Grouping, and Declination Pattern of Taiwan Mandarin Learners of English

Sally Chen, Janice Fon

National Taiwan University

sallychentw@gmail.com, jfon@ntu.edu.tw

Abstract

In this study, rhythmic measures, prosodic grouping, and extent of declination were investigated for native and nonnative English. L2 data were obtained from a standardized English oral test in Taiwan. Those who obtained the highest score were categorized as high proficiency (HL2), and those who barely passed low proficiency (LL2). Native speakers of English and Mandarin were also recruited to record the same/translated materials to serve as references. Results showed that L2 speakers assigned more accents and more intonational phrases (IPs) as compared to their native counterparts. Placement of IP boundaries was both type- and group-dependent. Speakers were in general unanimous in placing prosodic boundaries when they encounter disjunctures explicitly marked by punctuation marks. However, with disjunctures not explicitly marked, only the HL2 speakers consistently placed IP boundaries as the natives, and unlike native preference for falling contours, both HL2 and LL2 speakers preferred level tones instead. Finally, the extent of declination tended to be negatively correlated with English proficiency, and this did not come from negative L1 transfer. Considering all the data at hand, we proposed that prosodic grouping was easier to acquire, while native-like rhythmic and declination patterns were more difficult to learn.

Index Terms: L2 prosody, rhythm, prosodic grouping, declination patterns

1. Introduction

Intonation of a language plays a crucial role in regulating discourse. Thus, inappropriate intonation patterns bring up misunderstandings, which is commonly observed in second-language (L2) production [1, 2]. However, most L2 studies have been focused on the investigation of segments, as research has been scant in suprasegments, except for intonational errors [3] or rhythmic patterns [4].

To investigate suprasegmental features, the first step would be to observe how speakers segment speech stream into meaningful units, and how they address prominence. In this study, data description was based on the labeling of Tone and Break Indices (ToBI). Originally established on the phonology and phonetics of English intonation in the 1980's, ToBI has been considered one of the most representative systems with regard to prosodic transcription, specifically for prosodic grouping and tonal assignment [5, 6, 7]. Currently, the system has been applied to the investigation of prosody in various languages, such as German [8, 9], Greek [10], Japanese [11], and Korean [12], among others. In this study, we adopted this system to mark the phonological prominence and phonetic details of the recordings, based on which the comparison between L1 and L2 intonation became possible.

As Mennen proposes, the learning of prosody should be very similar to the process of segmental acquisition [2]. If that is true, the phonological learning of prosodic categories should be easier for L2 learners, as compared to phonetic fine-tuning details. In this study, phonological categorization patterns such as the placement and choice of boundary tones were included, and so were phonetic realization like accent density and declination.

It is assumed that prosodic features like rhythm are likely to be transferred from a source or native language into a target language (L2) [13]. If negative transfer does exist, more pitch accents should be observed in L2 production; furthermore, if L2 proficiency plays a role, then those with lower English levels would assign even more accents to their production than their higher-proficiency counterparts.

The other phonetic realization pattern investigated is declination. When perceiving speech, listeners tend to automatically compensate for such a downtrend [14]. Empirical studies further supported that declination is not just a by-product of physiology [15], but is linguistically controlled [16]. However, the issue of declination is more complicated than merely calculating the lowering of pitch accents [14]. The gradual decrease of F0 not only signals utterance ends, but decodes discourse structures [17]. Declining F0 trends are likely to occur on units of different sizes, varying from lower-level phrases to discourse structures [17, 18]. For example, the initial height was found positively correlated with utterance length, and this holds true cross-linguistically, as both Mandarin and English data showed a similar pattern [16]. Nativeness was also an influential factor for declination. Given that L2 learners speak target languages more slowly, the slopes of the best-fit trendlines of their intonational phrases are usually flatter than those of native speakers. For instance, the slope of Mandarin learners' English utterances was consistently flatter than that of native American English speakers', which had been speculated as the result of L1 transfer [19].

In this study, nativeness and proficiency effects were examined over the phonological categorization and phonetic realization of read speech of Mandarin learners of English and native English speakers. We hoped to validate whether the acquisition path of intonation is similar to that of segments, i.e., the phonological categorization acquired earlier than phonetic details.

2. Method

Nonnative data was selected from a test-based learner corpus, whereas native data of both English and Mandarin was recorded in a phonetic laboratory under a similar test scenario. Details were described as follows.

2.1. Subjects

2.1.1. Non-native Speakers

A set of 16 recordings was randomly selected from a learner corpus established by the Language Training and Testing Center. The corpus is test-based, with the data collected from the General English Proficiency Test (GEPT), a national English proficiency test of various levels regularly held in Taiwan. The data used in this study were drawn from the oral test of the intermediate level. Half of the recordings receiving a score of 3 out of a five-point scale (the passing score) was categorized as the Low-L2 group (LL2), and the other half receiving a 5 (the highest score) the High-L2 (HL2) group.

2.1.2. Native Speakers

As American accent has been the mainstream for learners in Taiwan, a group of 8 native speakers of American English was recruited for recording as the English control. In addition, another group of 8 native speakers of Taiwan Mandarin was also recorded, serving as the native Mandarin control. They were given exactly the same test materials (Mandarin translation for the Taiwan Mandarin group) and recorded under a similar test scenario in the sound-treated phonetic laboratory in National Taiwan University.

2.2. Materials

2.2.1. English Read-aloud

The read-aloud part of the proficiency test has two short passages, consisting of 165 English words in total. All the sentences in the two passages were statements, with punctuations clearly marked on the test sheet. The test sheets were given to the test takers/native speakers one minute prior to the recording for preparation. They were instructed to read at a comfortable speed within a limit of two minutes.

2.2.2. Mandarin Translation

The two English passages were translated into Mandarin following a comparable scheme: All the sentences were statements with punctuation marks, and the alternation between short and long phrases of the original text were also maintained.

Instead of using transliteration, the English personal or place names in the original text were replaced by local Mandarin personal names and place names commonly seen in Taiwan. This was to ensure that participants stay in the same language mode while being recorded. Moreover, in order to accommodate the fact that Mandarin is a tone language, and Tone 2 and Tone 3 in Mandarin do not contain a high tone [20], all the initial and final characters of each potential chunking site used words of either a Tone 1 or Tone 4, both of which contain a high tone. This makes cross-language top-line comparison possible. The same materials and instructions were given in Mandarin during the recordings.

2.3. Data Collection and Labeling

2.3.1. Equipment for Lab Recording

A SHURE SM10A head-mounted microphone connected to a KORG MR-1000 mobile recorder was used for the collection of native English and Mandarin data. The sampling

rate was set at 48kHz and later downsampled to 22.5 kHz for analyses.

2.3.2. Data Labeling

Two trained phoneticians independently labelled the data, following the English ToBI convention [6]. Whenever different opinions were held for a certain label, final agreement would be reached via discussion.

In addition to words, core tiers of a standard ToBI transcription include the Tones and the Break Indices. In the tones tier, pitch accents, or syllables receiving prosodic salience, phrase accents, and boundary tones (at utterance ends) were denoted. Single high tones or bi-tonal combinations with an H element were marked for pitch accents, and L-L% for the phrase accent and boundary tones of the final fall in a declarative utterance. In the breaks tier, levels of break indices (BIs) were marked according to the corresponding size of disjuncture between two adjacent words. BI4 represents the maximal level of disjuncture, and always appears at the boundary of intonational phrases (IPs) in the tones tier. An example of the ToBI labelling is shown in Figure 1.

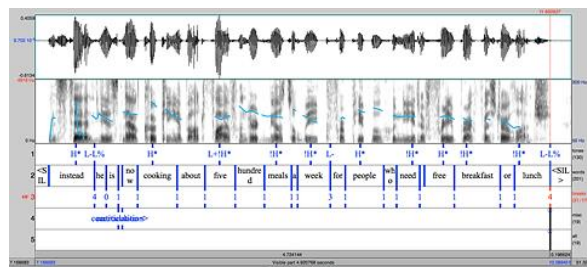


Figure 1: *Demonstration of English ToBI Labelling.*

3. Results

Results of phonological categorization and phonetic realization are reported as follows. For the former, position and tone type of the assignment of BI4 labels were analyzed for examining prosodic grouping. For the latter, BI4 ratio and pitch accent density were included for observing rhythmic features, and overall and sectional pattern of the topline of pitch contours were analyzed for declination.

3.1. Phonological Categorization

3.1.1. Boundary Tone Position

Two types of boundary positions were defined: Those with clear punctuation marks were always assigned a prosodic boundary, and were thus labeled “Syn+Pro”. In addition, words receiving seven out of eight BI4 assignments from the native speakers were termed “prosodic boundaries (Pro-only).” As shown in Figure 2, both HL2 and LL2 groups showed a very high level of consistency in assigning BI4s at Syn+Pro positions as native speakers. At Pro-only positions, however, only HL2 was found to act more like the native English controls. Effect of L2 proficiency played a role here.

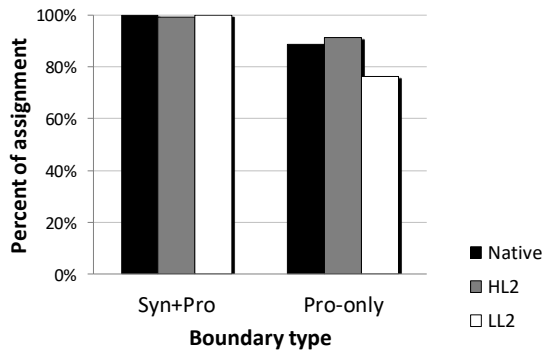


Figure 2: Assignment of boundary tone positions.

3.1.2. Boundary Tone Type

We then analyzed the tone types used at both kinds of boundary. As shown in Figure 3(a), for Syn+Pro boundaries, all groups preferred falling tones (L-L%), which is also the default type for statements. As shown in Figure 3(b), for Pro-only boundaries, the choice of falling tones or a continuation rises (L-H%) consisted of the majority of the data contributed by the native and HL2 groups. Interestingly, both nonnative groups showed a tendency for a higher percentage of H-L% and H-H% tones, and the preference of these seemingly “illegal” tone types was negatively correlated with their English proficiency—LL2 speakers even assigned more than half of their BI4s with H-L% tones. It was speculated that Mandarin learners of English might concentrated to the “non-low” feature of English boundary tones, and thus the confusion.

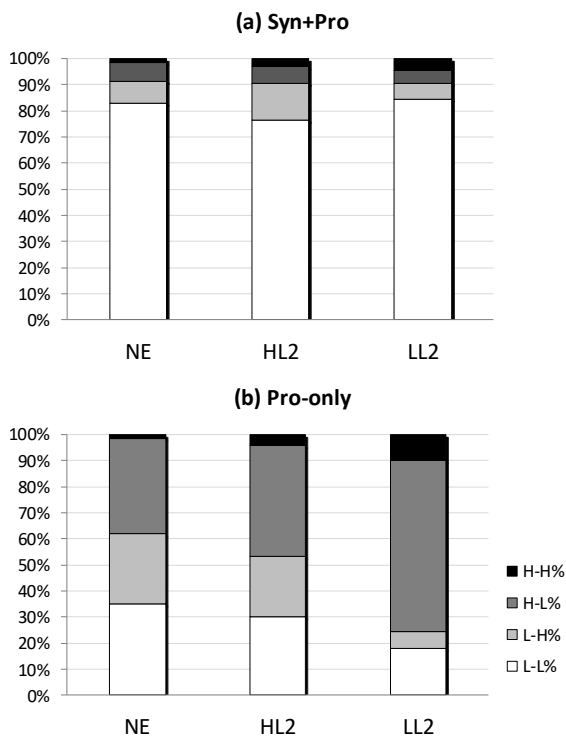


Figure 3: Boundary tones assigned at (a) syntactic-and-prosodic and (b) prosodic-only boundaries.

3.2. Phonetic Realization

3.2.1. The BI4 Ratio

As aforementioned, BI4s mark the boundaries of intonational phrases, the highest level of break indices in the ToBI system. As shown in Figure 4, the two non-native groups had a comparable percentage of BI4 assignment, which was higher than that of the native control [$F(2, 21) = 4.12, p < .05$]. No significant difference was observed between HL2 and LL2 groups, so the effect of nativeness, rather than that of L2 proficiency, served as the driving force.

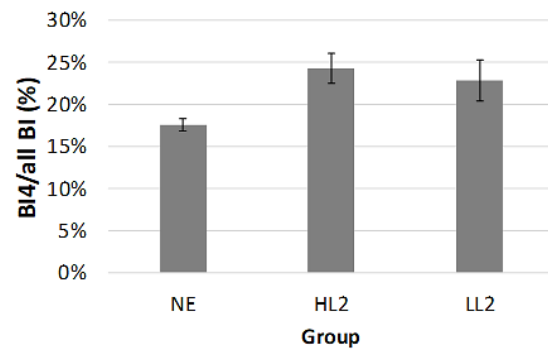


Figure 4: BI4 ratio (error bars stand for standard errors).

3.2.2. Accent Density

Accent density was computed with the total number of pitch accents labeled divided by the total number of words uttered in each recording. As can be seen in Figure 5, the results successfully reflected the impression that nonnative English “sounds heavier.” The higher the density level, the more pitch accents were assigned in speech [$F(2, 24) = 52.54, p < .001$]. Moreover, HL2 learners behaved more like native speakers. The two groups assigned fewer accents in English read speech, as compared to the LL2 group. Negative transfer of L2 might be the underlying driving force, as it is believed that as a syllable-timed language, Mandarin does not have stress distinction [21, 22], and thus each syllable should be pronounced with comparable prominence.

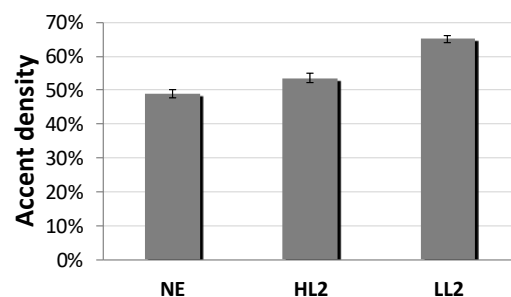


Figure 5: Accent density values of individual groups (error bars stand for standard errors).

3.2.3. Overall Declination

Extent of declination was defined as the pitch difference between the first and the last peaks of each IP. For this measure, the effect of proficiency was observed. As shown in

Figure 6, a likely developmental path was observed, with native English speakers employing the largest difference between their first and last peaks of IPs, followed by HL2 and LL2 groups, respectively. However, a one-way ANOVA test comparing English proficiency levels only revealed marginal significance. Rather than L2 proficiency, nativeness seemed to play a crucial role in this measure.

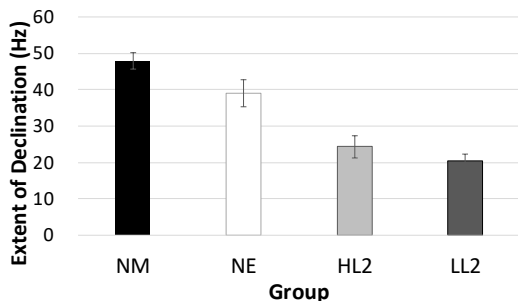


Figure 6: Extent of declination of different groups (error bars stand for standard errors).

3.2.4. Language-intrinsic Difference

Another statistic analysis was run to test for the language-intrinsic difference, and significant difference was found [$t(317) = 2.04, p < .05$]. As shown in Figure 6, compared with native English (NE), a larger declination level was observed for native Mandarin (NM). This supports a previous finding [16], and ruled out the speculation of another study [19] that smaller extent of declination in the production of Mandarin learners of English is likely to be driven by L1 influence. Mandarin in fact showed an even larger, rather than smaller declination extent than English.

3.2.5. Declination at Different Section of the Contours

In order to have a clearer picture about declination, we further picked out the IPs with three pitch accents and above and divided the pitch contours into two parts—the first half between the first two pitch accents, and the second half between the second pitch accent and the last one. Through examining the difference between the two section, we may observe more details in this phonetic measure.

As shown in Figure 7, native English showed a higher value of the first high target of the pitch accents than native Mandarin, and this difference in the first accent seemed to be grasped by the LL2 learners, as they raised their pitch in their English production to a comparable height as the native English speakers. On the other hand, the HL2 learners seemed to notice that the end of the contours should be lowered, as the two native groups in fact shared a comparable final low. With a successful manipulation of the end, however, the start of their first accent seemed to be much lower than the English norm. Our speculation was that attention of the L2 learners may be attracted to the extreme values of a contour first; when they become even more proficient, in addition to the beginning or the end, the learners should be able to manage the entire “process” in between and reached for an appropriate “range” for declination. Preliminary results of our follow-up work showed that indeed, learners with an even higher proficiency level than HL2 showed a larger overall pitch range than HL2 and LL2 groups here.

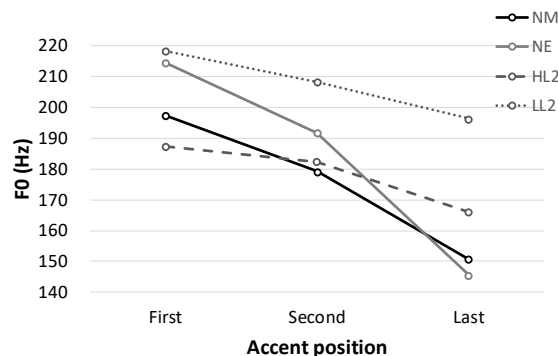


Figure 7: Different sections of the declination.

4. Discussion

Results of the above analyses showed that phonological measures seemed to be easier, and are thus acquired earlier by Mandarin learners of English. Either no difference was found among the three groups, or the HL2 acted more like the NE in the position or type of boundary tones. Effect of L2 proficiency was only observed in the choice of tone at boundaries that did not coincide with punctuation marks (Pro-only boundaries).

As for phonetic measures, the HL2 group was more like the NE only in accent density. On the other hand, the HL2 group showed a comparable performance like LL2 in B14 ratios, and only marginal difference was found between the HL2 and LL2 groups in overall declination. These results thus indicated that phonetic realization, which usually involves more fine-tuning details of a language, are harder to acquire.

Lastly, L2 learners showed a smaller extent of overall declination in their English read speech, but the general declination level of Mandarin, their native language, is in fact larger than that of English, the target language. This ruled out previous speculation that the difference of declination level between native and Mandarin learners’ English might result from negative L1 transfer [19].

5. Conclusion

In this study, the phonological categorization and phonetic realization of the prosody of L1 and L2 English read speech were examined. Based on the results obtained, acquisition of the phonological categories seemed to be easier, and is achieved earlier than that of phonetic details. More data from more L2 proficient levels are to be expected in the future to complete a more thorough story of L2 prosody learning in Taiwan EFL learners.

6. References

- [1] I. Mennen, “Phonetic and phonological influences in non-native intonation: an overview for language teachers.” *QMUC Speech Science Research Centre Working Paper*, vol. 9, pp. 1-18. 2006.
- [2] I. Mennen, I. 2007. “Phonological and phonetic influences in non-native intonation,” In *Nonnative Prosody: Phonetic Description and Teaching Practice*, J. Trouvain, and U. Gut (Eds). Mouton de Gruyter, pp. 53-76. 2007
- [3] I. Mennen, “Second language acquisition of intonation: the case of peak alignment,” In *Chicago Linguistic Society: The Panels*, vol. 34, no. 2, M. C. Gruber, D. Higgins, K. Olson, and T. Wysocki (Eds). 1998.

- [4] V. Dellwo, and P. Wagner, "Relations between language rhythm and speech rate," In *Proceedings of the 15th ICPHS*, D. Recasens, M. J. Sole, and J. Romero (Eds). Barcelona, Spain. 2003.
- [5] M. E. Beckman, and G. A. Elam, *Guidelines for ToBI Labelling*. The Ohio State University Research Foundation. 1993.
- [6] M. E. Beckman, J. Hirschberg, and S. Shattuck-Hufnagel, "The original ToBI system and the evolution of the ToBI framework." In *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.-A. Jun (Ed). Oxford University Press, pp. 9-54. 2005.
- [7] J. Pierrehumbert, *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation, MIT. 1980.
- [8] M. Grice, S. Baumann, and R. Benzmueller, "German intonation in autosegmental-metrical phonology," In *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.-A. Jun (Ed). Oxford University Press, pp. 55-83. 2005.
- [9] M. Grice, M. Reyelt, R. Benzmueller, J. Mayer, and A. Batliner, "Consistency in transcription and labeling of German intonation with GToBI," In *Proceedings of the 1996 International Conference on Spoken Language Processing*, Philadelphia, U.S.A., pp. 1716-1719. 1996.
- [10] A. Arvaniti, and M. Baltazani, "Intonational analysis and prosodic annotation of Greek," In *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.-A. Jun (Ed). Oxford University Press: pp. 84-117. 2005.
- [11] J. J. Venditti, "The J_ToBI model of Japanese intonation," In *Prosodic Typology: The Phonology of Intonation and Phrasing*. S.-A. Jun (Ed). Oxford University Press, pp. 172-200. 2005.
- [12] S.-A. Jun, "Korean intonational phonology and prosodic transcription." In *Prosodic Typology: The Phonology of Intonation and Phrasing*, S.-A. Jun (Ed). Oxford University Press, pp. 201-229. 2005.
- [13] I. Mennen, and E. de Leeuw, "Beyond Segments: Prosody in SLA," *Stud. Second Lang. Acquis.*, vol. 36, pp. 183-194. 2014.
- [14] C. Shih. *A declination model of Mandarin Chinese*. 2000.
- [15] A. Cohen, and J. t'Hart, "On the anatomy of intonation," *Lingua*, vol. 19, pp. 177-192. 1967.
- [16] J. Yuan, and M. Liberman, "F0 declination in English and Mandarin broadcast news speech," *Speech Communication*, vol. 65, pp. 67-74. 2014.
- [17] M. Swerts, E. Strangert and M. Heldner, "F0 declination in spontaneous and read-aloud speech," In *Proceedings of the 15th International Conference on Spoken Language Processing*, Philadelphia. 1996.
- [18] M. Swerts, and R. Geluykens, "Prosody as a marker of information flow in spoken discourse," *Language and Speech*, vol. 37, no. 1, pp. 21-43. 1994.
- [19] C-K. Liu, "MAE_ToBI reflects where EFL students fail in acquiring the intonation of English." *Languages, Literary Studies, and International Studies: An International Journal*, vol. 3, pp. 69-84. 2006.
- [20] J. Fon, W.-Y. Chiang, and H. Cheung, "Production and perception of two dipping tones (T2 and T3) in Taiwan Mandarin," *Journal of Chinese Linguistics*, vol. 32, no. 2, pp. 249-280. 2004.
- [21] L. Hyman, "On the nature of linguistic stress," *Studies in stress and accent*, vol. 4, pp. 37-82. 1977.
- [22] E. Selkirk, and T. Shen, "Prosodic domains in Shanghai Chinese," In *The Phonology-Syntax Connection*, S. Inkelas and D. Zec (Eds), pp. 313-337. 1990.