PROSODIC FEATURES OF NON-NATIVE ENGLISH PRODUCTION

Sally Chen & Janice Fon

Graduate Institute of Linguistics, National Taiwan University, Taiwan
d93142002@ntu.edu.tw; jfon@ntu.edu.tw

ABSTRACT

This study investigated the prosodic features of non-native spoken English. Sixteen recordings were extracted from an in-progress test-based learner corpus, each consisted of read-aloud and question-and-answer parts by a learner who had received a grade of 3 (passing) or 5 (highest) on a five-point scale. A group of 8 native speakers was recruited and recorded under the same test scenario as a control. Labeling followed the English ToBI convention. Results showed that liaisons only occurred among native speakers; non-native speakers were in general less fluent and assigned more tones in their production. In read speech, both non-native groups showed a high level of consistency in assigning BI4s at appropriate places; however, nativeness made a difference in the preference for tone types.

Keywords: L2 English, ToBI, prosodic grouping

1. INTRODUCTION

In addition to conveying linguistic information, intonation (or prosody in a broader sense) of a language plays an essential role in regulating discourse. As a result, inappropriate intonation patterns may bring up misunderstandings, and we also derive our impression about others from how they intonate [8, 9].

Inappropriate intonation is especially common among the population of which a target language, *e.g.* English, is acquired as a second or foreign language. Most second-language (L2) studies, however, focus on the investigation of segments, and research has been scant regarding suprasegments. Most related ones concentrated on intonational errors [7] or rhythmic patterns [4].

To investigate suprasegmental features of speech, how speakers segment the speech stream or address prominence along the way they speak is fundamental. Originally established on the phonology and phonetics of English intonation in the 1980's, Tone and Break Indices (ToBI) has been considered one of the most representative systems with regards to prosodic transcriptions, specifically for prosodic grouping and tonal assignment [2, 3, 10]. Currently, ToBI has been applied to the labeling of prosodic features of various languages, such as German, Greek, Japanese, and Korean [3].

However, the above applications have mainly focused on languages among native speakers. As Gut, et al. [6] put it, "crosslinguistic descriptions of the intonational system of languages are still few and far between". Recently, Bradlow and her colleagues [1, 5] have started on this issue by probing into the relationship between basic acoustic measures of segments and the resultant perceptual judgments among native speakers on the basis of a non-native English corpus with varying L1 backgrounds.

In the present study, we took the initiative to investigate the prosodic features of Mandarin speakers' English production in both read and spontaneous speech. Specifically, we examined whether L2 learners share the same prosodic grouping patterns and assign boundary tones as native English speakers do, as well as how they assign pitch accents.

2. METHOD

2.1. Subjects

Subjects include two groups of nonnative speakers with Taiwan Mandarin as their daily language, and a group of native English speakers as a control.

2.1.1. Non-native speakers

A set of 16 recordings was extracted from an in-progress learner corpus established by Language Training and Testing Center in Taiwan. The corpus is test-based, with all data collected from the General English Proficiency Test (GEPT), a national English proficiency test of various levels held regularly. Data drawn for this study were all from the intermediate level. Half of the recordings received a grade of 3 (the passing score), and the other half a 5 (the highest score) out of a five-point scale.

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 a control. They were given exactly the same materials and recorded under the same test scenario in a sound-treated booth in National Taiwan University.

2.2. Materials

Two different sets of materials were included:

2.2.1. Reading data

Materials for a read-aloud task were two texts consisting of 165 words in total. Sentences in these texts were all statements with punctuations clearly marked in the test sheet. The test sheets were given to the test takers/native speakers one minute prior to the recording for preparation. They were then instructed to read at a comfortable speed within a limit of two minutes.

2.2.2. Spontaneous data

The spontaneous data were the question-and-answer part of the same test. After listening to the recording of each question (10 in total, each repeated once), the test takers/native speakers had to provide an impromptu answer. Time limit for response was 15 seconds for Questions 1 to 5, and 30 seconds for Questions 6 to 10. Subjects heard a bell to signal the end of the response time.

2.3. Labeling

Following the English ToBI convention [2], the two authors independently labeled the data, and agreed on the final labels given in the core tiers: the "tones", the "words" tier, and the "breaks" tiers.

2.4. Analyses

Till this moment, we have finished the labeling of all the read speech data and one fourth of the spontaneous data. Based these data, we conducted three analyses: Given read and spontaneous speech were not equal in length, we calculated distribution of the percentages of BI0, BI4 and BI-p labels. Second, "accent density" was computed with the total number of pitch accents labeled divided by the total number of words uttered in the recording for both read and spontaneous data of each subject group. We also investigated the assignment of BI4

labels at appropriate prosodic boundaries, as well as their corresponding tone types.

## 3. RESULTS

3.1. Break indices

3.1.1. BI0

Native speakers merge the boundary between two adjacent words sharing the same phone at the boundary (e.g., *He is so….*), and they usually do not enunciate function words (*e.g.*, *and* is frequently pronounced as syllabic *n*) [11]. In contrast, L2 learners tend to enunciate each word clearly. This contrast was revealed via BI0 labels. As shown in Figure 1, the natives showed an average of 2% of liaisons in read-aloud and 6% in spontaneous speech, whereas the two non-native groups seldom applied this rule in their production, even in spontaneous speech [F (2, 24) = 26.7, p < .001].

Figure 1: Percentages of BI0 in read and spontaneous speech (the *y*-axis stands for percentages with a possible maximal value for all BI labels adding up to 100%).
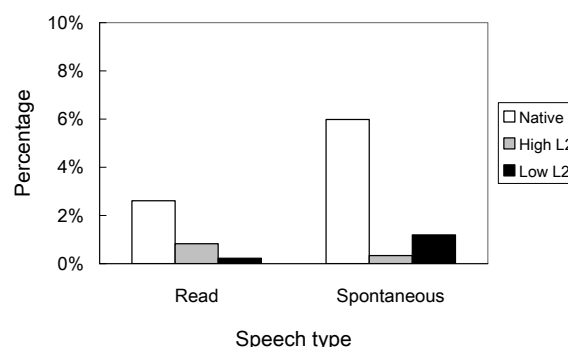
Figure 2: Percentages of BIp in read and spontaneous speech (the *y*-axis stands for percentages with a possible maximal value for all BI labels adding up to 100%).
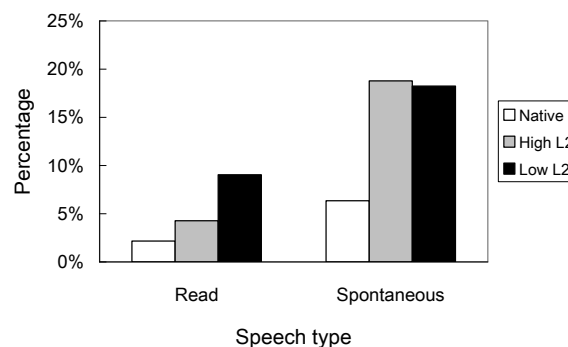
3.1.2. BIp

Since the distributional pattern of the three disfluency levels were alike, we collapsed the data together and used BIp to yield a more general picture. As expected, non-native speech is not as fluent as native's. Figure 2 shows a greater disfluency ratio for spontaneous data [F (1, 24) = 11.70, $p < .01$]. In addition, for both speech types, the two non-native groups showed significantly higher percentages for disfluency [F (2, 24) = 8.06, $p < .01$].

3.1.3. BI4

BI4s mark the boundary of intonational phrases, the highest level of break indices in ToBI. Along the process of labeling, the two labelers impressionistically heard more chunks in the non-native production, as compared to native speech. This was supported in the data. As shown in Figure 3, in read speech, the two non-native groups had a comparable percentage of BI4 assignment, which was higher than that of the native group [F (2, 21) = 4.12, $p < .05$]. In a more impromptu scenario like question-and-answer, the nonnative assignment of BI4s was negatively correlated with their English proficiency.

Figure 3: Percentages of BI4 in read and spontaneous speech (the *y*-axis stands for percentages with a possible maximal value for all BI labels adding up to 100%).
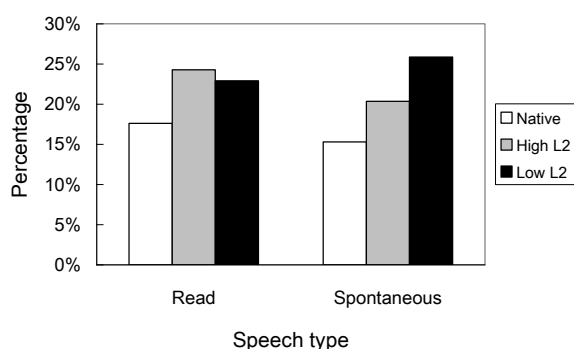

3.2.　Accent density

Accent density serves as a measure successfully reflecting the impression that nonnative English "sounds heavier": the higher the density level, the more pitch accents were assigned in the speech [F (2, 24) = 52.54, $p < .001$]. In addition, the more natural the speech type is, the lower the density level [F (1, 24) = 9.48, $p < .01$]. High-L2 learners behaved more like native speakers in read speech; however, they did not reduce their density level for spontaneous speech, indicating that they tended to hyper-articulate for most of the time.

Figure 4: Accent density in read and spontaneous speech (the *y*-axis stands for percentages of the words receiving pitch accents; the possible maximal value is 100%).
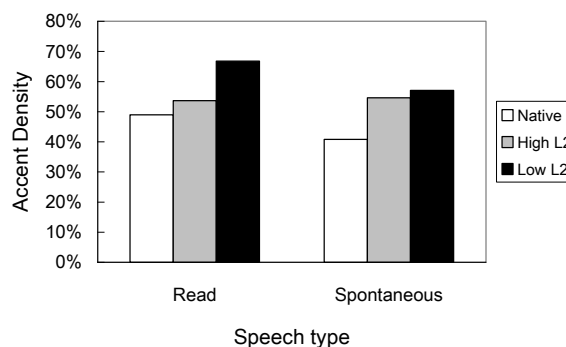

3.3.　Tone position and tone types at BI4 in read-aloud data

3.3.1. Positions

Any word in the texts receiving seven out of eight BI4 assignments from native speakers was defined as an appropriate place for intonational phrases (IPs). This always co-occurred with periods (*e.g., "… in New York City."*). As shown in Figure 5, under this definition, both non-native groups showed a very high level of consistency in assigning BI4s at these appropriate places.

Figure 5: The ratios of BI4 coinciding with appropriate boundaries for the two non-native groups.
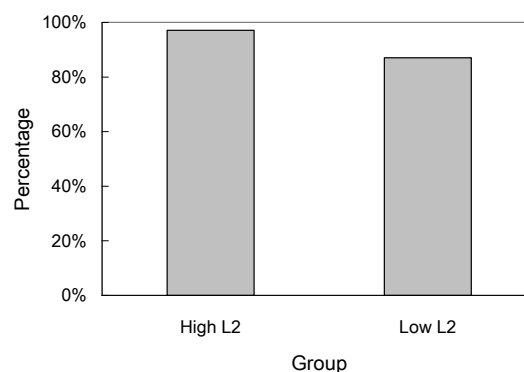

3.3.2. Tone types

We then further analyzed the tone types of these boundaries and compared the results with BI4s being randomly assigned to other inappropriate places in the texts (*e.g., "… in New York City."*).

As shown in the upper panel of Figure 6, all groups had the default tone type L-L% as the majority for IP boundaries. High-L2 learners also

showed a comparable percentage of L-H%, the default tone type for continuation, as native speakers did. Interestingly, both nonnative groups showed a tendency of higher percentages for H-L% and H-H% tones. The preference for these "illegal" tone types for statements was even more clearly observed from the BI4s assigned randomly to other inappropriate boundaries. As shown in the lower panel of Figure 6, the percentages of using H-L% were negatively correlated with English proficiency. Low-L2 speakers even assigned more than half of their BI4s with H-L%. Future comparisons with spontaneous data should provide a more complete picture of this potential preferential difference.

Figure 6: Tone types for BI4s at appropriate (upper) and inappropriate (lower) boundaries.
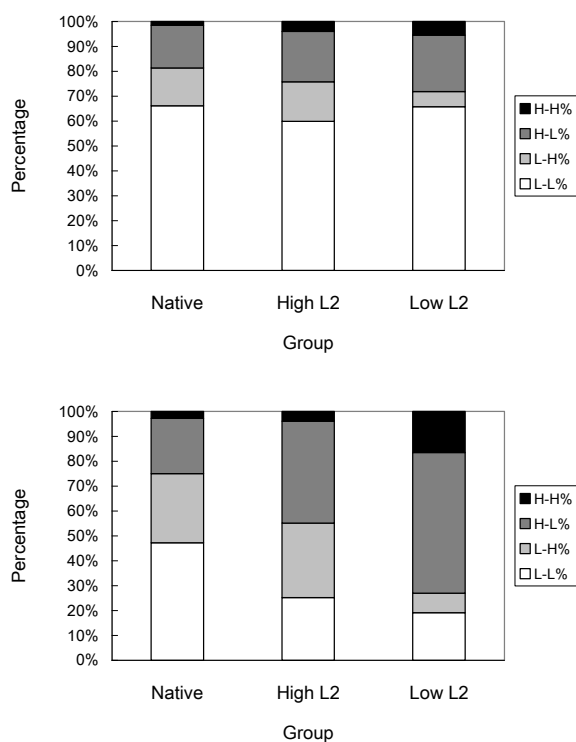
4. CONCLUSION

In this study, we found break indices faithfully confirmed our naive impression that only native speakers merge word boundaries; non-native speakers were less fluent, and chunked their speech into more IPs. In addition, accent density was found to be negatively correlated with English proficiency and speech type. Finally, in read speech, non-native speakers showed a high level of consistency in assigning BI4s at appropriate places; however, compared to native speakers, they had a different preference for tone types.

5. REFERENCES

[1] Baker, R.E., Baese-Berk, M., Bonnasse-Gahot, L., Kim, M., Van Engen, K., Bradlow, A.R. 2011. Word durations in non-native English. *Journal of Phonetics* 39, 1-17.

[2] Beckman, M.E., Elam, G.A. 1993. *Guidelines for ToBI Labelling*. The Ohio State University Research Foundation.

[3] Beckman, M.E., Hirschberg, J., Shattuck-Hufnagel, S. 2005. The original ToBI system and the evolution of the ToBI framework. In Jun, S.-A. (ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press, 9-54.

[4] Dellwo, V., Wagner, P. 2003. Relations between language rhythm and speech rate. In Recasens, D., Sole, M.J., Romero, J. (eds.), *Proceedings of the 15th ICPhS* Barcelona, Spain.

[5] van Engen, K.J., Baese-Berk, M., Baker, R.E., Choi, A., Kim, M., Bradlow, A.R. 2010. The Wildcat Corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech* 53(4), 510-540.

[6] Gut, U., Trouvain, J., Barry, W. 2007. Bridging research on phonetic descriptions with knowledge from teaching practice - The case of prosody in non-native speech. In Trouvain, J., Gut, U. (eds.), *Nonnative Prosody: Phonetic Description and Teaching Practice*. Berlin: Mouton de Gruyter, 3-21.

[7] Mennen, I. 1998. Second language acquisition of intonation: the case of peak alignment. In Gruber, M.C., Higgins, D., Olson, K., Wysocki, T. (eds.), *Chicago Linguistic Society* 34(2), The Panels.

[8] Mennen, I. 2006. Phonetic and phonological influences in non-native intonation an overview for language teachers. *QMUC Speech Science Research Centre Working Paper* 9, 1-18.

[9] Mennen, I. 2007. Phonological and phonetic influences in non-native intonation. In Trouvain, J., Gut, U. (eds.), *Nonnative Prosody: Phonetic Description and Teaching Practice*. Berlin: Mouton de Gruyter, 53-76.

[10] Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation, MIT.

[11] Raymond, W.D., Makashay, M.J., Dautricourt, R., Johnson, K., Hume, E., Pitt, M. 2001. *Variation in Conversation: An Introduction to the Buckeye Speech Corpus*. Presented at the *Annual ASA Meeting* Ft. Lauderdale, FL, U.S.A.