PROSODIC CUES AND DEGREE OF PERCEIVED FOREIGN ACCENTS IN LEARNER ENGLISH

[a]Bin Li, [b]Wenling Cao, [a]Xiu Yan

binli2@cityu.edu.hk;gracecao325@gmail.com;ctjackie@cityu.edu.hk

[a] City University of Hong Kong, Hong Kong S.A.R.
[b]University of York, U.K.

ABSTRACT

Foreign accent is judged by the degree to which the non-native (L2) production is perceived differently from production patterns by native speakers (NL) [1]. Both segmental and suprasegmental features are shown influential in accent perception by NL speakers or L2 learners [2]-[4]. However, roles of prosodic cues in the process are not well known yet. Our study aims to investigate effects of intonation and speaking rate in foreign accent perception. We modified properties of the two cues in English utterances produced by NL and L2 speakers. Mandarin and Cantonese learners of English judged on the degree of foreign accents. Results show that intonation had a stronger influence on accent rating in L2 than in native speech, while speaking rates affected judgement of L2 accents more. No differences were found between Cantonese and Mandarin L2 English samples, but Cantonese listeners seemed more sensitive to native English samples than Mandarin listeners did.

Keywords: foreign accents, prosodic cues, speech perception, Chinese ESL/EFL learners

1. INTRODUCTION

In perceiving accents in L2 English, listeners are influenced by factors such as familiarity of particular accents [5], L1s [6], age and length of stimuli [7]. Moreover, L2 accents are correlated with segmental and prosodic variations [2]-[4]. But prosody seems to have stronger influence in accent perception by L2 listeners [8], while NL listeners focus more on segmentals [9]. Much research is still needed to examine the contributing aspects in this superiority of prosody in accent perception.

1.1 The role of prosodic cues in accent Perception

Prosodic cues such as accent location and speaking rates have been shown to lead to perceptual differences [10]-[14]. Among them, intonation and speaking rate are of interests to the current study, as they are reported crucial in perceiving foreign accents in English: in distinguishing L2 from NL speech [15] [16], in identifying NL speech [6], and in improving native-likeness in L2 speech [17] [18].

1.2 The role of listeners' L1 in accent perception

L2 speakers of a given L1 identify their own L2 accents with equal or greater accuracy [19], but this role of familiarity may vary depending on tasks and types of phonetic cues contained in the L2 speech [6] [20]. On the other hand, there is also research suggesting that listeners L1 does not affect their perception of foreign speech, claiming a common perception pattern across language backgrounds as listeners would rely on acoustic cues than linguistic knowledge in accent perception [13] [18].

2. STUDY DESIGN

The current study aims to examine effects of phonetic cues in L2 accent perception by L2 learners. It also sets out to assess if listeners' L1s play a role in such L2 perception of accents.

2.1. Speech stimuli

Speech stimuli were spontaneous utterances from TV interviews involving native and bilingual speakers, and from consecutive interpretation by advanced L2 learners of English whose L1s were Cantonese or Mandarin. L2 speech in English was first evaluated by certified English teachers of interpretation. In total, forty utterances were selected.

To examine effects of prosodic cues on accent perception, we then modified the speech frequency and the speaking rate of the utterances in the following ways. First, a band-pass filter was designed based on individual speaker's pitch ranges extracted using Praat [21]. In general, the baseline of the filter was 50 Hz higher than the lowest point of a certain pitch range, and the top was 350 Hz higher than its peak. For example, the modification outcome yielded band-pass at 150-500 Hz with an 80 Hz width for a Mandarin female speaker with an original pitch range at about 100-285 Hz. For optimal perceptual effect, naturalness of all modified samples was assessed by the authors and 5 listeners other than the participants. Manual fine adjustment was used to ensure minimal lexical effect in the samples. As a result, all filtered samples were mumbling-like speech and unintelligible in terms of lexical content, but essential prosodic cues such as intonation, prominence, and stress remained.

Secondly, each talker's speech rate was calculated and normalized with a reference to the mean rate of the native and bilingual speakers of English. Last, in order to examine individual and combined effects of F0 and the speaking rate, three conditions were generated. Condition 1 (F0): only band-pass filters were applied. Condition 2 (SR): only speaking rates were modified. Condition 3 (F0+SR): both band-pass filters and modification to speaking rates were applied.

2.2. Procedure

Stimuli were randomized using E-Prime 1.1., and presented twice through earphones with computers. Participants were told that they need assess the degree of "native-likeness" of speech produced by people who were native or non-native speakers of English. They rated the degree on a 5-point scale, with 1 meaning "very non-native like" and 5 "very native-like". They were also told not to struggle with intelligibility of these utterances as some were changed on purpose. To avoid learning effect, listeners were randomly assigned to work with one manipulation condition only.

2.3. Listeners

Fifteen native speakers of Cantonese and Fifteen native speakers of Mandarin were recruited as listeners. They are all undergraduate or postgraduate students at the City University of Hong Kong, and have little knowledge about languages other than their mother tongue and English.

3. RESULTS

We tested 40 utterances by talkers from 4 backgrounds: English-Cantonese (EC) bilingual, Cantonese L2 learners of English (CE), Mandarin L2 learners of English (ME), native speakers of English (NSE). The prosodic cues modified were F0 and the speaking rate. Utterances were put in 3 conditions for perceptual judgment: F0 filtered, speaking rate modified, F0 filtered with speaking rate modified. Original utterances without any manipulation were also used as baseline condition. Our listeners were from different L1s: Cantonese and Mandarin.

3.1. General results

In general, stimulus from different talker groups received distinct ratings: those by NSE ($M= 3.86$, $SE= .023$) were rated higher than speech produced by EC bilinguals ($M= 3.42$, $SE = .046$, $p < .001$). Speech by these two groups were both judged more native-like than that by two learner groups (Stimulus from CE: $M= 2.58$, $SE= .044$; ME: $M= 2.46$, $SE=.037$). However, ratings for CE samples were not different from those for ME samples ($p= .229$).

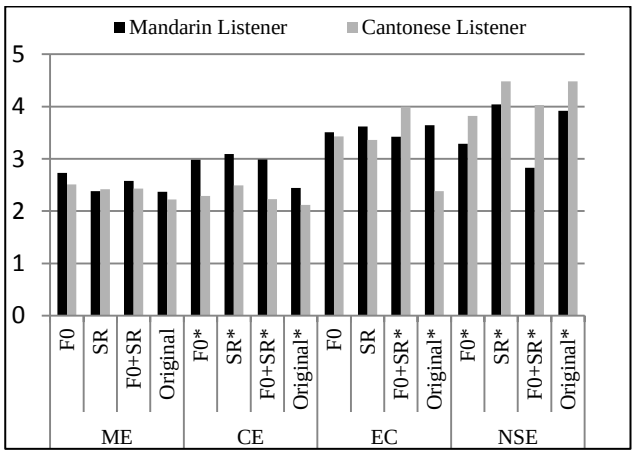
In terms of prosodic cues, types of modification affected the rating differently. In general, original samples ($M= 2.95$, $SE= .048$) were rated more accented than those with speaking rate modified (SR) ($M= 3.07$, $SE= .044$, $p < .001$), those from the F0 condition ($M= 3.07$, $SE= .044$), and also those from the F0+SR condition ($M= 3.06$, $SE= .048$). But the latter two conditions did not yield different ratings.

A three-way ANOVA was conducted on the accent rating data with the rating as dependent variable and modification types (F0, SR, F0+SR, Original) and talkers' L1 (ME, CE, CE bilingual, NSE) as within-subject factors, and listeners' L1 (Mandarin, Cantonese) as between-subjects factors. Statistics revealed main effects of talkers' L1 ($F (3, 3197) = 285.026$, $p < .001$), and of modification types ($F (3, 3197) = 8.809$, $p < .001$), but not of listeners' L1 ($F (1, 3199) = 3.174$, $p = .075$). Significant two-way interactions were observed between listeners' L1 and talkers' L1 ($F (3, 3197) = 45.634$, $p < .001$), between listeners' L1 and

modification type (F (3, 3197) = 7.023, p < .001), as well as between talkers' L1 and modification type (F (9, 3191) = 14.594, p < .001). There was also significant three-way interaction between Listener L1, talkers' L1, and modification types (F (9, 3191) = 7.436, p < .001).

Details of mean value across variables are shown in Figure 1. To further analyze the effects of between- and within-subject factors, a series of one-way ANOVA tests were implemented separately and the results were described by variables as follows.

Figure 4: Comparison of average rating scores between Mandarin and Cantonese groups for all talker groups and across modification types.



3.2. Ratings by talkers' L1s

Two One-way ANOVA tests with talkers' L1 as within-group factor were conducted in Mandarin listener and Cantonese listener groups respectively. For both listener groups, significant main effects were all observed across four conditions (Mandarin Listeners: F0 condition, F(3, 396) = 8.173; SR condition, F(3, 396) = 32.295; F0 + SR condition, F(3, 396) = 9.037; Original condition, F(3, 396) = 56.636. Cantonese listener: F0 condition, F(3, 396) = 42.531; SR condition, F(3, 396) = 100.844; F0 + SR condition, F(3, 396) = 66.093; Original condition, F(3, 396) = 117.057; all p < .001).

Post hoc tests revealed significant interaction between the variables. Mandarin listeners could not distinguish NSE from CE bilingual except in the F0+SR condition (p < .05). They performed well in identifying native and bilingual English from learner English, but failed to detect the difference between CE bilinguals and Cantonese in F0+SR condition (p= .06). Samples from NSE group were also rated more native-like than those from two learner groups in original and SR conditions (all p< .001). However, once the segmental information of learner samples were filtered out, Mandarin listeners could only distinguish native English from Mandarin accented English in F0 condition. Comparison between Cantonese samples and Mandarin samples suggested that Mandarin listeners could only separate these two in the SR condition (p < .001), as shown in Figure 1.

On the other hand, post hoc tests of Cantonese listener group revealed a different pattern as shown in Figure 2. Cantonese listeners failed to distinguish CE from ME in all conditions (P>.1). NSE samples were rated higher than both groups of learner samples in all conditions (p <.001). CE bilingual samples were rated in a similar way, except in the original condition where they were indistinguishable from either group of learner English. Unlike Mandarin listeners, Cantonese listeners were able to separate CE bilinguals from NSEs in original and SR conditions (p< .001), however, they could not do so once the segmental information was filtered out.

3.3. Ratings by modification types

One-way ANOVA tests with modification types as within-subject factor and listeners' L1 as between-subject factor were conducted for individual group of samples. As shown in Figure 3, main effect were found in Mandarin listeners' ratings on ME samples (F (3, 396) = 2.867, p < .05) and CE samples (F (3, 396) = 5.64, p= .001), and also in Cantonese listeners' ratings on CE bilinguals (F (3, 396) = 32.26, p < .001). Differences were also found between listener groups' ratings on NSE samples (Mandarin listener, F (3, 396) = 23.456, p < .001; Cantonese listener, F (3, 396) = 13.3, p < .001).

Post hoc tests revealed different degrees of influence by modification types. For ME samples, only marginal effect was found between F0 and original conditions (p= .078), and the F0+SR condition (p= .095) by Mandarin listeners. No significant difference was found in the ratings by Cantonese listeners. For CE samples, Mandarin listeners considered original samples more accented than all modified samples (p< .05). But Cantonese listeners rated all samples the same regardless the modification types. For CE bilingual samples, those in the original condition were rated the lowest (most accented, p< .05) while F0+SR

samples were rated the highest (most native-like, $p < .001$) by Cantonese listeners. Mandarin listeners, on the other hand, did not demonstrate difference in rating across the four conditions.

For NSE samples, both Cantonese and Mandarin listeners rated those in the original condition and SR condition more native-like than those in F0 condition and F0+SR condition ($p < .05$). Samples in F0 condition were rated more native like than F0+SR samples by Mandarin listeners; but sounded the same to Cantonese listeners.

3.4. Ratings by Listeners' L1s

ANOVA tests with listeners' L1 as between-subject factor revealed significant main effects on all conditions of CE samples (F0, $F(1, 199) = 15.683$, $p < .001$; SR, $F(1, 199) = 10.446$, $p = .001$; F0+SR, $F(1, 199) = 20.371$, $p < .001$; original, $F(1, 199) = 4.542$, $p < .05$), as well as on NSE samples (F0, $F(1, 199) = 9.678$, $p = .002$; SR, $F(1, 199) = 13.613$, $p < .001$; F0+SR, $F(1, 199) = 49.325$, $p < .001$; original, $F(1, 199) = 20.654$, $p < .001$).

4. DISCUSSION

Our study examined effects of prosodic cues and L1 on perceiving foreign accents in English. Our main finding shows that listeners' L1 has little effect on accent perception, as both Mandarin and Cantonese listeners could not distinguish between their peers' L2 production. Though Mandarin listener differed from Cantonese listener in identifying Cantonese-accented English, it remains unknown if the difference was caused by the samples as we did not include listeners who are native speakers of English. Also, the benefit of shared language background proposed by Bent and Bradlow [19] is not found in accentedness rating, as our Cantonese listeners did not show preference towards Cantonese-English bilinguals. This suggests that accentedness and intelligibility may involve different processing, as Munro and Derwing [16] reported that intelligibility, comprehensibility and accentedness are partially-independent dimensions. So, an utterance rated as highly intelligible and comprehensible can be perceived as moderately or heavily accented.

Though effects of listeners' L1s were not significant, further analysis across sample types suggests differences between Cantonese and Mandarin listener. For instance, Cantonese listener gave higher scores than Mandarin listener did when rating stimulus produced by native speakers of English. This may be attributed to Cantonese listeners' richer experience with English in Hong Kong compared with Mandarin listeners who only arrived in Hong Kong from the Mainland shortly before the experiment.

Another finding is that enhancing or softening of prosodic cues influences the perceived accentedness of nonnative speech. All modified Cantonese samples were rated more native like than original ones by Mandarin listeners, but not by Cantonese listeners. It is not clear why Mandarin listeners were more sensitive to such modifications. If the differences were caused by listeners' L1s, then Mandarin samples should show similar patterns. However, only marginal differences between F0 and original conditions were found for Mandarin samples.

Thirdly, an examination of the modification types shows clearly that intonation alone (with segmental cues filtered out) exerted a strong influence on successful detection of nonnative than native speech. The rating scores of NSE samples were lower (more accented) once the segmental information was filtered out, whereas the rating of non-native samples increased (more native-like) if only F0 was available. In other words, accents of non-native speech may be mainly caused by pronunciation of segments. Our results are in line with Winter and O'Brien's study [22] suggesting segments contribute more to perceived accentedness than intonations in foreign English speech. Speaking rates, on the other hand, do not seem important in perceiving native English samples, but they could reduce accents in non-native speech. This is consistent with findings from the study by Munro and Derwing (2001). Moreover, L2 listeners seem to benefit from speaking rates in distinguishing accented learner English from native speech.

5. CONCLUSIONS

This study examined roles of two prosodic cues in perceiving accented learner English by L2 learners themselves. We found that eliminating segmental information in nonnative speech could help reduce the perceived accentedness. Speaking rates also influence accent perception of non-native speech. Moreover, listeners' L1s do not contribute to accent perception of non-native speech. Further

explanation of findings in the study will require more listener groups, as well as speech produced by learners at various proficiency levels.

6. ACKNOWLEDGEMENT

Xxxx.

7. REFERENCES

[1] Fox, A. 2000. *Prosodic Features and Prosodic Structure*. New York: OUP.

[2] Almberg, J. and Husby, O. 2000. The Relevance of Some Acoustic Parameters for the Perception of a Foreign Accent", in James, A. and Leather, J (eds), *New Sounds*, 1-10, University of Klagenfurt.

[3] Mareüil, P. and Brahimi, B. and Gendrot, C. 2007. Role of Segmental and Suprasegmental Cues in the Perception of Maghrebian-accented French. *International Conference on Speech and Language Processing*, Korea.

[4] Trofimovich, P. and Baker, W. 2006. Learning Second Language Suprasegmentals: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech, *Studies in Second Language Acquisition*, 28: 1-30.

[5] Gass, S. and Varouis, E. 1984. The Effect of Familiarity on the Comprehensibility of Nonnative Speech. *Language Learning*, 34(1): 65-89.

[6] Grabe, E., Rosner, B., Garcia-Albea, J. and Zhou, XL. 2003. Perception of English intonation by English, Spanish, and Chinese listeners. *Language and Speech*, 46(4): 375-401.

[7] Jilka, M. 2000. *The Contribution of Intonation to the Perception of Foreign Accent*. University of Stuttgart. Ph.D. dissertation.

[8] Riney, T. and Inutsuka, K. 2005. Phonetic Parameters and Perceptual Judgments of Accent in English by American and Japanese Listeners. *TESLO Quarter*, 39:441-466.

[9] Sato, T. 1995. A Comparison of Phonemes and Prosody in the Evaluation of Spoken Japanese. *Japanese Language Education around the Globe*, 5:139-154.

[10] Lehiste, I. and Fox, R. A. 1992. Perception of Prominence by Estonian and English Listeners. *Language and Speech*, 35(4): 419 – 434.

[11] Beaugendre, F. and Hermes, D. 2001. Accentuation Boundaries in Dutch, French and Swedish. *Speech Communication*, 33: 305 – 318.

[12] Dupoux, E., Pallierm C., Sebastian, N., Mehler, J. 1997. A Distressing Deafness in French? *J. of Memory and Language*, 36: 406 – 421.

[13] Xu, B. and Mok, P. 2012. Intonation Perception of Low-Pass Filtered Speech in Mandarin and Cantonese, *Proc. Tonal Aspects of Languages*, Nanjing.

[14] Sebastian-Galles, N., Dupoux, E., Costa, A. and Mehler, J. 2000. Adaptation to Time-compressed Speech: Phonological Determinants. *Perception & Psychophysics*, 62(4): 834-842.

[15] Munro, M. 1995. Nonsegmental Factors in Foreign Accent: Rating of Filtered Speech, *Studies in Second Language Acquisition*, 17: 17-34.

[16] Munro, M. and Derwing, T. M. 1998. The Effects of Speaking Rate on Listener Evaluations of Native and Foreign-accented Speech. *Language Learning*, 48: 159–182.

[17] Munro, M. and Derwing, T. 2001. Modeling Perceptions of the Accentedness and Comprehensibility of L2 Speech. *Studies in Second Language Acquisition*, 23:451-468.

[18] Bent, T. and Bradlow, A. 2003. The Interlanguage Speech Intelligibility Benefit. *J. Acoust. Soc. Am*., 114: 1600-1610.

[19] So, K. 2005. The Influence of L1 Prosodic Background on the Learning of Mandarin Tones: Patterns of Tonal Confusion by Cantonese and Japanese Native Listeners. *Proc. of the Annual Conference of the Canadian Linguistic Association*.

[20] Wu, X., Tu, J. and Wang, Y. 2011. Native and Nonnative Processing of Japanese Pitch Accent. *Applied Psycholinguistics*: 1-19.

[21] Boersma, P. and Weenink, D. *Praat: doing phonetics by computer [Computer program] V 5.3.68*. http://www.praat.org/ accessed on 20 March 2014.

[22] Winters, S. and O'Brien, M. 2013. Perceived Accentedness and Intelligibility: The Relative Contributions of F0 and Duration. *Speech Communication*, 55:486-507.