

The perception of Mandarin focus intonation by native English speakers

Yi Liu, Jinghong Ning

The Hong Kong Polytechnic University, Hong Kong

yi.liu@polyu.edu.hk, choin.ning@polyu.edu.hk

Abstract

Based on cue theory and PAM-S, this paper examines how acoustic cues in Mandarin focus sentences are perceived by native English speakers. Acoustic analysis of the stimuli spoken by a native speaker reveals that in perceptual tests, temporal and intensity cues had an effect in determining focus prominence produced by non-native learners, whereas the influence of these phonetic cues was not detected in the perception of native speakers. In contrast, pitch variation is the cue that facilitates native listeners' perception of focus prosody in Mandarin. Moreover, pitch direction such as the interplay of tone and intonation, and emphasis locations in Mandarin are also factors affecting the focus perception of native English speakers.

Index Terms: perception, focus intonation, Mandarin, native American English speakers, acoustic cues, tones

1. Introduction

Focus is a commonly used oral intonation associated with the emphasis of information structure and prosodic realization. A large body of research has explored the properties of focus intonation. [1] proposed that, in contrast with unfocused syllables (or neutral statements), focused syllables have the following common prosodic features across languages: expanded fundamental frequency (F0) range, longer duration, and greater intensity. Post-focus compression (PFC) effects with prominent pitch lowering in the words following focus have been detected in English focus sentences ([2], [3], [4]).

[5] and [6] explored the properties of intonation in tonal languages. Mandarin has four lexical tones: Tone 1 (T1, 55), Tone 2 (T2, 35), Tone 3 (T3, 214), and Tone 4 (T4, 51). A Mandarin syllable consists of an optional initial, a final, and a lexical tone. Due to the tone sandhi, the first T3 syllables in T3 disyllabic words or combinations become T2 in continuous speech in Mandarin [7]. A substantial number of studies have been conducted to examine the characteristics of focus sentences in Mandarin as prosodic features of intonation ([8], [9], [4], [10], [11]). It is well documented that focused words present an enlargement of pitch range and an increase of duration and intensity, and that there is compression of PF words (in contrast to neutral-focus sentences) in Mandarin focus sentences ([8], [9], [4]). As a tonal language, Mandarin presents an interplay of intonation and tone ([12], [13]). Tone and focus intonation are simultaneously realized in the F0 pattern in Mandarin focus sentences [14]. The pitch patterns of neutral (without focus) and focus intonation produced by native Mandarin speakers are illustrated in Figure 1.

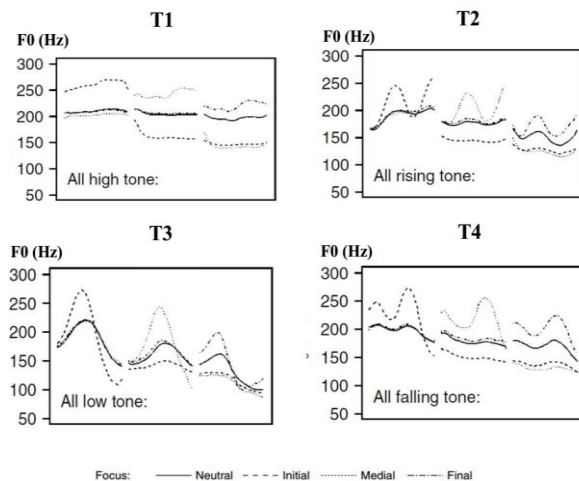


Figure 1. Pitch patterns for statements and focus in Mandarin (source: [8]).

The graphs in Figure 1 demonstrate the acoustic features of focus in Mandarin: regardless of tone type, the pitch rising and lowering of focus syllables, and PFC, relative to their counterparts in neutral sentences or statements. Figure 1 also shows that focused syllables maintain similar contours as those in neutral statements with larger pitch range. As well as the pitch adjustment shown in Figure 1, [8] also reported the phonetic features of duration and intensity in Mandarin focus intonation, the longer duration and stronger intensity of focus syllables and the shorter duration and weaker intensity of post-focus syllables than in neutral sentences.

Phonetic cues play an important role in the production and perception of native articulation [15]. Previous studies have proposed that acoustic cues affect the realization and perception of suprasegmental features (i.e., intonation) by native speakers [16] and non-native speakers [17]. Thus, the analysis of various cues is necessary in order to understand which phonetic features speakers/listeners are likely to realize/perceive at the suprasegmental level. Apart from pitch variation in English and Mandarin focus ([3], [8], [4]), duration and intensity cues have been identified with acoustic evidence. For instance, [18] found that in American English, duration and intensity are the acoustic cues in the production of stress, and duration ratio is a more effective cue than intensity ratio. [19] and [20] demonstrated that pitch prominence is the primary cue for English stress in natural and synthesized speech. In their study, [21] reported that F0 is the least reliable cue, compared with intensity, vowel quality and duration on perception of stress words in English. [22] reported that duration and intensity correlate with stress, and both serve as perceptual cues to facilitate stress perception

in Mandarin. Usually listeners attend to more than one cue in the perceptual processing of phonetic features, and are likely to make auditory judgments based on the interaction between phonetic cues [9]. The influence of phonetic cues can also be detected in the perception of segmental and suprasegmental features by non-native speakers during acquisition.

Differences in intonation between different languages, such as tonal and non-tonal languages, may pose difficulties for non-native acquisition [23]. Perceptual assimilation model (PAM; [24]) holds that the perception of second language (L2) segments depends on the similarity of the phonetic characteristics of the native language (L1) and L2 phonetic categories. Listeners tend to assimilate similar L2 segments to their own L1 categories, and develop a new category where they perceive totally new L2 segments. Recently, more and more attempts have been made to extend PAM to intonation prosody. According to the perceptual assimilation model for suprasegmentals (PAM-S, [25], [26], [27]), native speakers of both tonal and non-tonal languages assimilate L2 prosodic categories into their native prosodic categories [28]. [29] proposed that tonal listeners showed much greater categorization in the perception of Mandarin tones than non-tonal and accented listeners. [30], [27] found that Mandarin T1 and T3 were usually assimilated as statements, T2 as questions, and T4 as statements by English speakers and as exclamations by French speakers. Such assimilation was affected by both the overall descending pitch tendency in Mandarin sentences and the varying coarticulation of tones. Investigating native Australian English speakers' and French speakers' perception of Mandarin tones in continuous speech, [28] found that both groups categorized Mandarin tones by the phonetic similarities of the pitch contours between the Mandarin tones and their native intonation categories. They also reported that stress features of the two languages might influence the perception of non-native tones as English uses lexical stress, but French does not. Thus far, few studies have looked at how the interplay of tone and focus prosody in Mandarin speech affects the perception of non-native speakers, or investigated the acoustic cues which influence non-native learners' perception of focus sentences in Mandarin at suprasegmental level.

It has been documented that phonetic cues such as pitch, duration, and intensity may have influence on listeners' (including native and non-native speakers) perception of prominent focus prosody [9]. The aim of current study is to examine the effect of acoustic cues, tone, and intonation on non-native speakers' perception of Mandarin focus intonation. Following the observations of [18] and [17], the current study predicts that acoustic cues may affect native English speakers' perception of Mandarin focus locations, which is an understudied area. Analysis of perceptual accuracy and the acoustic features of stimulus could reveal how acoustic cues are perceived in Mandarin focus prosody by native English speakers and native Mandarin speakers.

As tone and intonation are suprasegmental features, PAM-S will be adopted to account for the auditory performance of native English speakers in perceiving Mandarin focus prosody. According to the conclusions of [14], [30] and [27], T2 can be assimilated as a question and T1, T3, and T4 can be perceived as statements by non-native speakers, and so we predict that

native American English speakers are less able to perceive focus sentences with T2 than those with T1, T3, and T4.

2. Experimental design

2.1. Subjects

Eighteen American students (9 female, 9 male; age $M=21.6$ years, $SD=1.5$) with an intermediate-advanced level of Mandarin (according to the proficiency guidelines developed by the American Council on the Teaching of Foreign Languages, ACTFL, 2012) participated in the perception task. They were all native speakers of American English. At the time of the test, they were studying at Peking University or The Hong Kong Polytechnic University. They had been studying Mandarin for four years on average ($SD=2.1$). The control group consisted of 18 (9 female, 9 male; age: $M=28$ years, $SD=7.4$) native speakers of Mandarin. None of the participants reported any hearing disorder.

2.2. Materials

The present study used the same set of materials as [8], including four kinds of ten-syllable tone sentences, high (T1), rising (T2), low (T3), and falling (T4). The sentences were designed with initial, medial, or final focus (the disyllabic combination) alternately, which responded on questions with a pronoun (*shuí* [who or whom]) or verb phrase (*gànmá* [do what?]) in Mandarin. A female native speaker of Mandarin with Grade 1 (the highest grade on the scale) in the Putonghua Proficiency Test recorded the target sentences as stimuli at a sound-treated laboratory of The Hong Kong Polytechnic University. There are total of 12 Mandarin focus sentences used as material in the perceptual task. As with [8], and [11], the current study used sentences with identical tones as material, which may neutralize the pattern of pitch movement.

2.3. Procedure

The subjects sat in front of a computer to complete the perceptual experiment. They were instructed to pay attention to the emphasis (the disyllabic combination) in the 12 Mandarin sentences they heard and to mark the focus (emphasis) locations in Mandarin sentences shown in Pinyin and tone marks on the answer sheet. Each sentence was played twice. The task was conducted in a quiet room at Peking University and The Hong Kong Polytechnic University. The interval between each item was ten seconds. The answer sheets were collected for further analysis.

3. Results and discussion

The perceptual data below (Table 1) show that American learners' overall identification rate was averagely 15% lower than that of Mandarin native listeners, with much wider standard deviation for subjects of 19.5% than Mandarin natives (5%). Statistically, linear mixed-effects models (LMM) was adopted with L1 difference (native and non-native), focus and tone conditions as fixed variables as well as subject and item difference as random variables, in order to detect the multiple effects on the perception of focus for non-native learners. The

fixed-effect results revealed a significant L1 difference in the perception results across different focus conditions and tone types ($\beta = -76.7$, $SE = 9.3$, $t(407) = -8.3$, $p < .001$, Bonferroni-adjusted), with the perception rate of native listeners was significantly higher than that of non-native learners. In addition, an interactive effect of L1, focus and tone conditions ($\beta = -5.67$, $SE = 1.6$, $t(407) = -3.6$, $p = .0009$, Bonferroni-adjusted) was reported with LMM, indicating that learners would behave differently according to different sentence environments. Furthermore, Post-hoc Tukey tests were conducted within subgroups of data in terms of different focus and tone types to detail the differences between the performances of native and non-native listeners. As shown in table 1, with corrected p-values, it was clear that learners showed comparable perception performance to native Mandarins in all tone types of final focus, and almost every tone type in the medial focus condition, with the exception of T2 tone type ($p < .001$). Also, significant L1 distinctions could be seen in the initial focus across tone types (T1: $p < .001$, T2: $p = .008$, T3: $p = .004$, T4: $p = .008$). These illustrated that American students had no problem identifying the final focus disyllabic combination, achieving a mean identification rate of 93.6% compared to that of natives (98.6%), irrespective of tone type. However, American learners obtained a much lower identification rate for the initial focus (averagely native = 98.9%, non-native = 73.6%) and medial focus with T2 (native = 99%, non-native = 70%) as they mismatched target with other focuses.

Table 1. Identification rate of focus sentences across different tone and focus conditions for non-native (American) and native (Mandarin) listeners (%), standard deviations (in the parenthesis), and post-hoc Tukey results P-values were Bonferroni-adjusted. Significance code: * = .05, ** = .001, *** = .0001.

| | | T1 | T2 | T3 | T4 |
|------------|---------|-----------|-----------|----------|---------|
| Non-native | Initial | 51(30)*** | 73(31)** | 80(30)** | 90(14)* |
| | Medial | 86(21) | 68(24)*** | 90(16) | 89(20) |
| | Final | 96(9) | 94(14) | 93(12) | 92(14) |
| Native | Initial | 99(5) | 99(5) | 100(0) | 98(6) |
| | Medial | 98(6) | 99(5) | 98(9) | 98(6) |
| | Final | 100(0) | 98(6) | 99(5) | 98(6) |

In further, to examine the environmental influence upon the perception of focus by non-native learners, again, LMM (L1 difference was removed) was run within non-native group. It reported that focus locations ($\beta = 25.8$, $SE = 4$, $t(195) = 6.4$, $p < .001$, Bonferroni-adjusted) weighed more over tone types ($\beta = 16.8$, $SE = 3.2$, $t(195) = 5.3$, $p < .001$, Bonferroni-adjusted), with larger variance resulted from the model, and the two factors interacted ($\beta = -6.2$, $SE = 1.47$, $t(195) = -4.2$, $p < .001$, Bonferroni-adjusted) with each other in the influence of learners' perceptual performance as well. The Post-hoc comparisons showed that T1 (T1-T2: $p = .045$, T1-T3: $p = .006$, T1-T4: $p < .006$, Bonferroni-adjusted) and T2 (T2-T1: $p = .024$, T2-T3: $p = .03$, T2-T4: $p = .05$) got the lowest identification rate among all the tone types with respective to the initial and the medial focus respectively (see table 1). That is, American listeners had the most difficulty in perceiving Mandarin focus in T1 initial and T2 medial sentences, which might be attributed to the influence of the acoustic properties in the focus stimuli by the Mandarin speaker.

Regarding the identification rates for the focus sentences, T1 in initial focus presented the lowest of all the perceptual results. In terms of acoustics, besides the range enlargement of focus disyllabic combinations realized in the focus speech of the native speaker, the intensity and duration of the focused position increased concurrently. Therefore, non-native learners as listeners may adopt pitch, intensity, and duration as perceptual cues to locate the target focus in an utterance. Figures 2 and 3 below indicate that T1 in initial focus and T2 in medial focus had shorter duration and weaker intensity than other tones in the same focus locations in terms of the speech of the native speaker. The intensity differential (ID), exhibited in Figure 2, further revealed that the T1 focus presented a much smaller increase in syllable intensity (.2 dB) than the other tone focuses (from .9 to 1.6 dB) in the initial position, as did the medial focus with T2 (.1 dB) compared with that of other tones (from .7dB to 1.5 dB) in the medial position. As for duration differential (DD) displayed in Figure 3, the speaker produced much shorter duration differential in the T1 initial focus (11 ms) than in other tone types (from 39 ms to 45 ms), and the shortest duration adjustment in T2 (7 ms) than other medial focuses (32 to 43 ms). Therefore, the distinction of duration in T1 initial and T2 medial focus results in the low identification rates of the learners. In addition, the mean ID and DD at final position of stimulus are much higher than that in initial and medial position. The perceptual results suggest that the phonetic properties of duration and intensity as perceptual cues, have influence on the differentiation of focused and unfocused disyllabic combinations in utterances by native American English speakers, and the auditory perception of focus prosody by non-native speakers is shaped by the acoustic features of the stimuli produced by a native speaker.

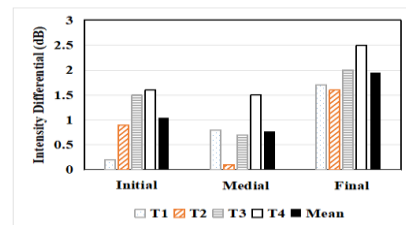


Figure 2. ID in the focus stimuli of a native speaker in the perception test.

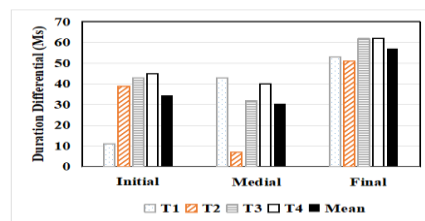


Figure 3. DD in the focus stimuli of a native speaker in the perception test.

For the control group the overall identification rate of focus locations was 98%, regardless of tone type. The data for the ID and DD in T1 initial and T2 medial indicate that they are distinct from other counterparts. However, unlike native American English speakers, such acoustic distinctions did not affect the perception of focus intonation by native listeners.

Based on perceptual results, it is clearly seen that American students were able to identify Mandarin focus in most focus and tone conditions as Mandarin natives did. However, American listeners still had difficulty in perceiving Mandarin focus in T1 initial and T2 medial sentences (refer to Table 1), which might be attributed to the influence of the acoustic properties in the focus stimuli by the Mandarin speaker. In addition, the results also suggest that final focus is more easily perceivable than initial or medial focus as non-native listeners pay more attention to the emphasis at the final location of utterances in the course of speech processing.

Previous studies examining the interaction and realization of intonation and tones in Mandarin have found that non-native speakers have difficulty perceiving Mandarin utterance types such as statements or questions when the pitch movement of the tones and intonation goes in opposite directions ([27], [30], [31]). This finding suggests that Mandarin statements with T4 are easily identified by non-native learners. Concerning the mean pitch contour of focus disyllabic combinations in Mandarin, mean_F0 contour shows that all the focus disyllabic combinations of the four tones maintain a similar pitch curve to neutral-focus sentences, but with range adjustment as shown in Figure 1. With the assumption that statements and focus sentences share the same pitch patterns in intonation, the intonation of both statements and focus utterance is to be treated as distinct from that of interrogatives.

Regarding the perception of emphasis in Mandarin by non-native speakers, the perceptual results show that the overall perceptual rate of focus locations was 84%, and the native American English speakers performed best on the focus of T4, with an identification rate of 90%, followed by T3 and T1 (besides initial focus). The high accuracy reflects the interplay of tone and intonation in Mandarin and the effect of tone on intonation as noted by [27] and [30], who proposed that L2 speakers tend to assimilate T1, T3, and T4 as statements and T2 as a question. Our finding supports PAM-S, the conclusions of [27] and [30], and the prediction of the current study regarding the effect of tone on the perception of intonation by non-native learners. In addition, our perception data may reveal that identification rates for final focus are higher than initial and medial locations, irrespective of tone type, suggesting that the perceptual results are location-specific for L2 learners.

Apart from the effects of tone and intonation, the acoustic results demonstrate that pitch adjustments were fully realized in the stimuli sentences by the native speakers [32]. Previous research has indicated that both temporal and intensity properties affect the auditory behavior of non-native listeners [9]. The ID and DD of T1 in initial focus and T2 in medial focus produced by a native speaker were weaker and shorter than that at the same focus location with other tones. This type of prosodic difference affected the perception of emphasis at sentences level as the non-native learners tended to rely on duration and intensity as perceptual cues to distinguish the focus intonation from neutral statements. The perceptual results show that the native American English speakers failed to identify the initial focus of T1 and medial focus of T2 because the intensity and duration were not salient elements in those focus positions. Thus, the auditory performance of the native American English speakers was shaped and determined by the acoustic features of the stimuli spoken by the Mandarin speaker. This finding suggests that the prosodic features of

duration and intensity serve as perceptual cues, facilitating non-native learners' differentiation of focus from neutral utterances [21]. Native speakers were 100% accurate in the perception of focus locations, although prominent prosody was not detected in the duration and intensity cues of the stimuli of T1 initial and T2 medial speech. These results imply that non-native speakers were affected by duration and intensity cues in Mandarin focus. This interesting finding suggests that duration and intensity are the perceptual cues to influence how native American English speakers perceive focus sentences in Mandarin. And these results need to be examined further with synthesized stimuli in the future to obtain more evidence to verify the findings.

In terms of perceptual performance, the native American English speakers were able to identify the focus locations in emphatic utterances in Mandarin. Final focus is easier to identify than initial and medial focus in natives' utterances, especially the focus disyllabic combination with T4. The identification rates for final-focus sentences ranged from 89% to 92% (regardless of tone type), which are clearly higher than those for initial focus. This perceptual performance of non-native speakers may be influenced by the acoustic features in the final focus prosody in Mandarin stimulus.

The results of the current study support the hypothesis proposed in the perceptual experiment. With their interplay of tone and intonation, Mandarin focus sentences with T2 are the hardest for non-native listeners to perceive. The acoustic data and perceptual results reveal that the effect of intensity and duration was observed in perception of the native American English speakers, but not the native speakers. [22] argued that duration and intensity are perceptual cues which correlate with stress perception in Mandarin for native speakers. However, the results of the current study indicate that the lack of prosodic salience of duration and intensity in the stimuli did not affect native speakers' performance in locating the focus prosody in Mandarin as the acoustic cue of pitch adjustment facilitated their identification of the focus intonation.

4. Conclusions

Based on the results of perceptual task, we conclude that the effects of duration and intensity cues were found in the identification of Mandarin focus by the native English speakers. The perceptual performance of native American English speakers was shaped by the intensity and duration features in the stimuli spoken by a native speaker. However, native speakers of Mandarin were not affected by the acoustic features of focus intonation produced by a native speaker, suggesting pitch was the cue used by native speakers to perceive Mandarin focus prosody. Moreover, the perceptual results were location-specific as the native American English speakers performed best in identifying final focus in Mandarin. In addition, the effects of tone on the perception of Mandarin focus by non-native speakers were observed: the focus sentences in which tone and intonation shared the same pitch movement yielded higher accuracy, and vice versa.

5. Acknowledgements

The authors would like to thank for the funding support of project 4-ZZHP from The Hong Kong Polytechnic University.

6. References

- [1] S. Frota, "The prosody of focus: A case-study with cross-linguistic implications", *Speech Prosody*, pp. 315–318, 2002.
- [2] W. E. Cooper, S. J. Eady and P. R. Mueller, "Acoustical aspects of contrastive stress in question-answer contexts", *Journal of the Acoustical Society of America*, vol. 77, pp. 2142–2156, 1985.
- [3] S.J. Eady and W. E. Cooper, "Speech intonation and focus location in matched statements and questions", *Journal of the Acoustical Society of America*, vol. 80, pp. 402–416, 1986.
- [4] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours", *Journal of Phonetics*, vol. 27, pp. 55–105, 1999.
- [5] M. Lindau, "Testing a model of intonation in a tone language", *Journal of the Acoustical Society of America*, vol. 80, pp.757–764, 1986.
- [6] V. Manfredi, "Spreading and downstep: Prosodic government in tone languages", In: H van der Hulst and K Snider (Eds.), *The phonology of tone*. Berlin: Mouton de Gruyter, pp. 133–184, 1993.
- [7] Y. R. Chao, *Mandarin primer*, Cambridge, MA: Harvard University Press, 1948.
- [8] F. Liu and Y. Xu, "Parallel encoding of focus and interrogative meaning in Mandarin intonation", *Phonetica*, vol. 62, pp. 70–87, 2005.
- [9] B. Wang Y. and Xu, "Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese", *Journal of Phonetics*, vol. 37, pp. 502–520, 2011.
- [10] Y. Xu, C. X. Xu and X. Sun, "On the temporal domain of focus", *Proceedings of International Conference on Speech Prosody*, Nara: Japan, pp. 81–84, 2004.
- [11] F. Shi and P. Wang, "Bianjiediao he jiaodiandiao [Boundary tone and focus tone]", *Journal of Chinese Linguistics*, vol. 42, no.1, pp. 93–108, 2014.
- [12] A. T. Ho, "Intonation variation in a Mandarin sentence for three expressions: Interrogative, exclamatory and declarative", *Phonetica*, vol. 34, pp. 446–457, 1977.
- [13] X. N. Shen, "Interplay of the four citation tones and intonation in Mandarin Chinese", *Journal of Chinese Linguistics*, vol. 17, pp. 61–69, 1991.
- [14] W. T. Gu, "Tone, intonation, and emphatic stress in L2 Mandarin speech by English and Cantonese learners", *Proc. 18th ICPHS, Glasgow: UK, 2015*.
- [15] A. Chrabaszcz, M. Winn, C. Y. Lin, and W. J. Idsardi, "Acoustic cues to perception of word stress by english, mandarin, and russian speakers", *Journal of Speech Language & Hearing Research Jslhr*, vol. 57, no. 4, pp. 1468-79, 2014.
- [16] V. Silber-Varod, H. Sagi and N. Amir, "The acoustic correlates of lexical stress in Israeli Hebrew", *Journal of Phonetics*, vol. 56, pp. 1–14, 2016.
- [17] B. Chandrasekaran, P. D. Sampath and P. C. Wong, "Individual variability in cue-weighting and lexical tone learning", *The Journal of the Acoustical Society of America*, vol.128, no.1, pp. 456–465, 2010.
- [18] D. B. Fry, "Duration and intensity as physical correlates of linguistic stress", *Journal of the Acoustical Society of America*, vol. 27, pp. 765–768, 1995.
- [19] D. Bolinger, "A theory of pitch accent in English", *Word* vol.14, pp. 109–149, 1958.
- [20] J. Morton and W. Jassem, "Acoustic correlates of stress", *Language and Speech*, vol. 8, pp. 159–181, 1965.
- [21] A. Sluijter and V. van Heuven, 'Spectral balance as an acoustic correlation of linguistic stress', *Journal of the Acoustical Society of America*, vol. 100, pp. 2471-2485, 1996.
- [22] X. N. Shen, "Relative duration as a perceptual cue to stress in Mandarin", *Language and Speech*, vol. 36, pp. 415–433, 1993.
- [23] S. A. Jun and M. Oh, "Acquisition of 2nd language intonation", *Proceedings of International Conference on Spoken Language Processing*, Beijing: China, pp. 76-79, 2000.
- [24] C.T. Best, "A direct realist view of cross-language speech perception", In: W Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, pp. 167–200, 1995.
- [25] C. So, "Cross-language categorization of monosyllabic foreign tones: Effects of phonological and phonetic properties of native language", In: T Stolz, N Nau and C Stroh (Eds.), *Monosyllables: From phonology to typology*. Berlin, Germany: Akademie Verlag, pp. 55–69, 2012.
- [26] C. So and C. T. Best, "Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences". *Language and Speech*, vol.53, pp. 273–293, 2010.
- [27] C. So and C. T. Best, "Categorizing mandarin tones into listeners' native prosodic categories: The role of phonetic properties", *Poznań Studies in Contemporary Linguistics*, vol. 47, no. 1, pp. 133–145, 2011.
- [28] C. So and C. T. Best, "Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories", *Studies in Second Language Acquisition*, vol. 36, pp. 195–221, 2014.
- [29] Y. C. Chang, P. Halle, C. T. Best and A. Abramson, "Do non-native language listeners perceive Mandarin tone continua categorically?" In *8th Phonetic Conference of China and the International Symposium on Phonetic proceedings, Frontiers: Beijing*, 2008.
- [30] C. So and C. T. Best, "Do English speakers assimilate Mandarin tones to English prosodic categories?" In *INTERSPEECH: the 9th Annual Conference of the International Speech Communication Association proceedings, incorporating SST*, Brisbane, 2008, pp. 22–26.
- [31] S. Luo and H. Lin, "English learners' perception and production of Mandarin intonation". In *18-ICPhS, Glasgow: Scotland, proceedings*, 2015.
- [32] Y. Zhang, S. L. Nissen and A. L. Francis, "Acoustic characteristics of English lexical stress produced by native Mandarin speakers", *The Journal of the Acoustical Society of America*, vol. 123, pp. 4498–4513, 2008.