10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

A Comparative study of rhythmic patterns in non-native Mandarin Speech By Russian, Japanese and Vietnamese learners

Yiran Ding, Wang Dai, Kaiqi Fu, Yanlu Xie, Jinsong Zhang

Research institute of international Chinese language education, BLCU
{yiranding.blcu blcudw, kaiq.fu}@gmail.com, {xieyanlu, jinsong.zhang}@blcu.edu.cn

Abstract

The influence of the first language (L1) on the second language's (L2) rhythm is intrinsically complicated. It is still questionable to make a declaration about how an L1 makes impacts on an L2's rhythm. As a trial, this study investigated rhythmic patterns of native and non-native Mandarin speech by Chinese, Russian, Japanese, and Vietnamese speakers. The three L2 mother tongues are regarded to have distinctly different rhythmic patterns, i.e., stress-, mora-, and syllable-timed. Data analyses were carried out using a total of 23,381 Mandarin utterances by 227 speakers in 4 L1s from the BLCU-SAIT speech corpus. Preliminary findings include: 1) all three kinds of L2 Mandarin speech showed, more or less, the pattern of stress-timed rhythm, which is different from the syllable-timed rhythmic pattern of Mandarin; 2) the rhythmic differences between L1 Mandarin and those by three L2s ranged differently with Vietnamese the least, and Russian the most. These findings might suggest that the rhythmic patterns in non-native Mandarin would be possibly conditioned by universal constraints prior to the influence of L1 diversity revealed.

Index Terms: rhythm, non-native Mandarin, universal trends, diversity

1. Introduction

The term speech rhythm refers to the temporal organization of language that could correspond to the auditory impression of isochrony, and languages can be classified into three rhythm classes: stressed-timed, syllable-timed, and mora-timed [1], [2]. Previous studies tried to find out the quantitative measurements through acoustic cues between languages by using metrics of speech rhythms [3], [4]. Based on the variability in the duration of consonant and vowels in a language, there is evidence that languages can be classified into three groups, and these three groups correspond to what has been traditionally delineated as rhythm classes [3].

Speech rhythm is closely related to language acquisition [5-7]. For the first language (L1) acquisition, infants are sensitive to rhythmic patterns and prefer to listen to a language with the rhythm of the mother's language [5]. It is one of the first attributes of the native in which language-users acquire [6] and is claimed as a precursor to language acquisition [7].

In the second language (L2) acquisition, previous studies have shown that rhythmic difference between native and target language would influence the learners' productions. The extent of influence is still in doubt[10]. Some studies have shown that the L1 background could have a (positive or negative) transfer effect on L2 prosody [8], [10], [12]. Ding [12] analyzed the production of L2 Chinese English by 10 Chinese learners, and Kawase [8] investigated the production of L2 English by 4 Japanese learners, their results showed the nonnative speeches located between L1 and L2. Other studies have emphasized the importance of universal tendency[11], [13]. Ordin [13] analyzed the production of L2 English by learners from German and French respectively, and the results showed that learners from rhythmically different L1 backgrounds still followed a similar developmental path, the L1 influence did not occur or even showed a pattern overshooting L2 – hence opposite to the L1 in some studies [9], [10]. The previous research shed light upon the potential interactions of L1 to L2, but this kind of inconsistency might attribute to many factors include both limitations of speech materials, speakers. Speech metrics obtained in relatively small speech materials might reflect both phonotactic and prosodic information, and the difference in speakers' L1 backgrounds, proficiency levels would also contribute to a different result [22]. Therefore, for further studies on L1 influence on L2 rhythm, a comparative study of non-native rhythm is desirable with an extensive speech corpus, which should also cover all rhythm classes in L1 backgrounds.

This study focuses on rhythmic patterns of native and non-native Mandarin. The rhythm class of Mandarin Chinese is often considered as syllable-timed, which has a simple syllable structure, and each syllable has a lexical tone [15]. In order to investigate the potential influence of L1 on L2 production, we selected Russia, Japanese, and Vietnamese learners of Mandarin. Russia is a typical stress-timed, characterized by its complex syllable structure like constant cluster, with the frequent use of vowel reduction [16], which is considered as a distinctly different rhythm compared with Mandarin. Whereas Japanese is a typical mora-timed language [17], a mora is a linguistic unit that is often smaller than a syllable, but in terms of speech rhythm, mora-timed is more similar to syllable-timed than to stress-timed [3]. Lastly, Vietnamese is also a typical syllable-timed language. It has approximately the same syllable structure and also has lexical tones, and thus it is nearly identical rhythm compared to Mandarin [18].

Under the circumstances, the rhythmic similarities between the L1 and L2 can be subdivided into distinctly different (Russian-Mandarin), similar (Japanese-Mandarin), and identical (Vietnamese-Mandarin). The comparisons of those three pairs let us investigate a.) rhythmic patterns of native and non-native Mandarin under different rhythmic similarities or L1 influence; b.) the reliability of the rhythm metrics to quantify the rhythmic difference between native and non-native speech in a large non-native speech corpus study; c.) the possible universal or diverse rhythmic proprieties in L2 Mandarin.

 10.21437/SpeechProsody.2020-110

2. Method

2.1. Rhythm metrics

Durational cues were most easily perceived, and the most frequently used to quantify the difference in rhythm [3], [4], [19],[21]. Accordingly, several measurement procedures have been proposed to investigate rhythmic differences between and within languages. Two primary metrics are pairwise variability index (PVI) and interval measures (IM), which are capable of exploiting syllable complexity, vowel reduction and stress-based lengthening to provide metrics of rhythm [3]. They are deriving from a simple segmentation of the speech string into vocalic and consonant intervals or entire syllables, measuring variability in these intervals. We used the most popular ones presented in the literature and tabulated them below. (See Loukina et al. [20] for a detailed review)

Table 1: *Summary of the rhythmic metrics*

Metrics	.description	author
%V	Percentage of vocalic intervals	[4]
ΔC/ΔV	Standard deviation of vocalic/consonant intervals	[4]
Varco(C,V)	ΔV(C) /mean vocalic/consonant duration	[19]
rPVI(C, V)	Raw PVI of vocalic/consonantal intervals	[3]
nPVI(C, V)	Normalized PVI of vocalic/consonantal intervals	[3]
ΔS	Standard deviation of syllables	[21]
VarcoS	ΔS/mean syllable duration	[21]
r/nPVI-S	Raw/normalized PVI of syllables	[21]

For IM metrics, Ramus [4] showed the typical stress-timed language that displayed a high ΔC, a lower %V. Conversely, a typical syllable-timed was characterized as relatively lower ΔC with higher %V, in mora-timed language, ΔC went lower and %V raised more. For PVI metrics, Grabe [3] showed that stress-timed language had relatively higher rPVI-c and nPVI-v values compared with syllable-timed, and there was a small difference between mora-timed and syllable-timed language in PVI scores. Besides metrics mentioned above, Mok [21] found that syllable-related metrics also gave a good separation between stress-timed and syllable-timed languages, especially for rPVI-s and ΔS. Stress-timed language like English might have higher values and the values of syllable-timed language are relatively lower.

2.2. Data collection and evaluation

Previous studies have used relatively small corpora during analysis by using rhythm metrics. There is increasing evidence shows that rhythm metrics can vary greatly between speakers or texts [22], [23], a large speech corpus, thereby becomes essential for an extensive rhythm study. Considered the critical reviews above, the following aspects were slightly refined in order to validate the reliability of experiments: a). total 227 speakers were selected to minimize the influence of between-speaker variability; b.) an extensive and identical reading material under the same language (Mandarin) was used to rectify the shortcomings of reflection-only of syllable structures and variations between different texts.

The data was selected from the BLCU-SAIT speech corpus of Mandarin [25]. The corpus consists of both native and non-native Mandarin speech with various L1 backgrounds. Our data included 66 Chinese native speakers (CN), 61 Vietnamese learners (VN), 48 Russian learners (RU) and 52 Japanese learners (JP) of Mandarin. The gender ratio was approximately 6:4, and each speaker has produced 103 identical sentences (1304 Chinese characters). All sentences were selected from their textbook in class and easily understood by L2 Mandarin learners at the intermediate level.

Table2: *Data setup.*

L1 backgrounds	Number of speaker	Sentences Each speaker	Total Sentences
Mandarin	66(F39/M27)	103	6798
Russian	48(F29/M19)	~	4944
Japanese	52(F36/M16)	~	5356
Vietnamese	61(F39/M22)	~	6283
	227(total)		23381(total)

Before IM and PVI metrics calculated, the first step was the segmentation of the speech signals into vocalic and consonant intervals. Segmentation was conducted by adapting HTK-based force-alignment. Different from previous works of Mandarin rhythm, we employed a relatively reasonable but straightforward segmentation method. In Mandarin, there is substantial coherence within the syllable, and relatively weak connection between syllables [24], thus compound vowels in Mandarin is hard to subdivide. For nasal finals, previous studies found out the second and third formant (F2, F3) in vowel parts (including vowel nucleus and nasalized vowel) is the primary perceptual cues for nasal finals [26]. Therefore, we treated the compound vowels included diphthong, triphthong as single vowel interval, and also treated nasal finals as an integrated vowel rather than separate it into vowel nucleus, nasalized vowel and nasal coda.

Finally, the boundaries of the syllables, vowels, and consonants were manually recalibrated during the related annotation task based on audio and visual inspection of the waveform and the spectrogram [27] by Praat [28]. All rhythmic metrics mentioned above were calculated for each sentence and then averaged for each speaker.

3. Result

3.1. Result of rhythm metric

Firstly, we carried out a series of ANOVAs for rhythm metrics mentioned above with the four L1 groups (CN, RU, JP, VN) as the between-group factor. Table 2 gave the statistical results of six basic rhythmic metrics among the four groups. The left showed the metrics and p-values for ANOVA, while the right showed the results of the Bonferroni post-hoc test.

Table 3: *ANOVA results with post-hoc test*

Rhythm	Bonferroni post-hoc test					
Metrics	CN/JP	CN/RU	CN/VN	JP/VN	JP/RU	VN/RU
%V*	***	1.000	1.000	***	***	1.000
ΔC*	***	***	***	.449	***	***
rPVI_c*	***	***	***	.681	***	***
nPVI_v*	***	1.000	***	.118	***	***
ΔS*	***	***	***	***	***	***
rPVI_S*	***	***	***	**	***	***

* indicates p < 0.05; while ** indicates p < 0.01; *** indicated p <0.001

As shown in Table3, all metrics show significant main effects (at p <.001). Most of the Pairwise comparisons (by Bonferroni) between the L1s show that they differed significantly from each other, and only several pairwise comparisons did not show the significant effects like *%V* in *three pairs (*CN/VN, RU/VN, VN-RU); *ΔC and rPVI-c in* JP-VNand *nPVI-v in two pairs (*CN-RU, JP-VN). To investigate those metrics directly, we compared the six metrics separately in the box-plot blow.

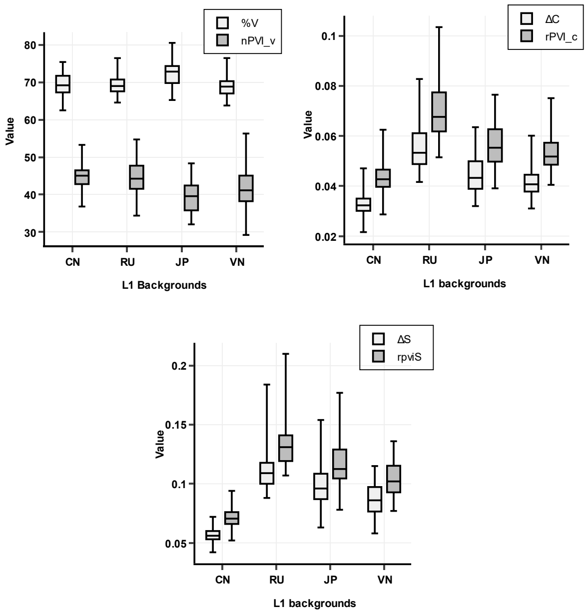

Figure 1: *plots of vowel/consonant/syllable metrics*

Two preliminary findings are observed in Figure 1. On the one hand, only *nPVI-v* scores are marginally confirmed general transfer hypothesis, the RU learners exhibit a relatively higher score than native speakers, and JP and VN learners show a relatively lower value than native speakers. On the other hand, *ΔC* and *rPVI-c* slightly conflict general transfer hypothesis, those values in all learners' groups are higher than native speakers, which may indicate the partial existence of a stress-timed pattern in L2 Mandarin regardless of their L1backgourds.

Furthermore, by the comparison of those syllable-related metrics, we can find out the substantially stabilized patterns in native Mandarin, which has low variation in syllable level. Instead, all values in the learners' groups present a relatively higher variation. This result suggests the existence of higher syllabic variation in non-native Mandarin

3.2. Discrimination of the groups

In confronting the difficulty of encapsulating the cross-influences in rhythm with just a pair of metrics, the secondary analysis aimed at understanding the distribution of the rhythmic patterns of speakers according to their L1 backgrounds with the help of Principal Component Analysis (PCA). The idea was to reliably figure out the correlation or distance between different groups (CN, RU, JP, VN) by using the 13 rhythmic parameters presented above. The illustration of the 2d PCA result was plotted below.

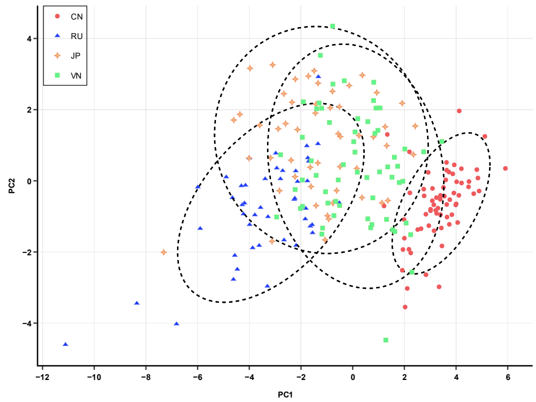

Figure 2: *Illustration of PCA result (Proportion of Variance: PC1 = 0.612; PC2 = 0.167)*

As shown in Figure 2. Group Mandarin (in red circle) locate in the middle right. Group Russian learners (in blue triangle) locate at the middle left. The group Japanese learners (in auburn star) and Vietnamese learners (in green square) distribute between them with the partial mixtures.

The result suggests at least the distance between native and non-native is distinguishable, especially for RU and CN. It partially supported that the distinctly different rhythm between L1 and L2 might have a negative transfer to cause the bigger accentness of their production. Within non-native groups, despite the L1-L2 rhythmic similarities in non-native groups were different, the distance between each group compared with native is not clear, each group kept a certain degree of confusion between others, especially for JP and VN. We might preliminarily assume that learners still could not acquire the native-like rhythm directly by the benefit of positive transfer, no matter the L1 and L2 rhythm is similar or even identical.

3.2.1. Classification

In order to confirm this consistent tendency of 'stress-timed' in all learners' groups, we conducted several classification procedures to drivel them into different numbers of clusters by adopting those features (metrics) mentioned above. Firstly, it was a three-class classification task between Russia, Japanese learners of Mandarin, and native speakers, which corresponding to three typical rhythm types in their native speech. Secondly, it was a four-class classification task with the addition of the Vietnamese group. We used three different statistical models to train and test. All features were normalized by z-score and tested under 5-fold cross-validation.

Table 3: *Classification accuracy*

Method	**3-class**	**4 class**
Logistic regression	87.20%	***74.15%***
SVM	***90.27%***	73.3%
LightGBM	89.41%	72.60%

As shown in Table 3, the overall accuracy is decent enough with limited features and showed the partial effectiveness of rhythm metrics in accentness classification tasks. All models reach an acceptable accuracy in the 3-class task, but the accuracy drops almost 15% when countered into the 4-class task. Then we took the following procedure by investigating the confusion matrix in two tasks to figure out the possible reason that might contribute to this decline. We manually divided data into five pairs of train and test data with balanced speakers with different L1 backgrounds and tested again by using LightGBM and added up all confusion matrix.

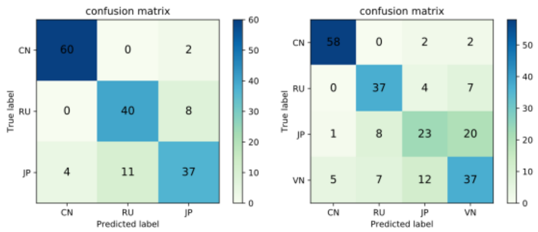

Figure 3: *confusion matrix of 3/4 classification task*

As shown in Figure 3, in the three-class classification task, speakers in the Mandarin group classified approximately correct, only two speakers were misplaced into the Japanese group. For speakers in the Japanese group. Fifteen speakers misclassified into Mandarin and Russian. For Speakers in the Russian group, eight speakers mismatched into the Japanese group. Briefly, there was no obvious confusion between native and non-native speeches and little confusion between Russian and Japanese. In the four-class classification task, the Mandarin group held relatively higher accuracy compared to non-native groups. Instead, the accuracy of classification in non-native groups was not decent enough. Serval speakers misclassified into Russian, and most of the mistakes were occur between Japanese and Vietnamese groups. The confusion matrix suggested that the rhythmic patterns were distinctly different between L1/L2, but kept kinds of confusion in L2 groups. Following the result of the rhythm metric, it might indicate the existence of more or less stress-timed patterns in L2 speech.

4. Discussion

In order to reveal and compare the potential L1 influence on L2 rhythm, we analyzed an extensive corpus with native and three different non-native Mandarin speakers. Based on the results of rhythm metrics analysis and discrimination tasks, the three preliminary findings are worth discussing more.

Firstly, the identical text was produced by all the speakers. There is no difference in rhythmic structure (syllable structure) between their productions, and the significant difference between each group was contributed to the L1 particularity indeed. We found out the distinctly different between native and non-native speech reflected by both rhythm metrics and discrimination tasks, which is in line with our auditory perceptions. It is not hard for native speakers to perceive this discernible accentness beyond segmental information.

Secondly, by the comparison of rhythmic metrics and results of discrimination tasks, our studies suggested that universal constraints would possibly condition the rhythmic patterns in non-native Mandarin. All learners presented more or less stress-timed features in their non-native speech regardless of L1 backgrounds. Moreover, the result of the discrimination task presented quite a confusion between each non-native group. It has slightly conflicted with the general hypothesis (i.e., L1 transfer effect) that the rhythmic pattern of nonnative speech should be intermediate between L1 and L2. Previous research has also reported this phenomenon in non-native Mandarin, speech rate, hypercorrection, or selective lengthening in different prosodic structure have been used for explanations [8], [29]. Thus, it might be a relatively common issue in the L2 Mandarin rhythm. These findings let us suggest that acquisition of Mandarin rhythm might support the central tenet of the Structural Conformity Hypothesis (SCH)—universal trends in the linguistic development that hold for primary languages also hold for interlanguages [30]. L2 rhythm acquisition is claimed to be a multisystemic process [31] that requires the simultaneous attainment of several language-specific features, both phonotactic and prosodic. In the phonological studies of L2 Mandarin acquisition, most learners would always confront difficulties during the acquisition of retroflex and alveolo-palatal consonants; alveolar and velar nasals; Tone3 and neutral tones, these phonemes and tones are prone to systematic errors and always take a long time to become stabilized. Their intrinsic durations might contribute to the difference of syllable duration directly hence to take an impact in rhythmic patterns. This phenomenon does coincide with our auditory impression of nonnative Mandarin speech – generally syllable-timed but meanwhile accompanied by some systemic mispronunciation and misapplication of reduction or lengthening strategy.

Finally, despite the mutual existence of partial stress-timed features, each group still showed the difference to some degrees. By the benefit of the extensive corpus analysis, each metric in each group approximately distributed in normality. In the Russian group, some syllable-related metrics kept quite consistently high values compared with native. Conversely, in both Japanese and Vietnamese groups, a certain percentage of speakers did reach native-like values, and other speakers' values lay between Russian and Mandarin groups' values. We could further speculate that the influence of L1 specific would still reveal in rhythm after it was conditioned by universal constraints. The distinctly rhythmic differences between Russian and Mandarin might delay or halt their rhythmic development by the negative transfer and cause a kind of fossilization. The similar rhythm between Japanese and Vietnamese compared to Mandarin might accelerate or promote their rhythmic development and reach a native-like rhythm earlier by the benefit of positive transfer.

5. Conclusion

This work compared the rhythmic patterns of Mandarin speech between native speakers and three groups of L2 speakers who were native in Russian, Japanese and Vietnamese, respectively. The detailed results obtained from the analysis of the extensive L2 data reconfirmed the usefulness of durational metrics in the distinction between native and non-native speech and might shed light on the universal constraints in non-native Mandarin rhythmic patterns before the L1 diversity revealed. Some specific properties link to the rhythm like syllabic duration are higher learnable and teachable, and the combination of classes and computer-assisted language learning systems will be helpful to facilitate the naturalness of learners with adequate exercises and correct feedbacks.

6. Acknowledge

This study was supported by Advanced Innovation Center for Language Resource and Intelligence (KYR17005), the Science Foundation and Special Program for Key Basic Research fund of Beijing Language and Culture University (the Fundamental Research Funds for the Central Universities) (20YJ040002, 16ZDJ03), and the project of "Intelligent Speech technology International Exchange." Jinsong Zhang is the corresponding author.

7. References

[1] P. Ladefoged and K. Johnson, *A Course in Phonetics (6th ed)*. 2011 Wadsworth: Cengage Learning.

[2] D. Abercrombie, *Elements of general phonetics*. Edinburgh: Edinburgh Univ. Press, 2006.

[3] E. Grabe and E.L. Low, "Durational variability in speech and the rhythm class hypothesis", in N. Warner and C. Gussenhoven [Eds], Papers in Laboratory Phonology, 7: 515-546, Berlin: Mouton de Gruyter, 2002.

[4] F. Ramus, M. Nespor and J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition*, vol. 73, no. 3, pp. 265-292, 1999.

[5] J. Kim, C. Davis and A. Cutler, "Perceptual Tests of Rhythmic Similarity: II. Syllable Rhythm," *Language and Speech*, vol. 51, no. 4, pp. 343-359, 2008.

[6] L. Murty, T. Otake and A. Cutler, "Perceptual Tests of Rhythmic Similarity: I. Mora Rhythm," *Language and Speech*, vol. 50, no. 1, pp. 77-99, 2007.

[7] U. Goswami, "Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective," *Annals of the New York Academy of Sciences*, vol. 1453, no. 1, pp. 67-78, 2019.

[8] S. Kawase, J. Kim, C. Davis, "The influence of second language experience on Japanese-accented" *Proc. 7th International Conference on Speech Prosody 2016*, 2016, 746-750.

[9] English rhythm W. Gu and K. Hirose, "Rhythmic Patterns in Native and Non-Native Mandarin Speech," *Proc. 7th International Conference on Speech Prosody 2014*, 2014, pp. 592-596.

[10] L. White and S. Mattys, "Calibrating rhythm: First language and second language studies," *Journal of Phonetics*, vol. 35, no. 4, pp. 501-522, 2007.

[11] M. Ordin and L. Polyanskaya, "Acquisition of speech rhythm in a second language by learners with rhythmically different native languages," *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 533-544, 2015.

[12] H. Ding and X. Xu, "L2 English Rhythm in Read Speech by Chinese Students," *In Proc. Interspeech 2016*, pp. 2696-2700.

[13] P. P. Mok, and V. Dellwo, "Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English," *Proceedings of Speech Prosody*, 2008.

[14] A. Tortel, and D. Hirst, "Rhythm metrics and the production of English L1/L2." *Proceedings of Speech Prosody, 2010*.

[15] H. Lin and Q. Wang, "Mandarin rhythm: An acoustic study", *Journal of Chinese Language and Computing*, vol. 217, no.3, pp. 127-140, 2007.

[16] R.M. Dauer, "Stress-timing and syllable-timing reanalyzed", *Journal of Phonetics*, vol. 11, no. 1, pp. 51-62, 1983.

[17] L. White, "Acquiring rhythm: A comparison of L1 and L2 speakers of Canadian English and Japanese", *In Proceedings of the 32nd Boston University Conference on Language Development*, 2008, pp. 155-166.

[18] A. Romano, P. Mairano and L. Calabròc, "Measures of speech Rhythm in East-Asian tonal Languages," *Proc. 17th ICPHS*, 2011, pp. 2693-2696.

[19] V. Dellwo, "Rhythm and speech rate: A variation coefficient for ΔC", in P. Karnowski, & I. Szigeti [Eds], *Language and language Processing*, 231-241, Frankfurt: Peter Lang, 2006.

[20] A. Loukina, G. Kochanski, B. Rosner, E. Keane and C. Shih, "Rhythm measures and dimensions of durational variation in speech," *The Journal of the Acoustical Society of America*, vol. 129, no. 5, pp. 3258-3270, 2011.

[21] P. Mok, "On the syllable-timing of Cantonese and Beijing Mandarin" *Chinese Journal of Phonetics*, pp. 148-154, 2009.

[22] A. Arvaniti, "Rhythm, Timing and the Timing of Rhythm," *Phonetica*, vol. 66, no. 1-2, pp. 46-63, 2009.

[23] L. Wiget, L. White, B. Schuppler, I. Grenon, O. Rauch, and S. L. Mattys, "How stable are acoustic metrics of contrastive speech rhythm," *The Journal of the Acoustical Society of America*, vol. 127, no. 3, pp. 1559–1569, 2010.

[24] Z.-J Wu and M.-C Lin. *Essentials of Experimental Phonetics[M]*, Higher Education Press, Beijing, 1989

[25] B. Wu, Y. Xie, L. Lu, C. Cao and J. Zhang, "The construction of a Chinese interlanguage corpus," *Conference of The Oriental Chapter of International Committee for Coordination and Standardization of Speech Databases and Assessment Techniques (O-COCOSDA)*, Bali, 2016, pp. 183-187.

[26] Y. Li, Y. Xie, L. Feng and J. Zhang, "The perceptual cues for nasal finals in standard Chinese," *10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Tianjin, 2016, pp. 1-5.

[27] W. Wang, W. Wei, Y. Xie, M. Guo and J. Zhang, "Improve the Accuracy of Non-native Speech Annotation with a Semi-automatic Approach," *11th International Symposium on Chinese Spoken Language Processing (ISCSLP),* Taipei City, Taiwan, 2018, pp. 116-120.

[28] P. Boersma and D. Weenink. Praat: doing phonetics by computer [computer program]. Version 6.0.50, retrieved 03 August 2018 from http://www.praat.org

[29] S. Shuju, X. Yanlu, F. Xiaoli and Z. Jinsong, "Automatic detection of rhythmic patterns in native and L2 speech: Chinese, Japanese, and Japanese L2 Chinese," *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Tianjin, 2016, pp. 1-4.

[30] F. R. Eckman, "The Structural Conformity Hypothesis and the Acquisition of Consonant Clusters in the Interlanguage of ESL Learners," *Studies in Second Language Acquisition*, vol. 13, no. 1, pp. 23–41, 1991.

[31] A. Li and B. Post, "L2 Acquisition Of Prosodic Properties Of Speech Rhythm," *Studies in Second Language Acquisition*, vol. 36, no. 2, pp. 223–255, 2014.