SUPRASEGMENTAL DIFFERENCES BETWEEN CHINESE AND INTERNATIONAL ENGLISH INSTRUCTORS IN AN EFL CONTEXT

Xianghui Meng, Beijing International Studies University
Difei Lu, Beijing International Studies University

The comparison of the suprasegmental properties between native English speaker (NES) speech and second/foreign language learner speech exists in abundance. However, little research has been conducted to analyze the pronunciation features of language instructors, not least in an EFL setting. This study aimed to find the suprasegmental differences between Chinese English instructors and international instructors in an EFL classroom context. In order to evaluate this, we compared a two-minute spontaneous sample extracted from an actual lecture given in English by 10 college instructors from China (Mandarin speakers) and 10 college instructors from different English-speaking countries (native English speakers). The speech samples were analyzed in terms of eleven features of interest: syllable per second, articulation rate, mean length of run, phonation time ratio, number of silent pauses, mean length of pauses, number of filled pauses, mean length of filled pauses, number of prominent syllables per run, proportion of prominent syllables, and overall pitch range. The independent T-Tests performed on each variable separately revealed that no significant differences existed between the two groups of instructors except for mean length of pauses. Despite the lack of significantly statistical differences in the majority of the features examined, this study uncovered some revealing findings that are instrumental in enhancing the classroom speech quality for both Chinese and international instructors.

INTRODUCTION

Pronunciation teaching is key to the development of speaking abilities (Gilakjani & Sabouri, 2016). Unfortunately, in many ESL and EFL classrooms, pronunciation teaching receives a paucity of attention compared to other skills involved in foreign language teaching and learning (Abbas, 2011). The reasons for EFL instructors ignoring pronunciation teaching are many and varied. In a study that examined Iranian teachers' view toward pronunciation instruction, Abbas and Narjes (2016) concluded that lack of time, motivation, resources, materials and educational facilities all precluded teachers from teaching pronunciation in class. However, teachers' neglect of this area of instruction does not belie the students' desire to improve their pronunciation with the help of their instructors. In fact, a great number of EFL leaners recognize the importance of pronunciation (Alghazo, 2015) and express their willingness to reduce their accent so as to sound native-like (e.g., British RP or General American). Chinese EFL teachers play an indispensable role in shaping the pronunciation of their students, because for most Chinese students learning happens mostly in the classroom (Chen, 2011), where the teacher usually speaks throughout the whole class. A teacher with a clear and highly intelligible pronunciation could exert a good influence on the students, inspiring and motivating them to

improve their own pronunciation. By contrast, a teacher with a rather mediocre pronunciation tends to produce the opposite effect. It is for this reason that EFL instructor speech is worth investigating.

Literature Review

A fairly large body of research has been dedicated to the suprasegmental differences between ESL and EFL learners and native English speakers (NESs). In a speaking rate consistency study in native and nonnative speakers of English, Baese-Berk and Morrill (2015) analyzed read speech by L1 Korean, L1 Mandarin and NESs and calculated the mean speaking rates (syllable per second) of those three groups of speakers. Their results lent support to a previous study that demonstrates that L2 speakers tend to speak more slowly than do native speakers (Munro and Derwing, 1995). Pausing is another major feature that distinguishes native and nonnative speech patterns. Anderson-Hsieh and Venkatagiri (1994) investigated the speech of Chinese ESL speakers of varying proficiency levels and discovered that intermediate proficient speakers were less fluent for having longer and more frequent pauses than higher proficiency speakers. Tavakoli (2011) examined the pausing patterns of L2 learners and native speakers narrating picture stories and revealed that in addition to more frequent and longer duration of pauses that are generally associated with nonnative speech, L2 speakers paused more frequently in the middle of clauses. Chen et al. (2001) compared the sentence stress production between 40 Mandarin speakers and 40 American native speakers. Their results indicated that Chinese speakers produced stressed vowels with shorter duration and unstressed vowels with greater volume than NSs. In a more recent study on the production of English contrastive focus, Yeung et al. (2019), after comparing 26 Mandarin speakers and 21 NESs, came to the same conclusion that Mandarin speakers produced focused words with shorter durations. In addition to speaking rate, pausing and stress, pitch has also been identified as a prominent feature distinguishing native speaker (NS) and nonnative speaker (NNS) speech. Li and Lan (2011) compared speech read by two dialects of Chinese speakers—Mandarin and Cantonese—to that of NSs and revealed two major findings: first, Chinese speakers' mean syllable durations were much longer, suggesting that they had a slower speech rate than NSs; secondly, NSs had a much wider pitch range than both groups of Chinese speakers. The narrow pitch range of NNSs is consistent with other research findings. Pickering's (2004) comparative analysis of the instructional discourse of native speaker teaching assistants and international teaching assistants (ITAs) in a North American university also established a reduced pitch range as a characteristic of NNS speech.

As can be seen from the review, previous research into the acoustic differences between native and nonnative speech have largely been restricted to ESL and EFL learners. Very few attempts have been made to explore the speech properties of native and nonnative English instructors in an EFL classroom setting.

The Current Study

This research sets out to look at the suprasegmental properties of EFL instructors in a Chinese university. The subjects are both local Chinese teachers and international teachers from four English-speaking countries. If it is postulated that Chinese students in general favor native English accents over nonnative English accents (e.g., Mandarin-accented English), then comparing and contrasting the suprasegmental differences between Chinese and international English instructors might yield revealing findings regarding how the two groups vary in terms

of speech rate, pause, prominence and pitch range. The results could also have pedagogical implications for Chinese teachers who wish to refine their own pronunciation.

The study is guided by the following research question:

RQ What suprasegmental features of classroom speech are significantly different between Chinese instructors and international instructors?

METHODS

Participants

Ten Chinese teachers (1 male and 9 female) and ten international teachers (1 female and 9 male), who all taught English to first and second-year college non-English majors, were recruited in this study. The Chinese teachers had a teaching experience ranging from five to twenty-six years. The speech samples were collected from various English courses they were teaching, i.e., intensive reading, public speaking, writing, and business English. The international teachers all came from English-speaking countries, i.e., the US, the UK, Canada and Australia, each with their characteristic accent. They all had a minimum of two-year experience of teaching English as foreign language, which is one of the conditions of employment as an international language teacher in this country.

Suprasegmental Measures

The analysis was carried out on four aspects of suprasegmental measures: speech rate, pause, stress and pitch. Adopting the analysis framework used by Kormos and Dénes (2004) and by Kang and Pickering (2010), the following table presents a complete list of these measures along with their operational definitions.

Table 1

Suprasegmental Measures

Measures	Submeasures	Descriptions
Rate	1) Syllable per second	Calculated by the total number of syllables divided by the total length of speech
	2) Articulation rate	Calculated by the total number of syllables divided by the total length of speech excluding silent pauses
	3) Mean length of run (MLR)	Calculated by total number of syllables divided by the number of utterances produced between .1 second and above
	4) Phonation time ratio	Calculated by the amount of time speaking excluding silent pauses divided by the total length of the speech
Pause	5) Number of silent pauses	Identified by the pauses that are .1 second long and above
	6) Mean length of pauses (second)	Calculated by the total length of silent pauses divided by the total number of silent pauses
	7) Number of filled pauses	Identified by hesitation markers, e.g., "um", "well", etc. but does not include repetitions, restarts or repairs
	8) Mean length of filled pauses (second)	Calculated by the total length of filled pauses divided by the total number of filled pauses
Stress	9) Number of prominent syllables per run (pace)	Calculated by the total number of stressed words divided by the total number of utterances produced between .1 second and above
	10) Proportion of prominent syllables (space)	Calculated by the total number of stressed words divided by the total number of words
Pitch	11) Overall pitch range (-) (Hz)	Calculated by the difference between points of F_0 maximum and minimum on prominent syllables

Sample Collection
Because of the nature of this study, the speech samples collected had to be spontaneous and continuous, uninterrupted by extraneous pauses induced by board work, text reading, and planned or unplanned interactions with students. However, to obtain a sample of such specification in a typical EFL class where interaction between learners and teachers was frequent proved to be a tall order. In addition, Chinese instructors' use of L1 was yet another unwanted interference that needed to be ironed out before sample collection. L1 use in an EFL classroom setting is common, but it is unsuitable for the purpose of this research. As often as not, the Chinese instructors would lecture in English for about one minute, and then, naturally and unwittingly, switch to Mandarin to explain or supplement what they had described in English. To ensure that the speech samples collected were qualified for analysis and that the collection process was as efficient as possible, all instructors had been informed beforehand that they needed to speak on a subject of learning, e.g., vocabulary knowledge, for at least two minutes without interacting with students in any way or explain what they have said in Mandarin (as in the case of Chinese instructors). A 45-minute session was recorded of each instructor on a computer software Praat, and then a 2-minute continuous and unedited speech sample was extracted from the 45-minute-long recording to make sure that the sample to be analyzed was natural and spontaneous.

Sample Analysis
The samples were rated by the two authors of this study. First of all, each of the samples was transcribed, and the number of syllables and words were calculated using the same syllable counter. Secondly, each speech file was loaded onto Praat, a software that helps identify the location and length of pauses. All the silent or filled pauses (over .1 second) were marked by double slash followed by parentheses where the lengths were indicated. In the meantime, prominent syllables or words were identified as those that are higher in pitcher, longer in duration and louder in volume (Kang, 2010), and were boldfaced with their pitch value marked on top. The second author rated 2 speech samples (one Chinese and one international) on the length of pauses, identification of prominent words and their respective pitch values. The calculation of interrater consistency on the identification of prominent words was based on percentage exact agreement. Firstly, the words that were marked as prominent by both raters were pinpointed and their number was counted. This number was then divided by the total number of words that were marked as prominent by either rater. Length of pauses and pitch value were examined by Pearson correlation. Results are presented in Table 2.

Table 2
Interrater Consistency of Variables

Variable	Chinese instructor	International instructor
Silent pauses	.945	.757
Filled pauses	.762	N/A
Number of prominent words	59%	72%
Pitch value	.871	.956

RESULTS

The results of the suprasegmental features of the two groups of instructors are presented in Table 3 and Table 4. The statistics of these two tables were then subjected to independent T-tests, which yielded the suprasegmental differences presented in Table 5. As there are multiple features being compared simultaneously, the alpha value is Bonferroni-corrected to .005 (i.e.,.05/11≈.005).

Table 3

Suprasegmental Measures of Chinese Instructors

Measures	Sub-measures		Gu	Lu	Mengli	Qing	Qiong	Suli	Xiaobo	Xiaolin	Yuange	Zhencong	Avr	SD
Rate	1)	Syllable per second	2.64	2.66	3.50	2.97	3.82	3.32	3.47	3.44	3.49	2.66	3.20	0.43
	2)	Articulation rate	4.02	3.27	4.44	3.67	4.45	3.68	3.99	4.10	4.23	3.31	3.92	0.42
	3)	Mean length of run (MLR)	5.28	6.61	5.47	5.46	7.19	8.70	7.03	7.50	8.00	5.71	6.70	1.19
	4)	Phonation time ratio (%)	65.56	81.39	78.76	80.98	85.76	90.10	86.95	83.95	82.55	80.39	0.82	0.07
Pause	5)	Number of silent pauses	61	50	77	66	65	47	60	57	53	57	59.30	8.71
	6)	Mean length of pauses (second)	0.68	0.45	0.33	0.34	0.26	0.25	0.26	0.34	0.39	0.41	0.37	0.13
	7)	Number of filled pauses	8	6	22	12	4	0	18	2	1	14	8.70	7.54
	8)	Mean length of filled pauses (second)	0.32	0.38	0.36	0.41	0.33	0.00	0.37	0.41	0.38	0.47	0.34	0.13
Stress	9)	Number of prominent syllables per run (pace)	0.95	1.73	1.16	1.11	2.19	2.48	1.85	1.96	1.58	1.57	1.66	0.49
	10)	Proportion of prominent syllables (space)	0.24	0.35	0.34	0.28	0.42	0.41	0.39	0.38	0.30	0.37	0.35	0.06
Pitch	11)	Overall pitch range (-) (Hz)	257.50	272.70	220.20	209.80	345.20	223.40	300.10	303.00	232.00	96.20	246.01	68.30

Table 4

Suprasegmental Measures of International Instructors

Measures	Sub-measures	Chris	David	Graham	James	John	Kash	Ling	Michael	Phil	Tristan	Avr	SD
Rate	1) Syllable per second	3.49	3.76	2.75	3.81	3.00	3.00	3.59	2.71	2.34	2.99	3.14	0.49
	2) Articulation rate	4.78	4.31	4.38	5.38	4.19	3.96	4.59	3.44	4.79	4.35	4.42	0.52
	3) Mean length of run (MLR)	7.00	12.66	6.73	7.13	5.49	7.80	7.55	6.71	5.22	7.22	7.35	2.04
	4) Phonation time ratio (%)	73.03	87.23	62.77	62.77	70.78	71.71	75.82	78.82	48.86	68.63	0.72	0.10
Pause	5) Number of silent pauses	61	36	50	65	66	0	57	49	55	51	49.00	19.33
	6) Mean length of pauses (second)	0.53	0.42	0.89	0.54	0.51	0.62	0.45	0.51	1.12	0.74	0.63	0.22
	7) Number of filled pauses	0	11	0	8	10	0	4	0	0	4	3.70	4.47
	8) Mean length of filled pauses (second)	0.00	0.41	0.00	0.44	0.52	0.00	0.42	0.00	0.00	0.42	0.22	0.23
Stress	9) Number of prominent syllables per run (pace)	1.30	1.69	1.53	1.13	1.28	1.74	1.75	1.42	1.35	1.38	1.46	0.21
	10) Proportion of prominent syllables (space)	0.29	0.20	0.29	0.21	0.31	0.34	0.38	0.21	0.33	0.24	0.28	0.06
Pitch	11) Overall pitch range (-) (Hz)	145.10	125.00	118.80	166.66	176.20	84.90	224.90	131.30	95.20	161.30	142.94	41.39

Table 5

Suprasegmental Differences Between the Two Groups of Instructors

Measures	Sub-measures	t	df	Sig.
Rate	Syllable per second	.253	18	.803
	Articulation rate	-2.351	18	.030
	Mean length of run (MLR)	-.879	18	.391
	Phonation time ratio	2.588	18	.019
Pause	Number of silent pauses	1.536	18	.142
	Mean length of pauses (second)	-3.211	18	.005
	Number of filled pauses	1.803	18	.088
	Mean length of filled pauses (second)	1.424	13.895	.177
Stress	Number of prominent syllables per run (pace)	1.196	12.295	.254
	Proportion of prominent syllables (space)	2.436	18	.025
Pitch	Semitone range	.513	18	.614

Rate

From the analyses of the twenty English teacher speech files, we can see that both Chinese and international instructors are fluent speakers, but it may be interesting to see that Chinese instructors were found to have a slightly higher average speech rate (3.2 syllables per second) than international instructors (3.14 syllables per second). Independent T test also showed that no significant difference existed between the two groups in this parameter (t=.253, df=18, p>.005).

**Figure 1**

Speech rate (syllable per second) of both Chinese and International instructors
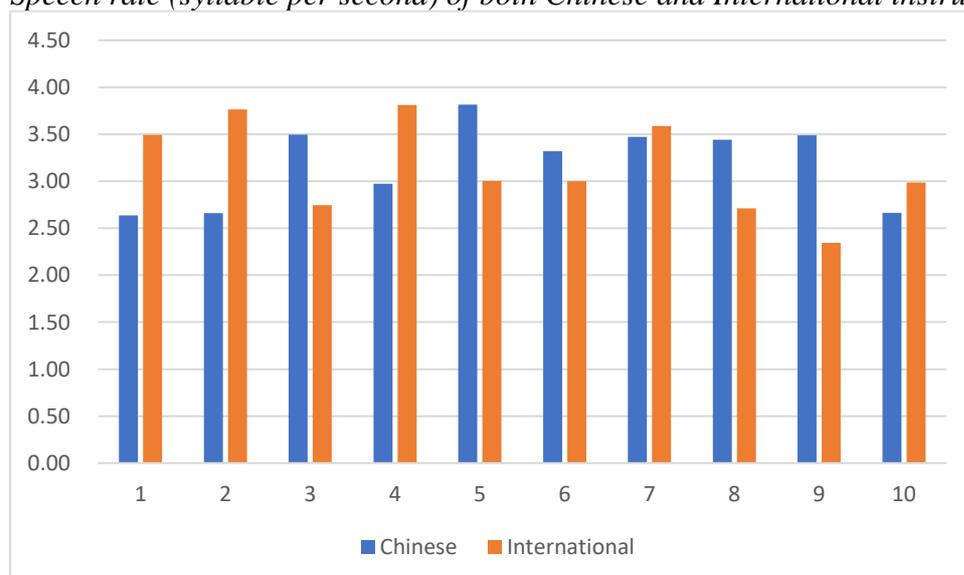
Pause

One acoustic measure where significant difference was found between the two groups was mean length of silent pauses (*t=-3.211, df=18, p<0.05*). It turns out that international instructors tend to pause longer than Chinese instructors (see Figure 2).

Figure 2

Mean length of pauses of both Chinese and international instructors
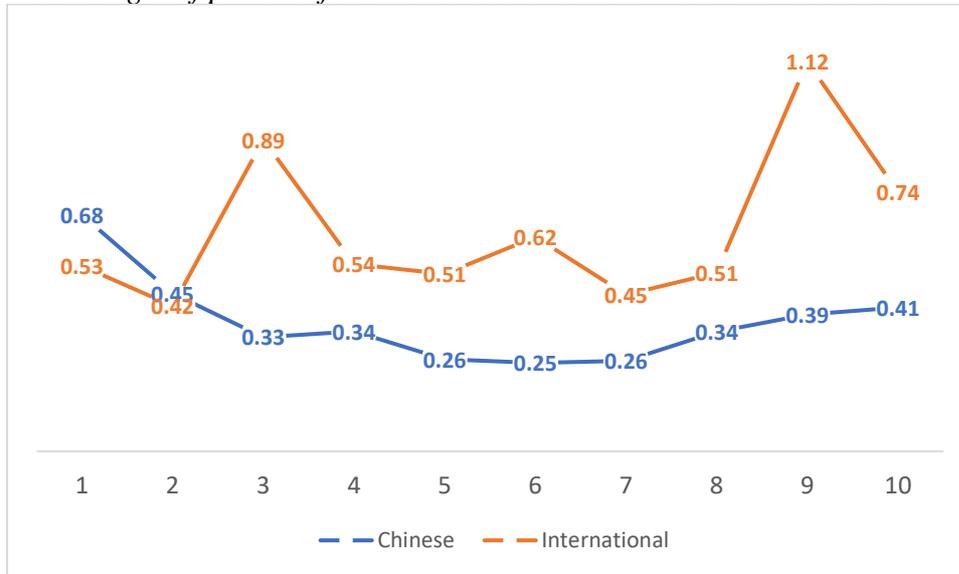
Despite the fact that the Chinese instructors had shorter pauses, further enquiry into the distribution of pauses demonstrated that they also tended to pause more frequently within T units (see Figure 3), with a mean rate of 2.2 pauses per unit (calculated by the total number of pauses within T units divided by the total number of T units), which is significantly higher than the international instructors (t=2.262, df=10.418, p=.005).

Figure 3

Pauses (indicated by the blank areas) within a T unit produced by a Chinese instructor
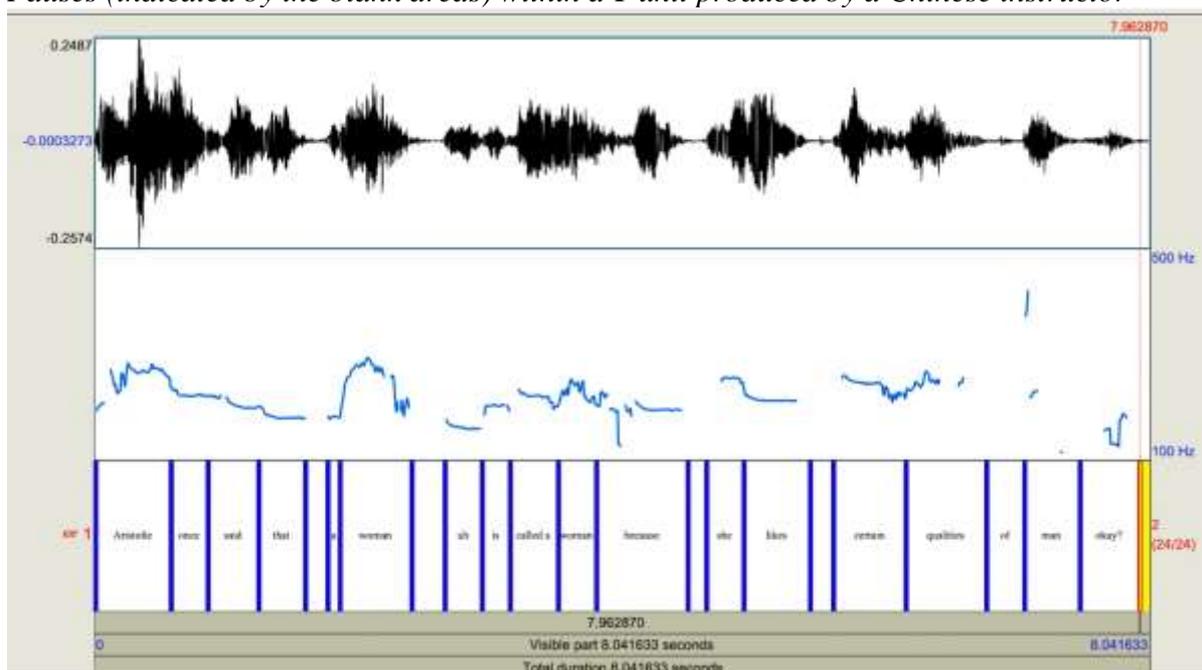
Stress

Significant difference was found in neither the number of prominent syllables per run (pace) ($t=1.196$, $df=12.295$, $p>.005$) nor the proportion of prominent syllables (space) ($t=2.436$, $df=18$, $p>.005$) produced by the two groups of instructors.

Pitch

As gender was not evenly distributed in both Chinese (1 man and 9 women) and international (1 woman and 9 men) instructors, comparing male and female speech files will naturally lead to the result that the former have a wider pitch range than the latter. To avoid this issue, pitch value was converted into semitone interval D, using the formula $D = \frac{12}{\log_{10} 2} \cdot \log_{10} \frac{f_1}{f_2}$ (Hart, Collier and Cohen, 1990), where f_1 represents the maximum or minimum pitch value and f_2 represents the average pitch value of each speaker. The comparison of D values showed that no significant difference existed between the two groups in terms of overall pitch range. ($t=.513$, $df=18$, $p>.005$).

DISCUSSION

Chinese instructors' ostensibly faster speech rate can be expounded from two aspects. Firstly, in an EFL context, NS instructors would talk at a slower-than-normal rate lest students have trouble understanding them. They would also adjust their speech rate based on the perception of their students' level of proficiency. For example, NS grade-school teachers speak at an even slower rate to children. On the other hand, the Chinese instructors' marginally faster speech rate could be explained by their familiarity with the content they teach. Fluency is a task-dependent variable (Lam & Wendy, 1994). The more involved and unfamiliar the speaking task is, the less fluency with which the speaker will perform. All the Chinese teachers involved in this study had had years of experience teaching the same content, which enables them to deliver their classes in English with much less conceptualizing difficulty that would otherwise contribute to dysfluency.

The pause difference does not indicate that the international instructors are less fluent than their Chinese counterparts. Listeners are sensitive to the location of pauses, and pauses within clauses are reflective of reduced cognitive fluency (Kahng, 2020). As a result, a faster speech rate and shorter pauses do not necessarily contribute to a speaker's fluency.

Stress or prominence is a salient suprasegmental feature that sets NSs and NNSs apart. Tyler, Jefferies and Davies (1988) found that NSs were able to avail themselves of extensive prosodic devices that would help their listeners better differentiate the important message of their speech, and simultaneously signal the interconnections between their ideas. However, such abilities were inadequate among NNSs. In fact, they found that Chinese and Korean ITAs assigned more or less equal stress to all words in an utterance. The current study, however, reveals no significant differences between the stress measures (i.e., pace and space) of the Chinese and NS instructors.

Apart from stress, pitch range is regarded as yet another distinguishing feature of native and nonnative speech. Study by Kang (2010) concluded that the wider the pitch range, the less accented the speaker would be. Besides, previous research suggests that a

narrow pitch range is characteristic of EFL speech. For instance, Grazia Busà and Urbani (2011) compared the differences in pitch range between American NSs and Italian NNSs and partially confirmed their presumption that Italian speaker had a narrower pitch and less pitch variations. However, the result yielded by the current study on Chinese and native English instructors' pitch range fails to lend support to previous findings. The inconsistency could be attributed to the speech style, i.e., formal vs. informal speech, of NS teachers, who often feel the need to speak "unnaturally" in an EFL setting, for example, with a slower-than-normal speech rate, to ensure that they are understood by every student.

CONCLUSION

This study probes into the suprasegmental differences between Chinese instructors and international instructors in an EFL context, and reveals that only the mean length of pauses is significantly different between the two groups of instructors. It is possible that segmental features, e.g., Voice Onset Time (VOT), vowel formants, vowel duration, etc. contribute to a larger part of the differences between the two groups of instructors than suprasegmental features, which is another research area worth exploring. Nevertheless, this study reveals findings that could improve the speech quality of both native and Chinese English instructors. Firstly, the research shows that NS EFL teachers have a slightly slower speech rate than NNS Chinese teachers. Speech rate is a matter of concern for some international EFL teachers, who are inclined to believe that talking at a natural rate will render their instructional speech less intelligible to students. The finding of this study could in a way alleviate their concern and prompt them to think that a natural speech rate is conceivably more conducive to college students' English learning. Secondly, Chinese instructors can work on their pause distribution to improve their perceived fluency. Despite the tendency to speak faster and pause shorter than NS instructors, their speech displays way more interpolations within clauses, which undermines their perceived fluency.

ABOUT THE AUTHORS

Xianghui Meng is a lecturer of English at Beijing International Studies University. His current research interest centers around second language pronunciation acquisition, L2 learners' perception of native and nonnative English accents, and World Englishes in relation to China English.
Email: mxhrobert@126.com

Difei Lu is a lecturer of English at Beijing International Studies University. Her research focuses on the acquisition of second language by adult learners, with a special interest in L2 speech perception between World Englishes and American tongue in university English courses.
Email: ludifei2017@126.com

REFERENCES

Abbas, P. G. (2011). A study on the situation of pronunciation instruction in ESL/EFL classrooms. *Journal of Studies in Education*, *1*(1), 1-15.

Abbas, P. G. & Narjes, B. S. (2016). Why is English pronunciation ignored by EFL teachers in their classes?. *International Journal of English Linguistics*, *6*(6), 195-208.

Alghazo, S. M. (2015). Advanced EFL learners' beliefs about pronunciation teaching. *International Education Studies*, *8*(11), 63-76.

Anderson-Hsieh, J. & Venkatagiri, H. (1994). Syllable duration and pausing in the speech of Chinese ESL speakers. *TESOL Quarterly*, *28*(4)

Baese-Berk, M. M. & Morrill, T. H. (2015). Speaking rate consistency in native and non-native speakers of English. *The Journal of the Acoustical Society of America*, *138*(3)

Chen, H.C. (2011). Judgments of intelligibility and foreign accent by listeners of different language backgrounds. *Journal of Asia TEFL*, *8*(4), 61-83.

Chen, Y, Robb, M. P, Gilbert, H. R. & Lerman, J. W. (2001). A study of sentence stress production in Mandarin speakers of American English. *Acoustical Society of America*, *109*(4), 1681-1690.

Gilakjani, A. P. & Sabouri, N. B. (2016). How can EFL teachers help EFL learners improve their English pronunciation?. *Journal of Language Teaching and Research*, *7*(5), 967-972.

Grazia Busà, M. & Urbani, M. (2011). A cross linguistic analysis of pitch range in English L1 and L2. *17th International Congress of Phonetic Sciences* (pp. 380-383).

Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody (Cambridge studies in speech science and communication)*. Cambridge: Cambridge University Press.

Kahng, J. (2020). Explaining second language utterance fluency: contribution of cognitive fluency and first language utterance fluency. *Applied Psycholinguistics*, *41*(2), 457-480.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*(2), 301–315.

Kang, O., & Pickering, D. R. (2010). Suprasegmental measures of accentedness and judgements of language learner proficiency in oral English. *Modern Language Journal*, *94*(4), 554-566.

Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145–164.

Lam, Wendy Y. K. (1994). Investigating the oral fluency of 15 EFL teachers: a quantitative approach revisited. *International Language in Education Conference,* December 16, 1994*,* Hong Kong.

Li, B. & Lan, S. (2011). Suprasegmental features of Chinese-accented English. *Journal of the Acoustical Society of America*, *129*(4)

Munro, M.J. & Derwing, T.M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*(3), 289-306.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes*, *23*(1), 19-43.

Tavakoli, P. (2011). Pausing patterns: differences between L2 learners and native speakers. *ELT Journal*, *65*(1), 71-79.

Tyler, A. E, Jefferies, A. A. & Davies, C. E. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teacher's spoken discourse. *World Englishes*, *7*(2), 101–110.

Yeung, Alex, Baek, Hyunah, Takahashi, Chikako, Duncan, Joseph, Benedett, Sharon, Hwang, Jiwon, Broselow, Ellen (2019). Pitch range, intensity, and vocal fry in non-native and native english focus intonation. *The Journal of the Acoustical Society of America*, *145*, 1826 - 1826.