Prosodic Characteristics in the Speech of Chinese EFL Learners.

Veronika Makarova and Xia Zhou

Department of Languages and Linguistics University of Saskatchewan, Canada

Abstract

This study reports some prosodic characteristics in the quasispontaneous classroom speech of Chinese EFL learners. Recordings of ten dialogues produced by twenty second-year non-English majors were analyzed to extract the following features: durations of inter- and intra-turn pauses, duration of filled-in pauses, numbers of words per tone unit, tone unit durations, speech rates and pitch accent type (tone) statistics. The deviations from standard native speech in the areas of tonality and tonicity are also considered. The paper offers some practical suggestions aimed at improving the prosodic characteristics of the English speech of Chinese EFL learners.

1.Introduction

The development of global economy and the advancement of English as the International language impose strong demands on EFL teachers to assist qualified specialists in achieving high English communication proficiency. It has been argued that the capability to deliver intelligible speech is the key to success in professional communication [1] and that pronunciation of non-native speakers of English is 'the area that most threatens intelligibility' [2, p.1]. Therefore most EFL/ESL teachers now agree that pronunciation training should be an essential part of an oral communication course [3, 4].

In many Asian countries, such as Korea, EFL teaching is beginning to turn towards the emphasis on oral English training with some attention being paid to aspects of pronunciation [5]. Another example of this trend is the suggested goals of Japanese EFL which include the development of 'the empirical research capacity to evaluate and compare the efficiency of different pronunciation tasks, activities and methods' [6, p. 117].

Unfortunately, the teaching of oral English has not yet become a focal part of language curricula in China, where still much emphasis is put on written texts, vocabulary, grammar and preparation for written tests. Consequently, pronunciation training still remains 'Cinderella' [7] of the Chinese EFL, and hardly ever gets any attention. It is worth examining how Chinese EFL learners perform under these circumstances, and whether they manage to acquire adequate pronunciation skills without explicit pronunciation training.

The aspect of pronunciation acquisition which has recently been attracting attention of researchers is prosody and intonation, since the entire spoken language acquisition (both L1 and L2) pivots on pronunciation and particularly prosody [8]. The role of intonation is essential for many communicative goals, such as to signal interrupting, asking for clarification, taking the floor, changing the subject, concluding an argument, or constraining a hearer to reply [9, p. 295]. Some typical prosodic errors of EFL learners are described, for example, in [10]. However, very little information is available on the prosodic characteristics of the English speech of Chinese learners, and most available sources focus on segmental qualities or on re-stating the importance of general pronunciation teaching techniques in Chinese EFL context [1, 11, 12, 13]. The goal of current study therefore was to obtain some experimental evidence of the prosodic parameters in the speech of Chinese EFL learners.

In particular, this study focuses on dialogue and turn durations, speech rates and durations of pauses. Low speech rates and long pauses are known to contribute to the impression of the foreign accent, as well as to be linked to lower listening comprehension and general L2 ability [14, 15]. The study also reports some preliminary findings for tonal statistics, the choice of the accent placement and chunking of speech flow into tone units, since errors in these parameters have been reported as common characteristics of non-native speech [10].

2. Materials and methods

Subjects

Twenty non-English major second-year students (6 females and 14 males) participated in the study. The students are enrolled in the School of Automation in one of the national Chinese universities. The students can be considered advanced learners, since they have been studying English for over 6 years by the time of the recording, and successfully passed CET-4 or CET-6 Chinese English Proficiency tests. Besides the formal grammar-translation style English class, they have also been enrolled in an advanced Listening and Conversation class. None of the subjects have ever received any specific training in English pronunciation skills.

Methods

The experimental recording took part during a routine language class. The learners were given an oral task of making a conversation in pairs on a given topic (refusing an invitation to a birthday party or asking for help with something). The learners had only a few minutes to prepare, so their speech can considered quasi-spontaneous. The resulting be ten conversations were recorded and videotaped. The video record allowed to investigate the gesticulation and facial expressions in addition to the speech (the paralinguistic data are not included into the current paper). The sound files were digitized at 16000 Hz sampling rate, transcribed, and analyzed with Praat software package to obtain the characteristics of durations and pitch. The transcription included prosodic (tonetic) marking of tone-unit boundaries, stresses and tones in accented syllables. This paper reports the following set of prosodic characteristics: dialogue durations, turn durations, speech rates (including and excluding pauses), durations of inter-turn pauses, durations of intra-turn pauses, duration of filled-in pauses, tonal statistics (types of pitch accents), errors

in the tonality system (splitting speech into tone units), errors in the tonicity system (choosing the accented word).

3. Results.

Dialogue durations, speech rates and turn durations

These parameters are reported below in Table 1. All the dialogues were quite short, ranging from 21 to 80 seconds. The number of turns in the conversations varied from 4 to 15. The speech rates are extremely low, ranging from 1.46 wps (87.6 wpm) to 2.65 wps (159 wpm), with the average speech rate of 1.96 wps (117.5 wpm). However, as will be shown later, the low speech rates are explained by long pauses, but not by the subjects' inability to talk fast.

 Table 1. Dialogue durations, speech rates, and turn durations in the conversations

Dia log ues	Dialogue durations, sec	speech rates (words per second)	average turn durations, sec*		
1	36	2.03	A:2.7; B:1.9; 2.2		
2	44	2.2	A:3.2; B:1,0; 2.6		
3	80	1.94	A:4.7; B:3.1; 3.8		
4	61	2.52	A:2.2; B:2.1; 2.2		
5	60	1.46	A:3.5; B:4.1; 3.8		
6	37	1.54	A:2.8; B:3.7; 3.1		
7	21	2.28	A:2.3; B:2.0; 2.1		
8	66	1.88	A:2.9; B:3.8; 3.2		
9	47	1.83	A:2.1; B:2.1; 2.0		
10	28	2.65	A:1.7; B:3.8; 2.5		

Note: * A indicates the first, B the second speaker, the average speech rate for the dialogue follows.

Inter-turn pauses

Inter-turn pauses are pauses between the speakers' turns in the conversations. The average number of inter-turn pauses in a conversation is 7.4 with the average duration of 759ms. The pauses can last for up to 5 seconds, and there were only five cases of zero pauses between turns (i.e., no pause) (Ref Table 2 for details).

Intra-turn pauses

The intra-turn pauses are pauses within each turn excluding filled-in pauses, which are reported separately below. Intraturn pauses last for about 303 ms, which makes them 2.5 times shorter than intra-turn pauses. They range from 0 (no pause between tone units) to 3.7 sec. The durations are given in Table 3 below. About every fourth or fifth tone unit is separated from the neighboring one not by an actual silent pause, but by other means (such as tone boundary or filled-in pause), which is reflected in the '0' column of the table (giving numbers of 0 pauses between tone-units within a turn).

Table 2 Durations of Inter-turn pauses, ms

Dial	Ν	Mean	St	Media	Max	Min	0
	ра	dur,	Dev	n	dur	dur	*
	use	ms					
1	8	393	318	378	1119	114	0
2	7	637	425	430	1364	143	0
3	13	452	354	419	1413	0	2
4	10	496	290	465	894	0	1
5	10	528	478	444	1387	0	2
6	5	1578	1956	647	5000	329	0
7	3	1160	385	1081	1579	820	0
8	5	668	146	640	907	530	0
9	8	1158	1072	1047	263	3566	0
10	5	522	308	371	236	860	0
All	7.4	759	573	455	5000	0	0
dial							

*0 column gives the numbers of zero pauses between turns in dialogues.

Table 3. Durations of intra-turn pauses by speaker, ms

	Ν	Mean	Medi				
	pau		St	an,	Max	Min	
Speaker	ses	ms	Dev	ms	dur	dur	
1 D1A	5	190	211	124	552	0	
2 D1B	9	167	172	125	431	0	
3 D2A	24	368	323	314	1181	0	
4 D2B	0	0	0	0	0	0	
5 D3A	34	480	417	391	1598	0	
6 D3B	16	449	255	370	1108	140	
7 D4A	9	143	118	170	295	0	
8 D4B	20	138	170	40	480	0	
9 D5A	19	388	376	327	1377	0	
10 D5B	19	462	386	437	1592	0	
11 D6A	9	222	245	179	733	0	
12 D6B	14	252	217	152	580	0	
13 D7A	5	99	138	60	333	0	
14 D7B	7	404	581	128	1650	0	
15 D8A	21	541	798	502	3794	0	
16 D8B	17	542	365	402	1222	0	
17 D9A	13	247	248	222	863	0	
18 D9B	5	318	167	363	547	153	
19D10A	6	228	159	222	507	72	
20D10B	7	419	256	473	731	60	

Tone-unit duration

The average length of a tone-unit in words, syllables, and in ms is reported below in Table 4. The table also provides 'intratone-unit speech rates', which are calculated as the total number of tone-units uttered by one speaker over the total duration of tone-units, whereby all the pauses and filled-in pauses are excluded from the analysis.

The table demonstrates that the number of words in a tone-unit is very small. The by-speaker values range from 1.67 to 3.7 words per tone-unit. The duration of tone-unit in ms varies between 492 and 1115ms. This is explained by the high frequency of one-, two- and three-word tone units, as shown by median values. Most words are monosyllabic, whereby the average number of syllables per tone unit is 3.57 (ranging from 2.72 to 4.73 syllables per tone unit.

In contrast to overall speech rates given above in Table 1, the intra-tone-unit speech rate values in Table 4 are quite high, with the average of 3.33 wps (or 199.8 wpm), and approach the ones reported for native English speech [14].

Table 4. Duration of tone-units (TU) in words (w) and ms; intra-TU speech rates (wps)

	TU dur	ation, w	ords		TU dur, ms	Intra TU sp rate	TU dur, syll
Speaker	Mean	Med	Min	Max	Mean	Mean	Mean
1 D1A	2.57	2	1	6	928	2.77	3.29
2 D1B	2.69	2	1	7	924	2.91	3.85
3 D2A	2.79	2	1	7	757	3.68	3.17
4 D2B	1.67	1	1	4	590	2.82	2.33
5 D3A	3.74	3	1	7	972	3.84	4.05
6 D3B	2.85	2	1	7	858	3	3.23
7 D4A	3.69	4	1	8	1115	3.31	4.62
8 D4B	3.12	2.5	1	8	861	3.62	4.19
9 D5A	2.29	2	1	4	746	3.06	3.14
10 D5B	2.22	2	1	7	738	3.01	2.72
11 D6A	2.17	1	1	5	492	3.3	3
12 D6 B	2.71	2	1	8	665	4.08	4
13 D7A	2.43	3	1	4	727	3.34	3.7
14 D7B	2.78	3	1	6	719	3.86	3.78
15 D8A	2.42	2	1	8	762	3.17	3.08
16 D8B	2.44	2	1	4	714	3.41	2.84
17 D9A	2.7	2	1	7	737	3.66	3.31
18 D9B	2.73	2	1	6	948	2.88	4
19 10A	3.27	3	1	5	994	3.62	4.73
20 10B	3.63	3	1	11	1004	3.61	4.45

Parameter interactions

ANOVA and correlation analysis of selected parameters show some interactions:

Dialogue duration vs N pauses : p<0.001; cor=0.77 Dialogue duration vs speech rate : p<0.001; cor=-0.31 Dialogue duration vs TU duration: p<0.001; cor=-0.31

Filled-in pauses

Most speakers (14 out of 20) use an abundance of filled-in hesitation pauses. The total number of hesitation pauses in nine conversations is 42, with the average duration of 392 ms (dialogue 4 produced by two female speakers contained no filled-in pauses). The maximum number of hesitation pauses is found in the turns of Speaker D5B (male), who used 9 hesitation pauses with the average duration of 435ms. The hesitation pauses have mostly $[\Box]$ –like quality, with the exception of 3 pauses which were bilabial nasals [m].

Tonal statistics

The analysis of the pitch accents (tones) showed that the Chinese speakers of English employed the total of 342 tones in all the conversations. The terminal and non-terminal tones included:

 $\begin{array}{l} HF & - \ 67 \ (19.5\%), \\ HR & - \ 65 \ (19\%), \\ MF & - \ 58 \ (17\%), \\ RF & - \ 51 \ (15\%), \\ LF & - \ 45 \ (13\%), \\ Level & - \ 45 \ (13\%), \\ LR & - \ 12 \ (4\%) \end{array}$

This distribution of tones strongly differs from the one found in native English speech [16]. In particular, the high numbers of high rises, levels and rise-falls, and the absence of fall-rises draws attention.

Errors in tonality and tonicity domains

Following Halliday [17], three major domains of intonation can be differentiated: tonality (chunking of speech into toneunits), tonicity (choosing the most important word in a TU and assigning a tone to it) and tone (selecting a correct tone).

Chinese learners of English make a number of errors in all these domains. English tonality rules are violated by isolating pronouns, post-positions, 'but' and other grammar words into separate tone-units, e.g., Dialogue 1SA turn 1: 'I have something / very happy to tell/you', Dialogue 1SB turn 6: 'my/terminal examination is coming/ I have/ to get ready for it/so...', Dialogue 2 SA turn 3 /I/err/I want to/.../ask you for help/'.

Most common errors in tonicity include the placing of a pitch accent on a pronoun or preposition. E.g., Dialogue 2 Speaker A turn 5: 'so/I ask for <u>you</u>/for help/', /can <u>you</u>/help me <u>to</u>/ finish it/; Dialogue 4 SB turn 2 (tonality, tonicity and tone errors): / <u>But</u>/ II am <u>lso >sorry</u> / <u>But</u>/ I have a very important appointment tomorrow'.

Errors in tone assignment include the frequent use of high rises and level tones, insufficient use of low rises for questions, the use of rise-falls with non-final tone-units, lack of rise-falls.

Other errors include the frequent use of monotone and flat TU contours, small excursions of falls, small or no pitch movement in the accented syllables with the falling tones, insufficient or no vowel reduction, lack of prominence contrast between stressed and unstressed syllables (i.e. syllables in a TU may have relatively equal lengths of tonebearing units and little or no prominence contrast), lack of liaison, and 'separate' articulation of words. These features will be addressed in fuller details elsewhere.

4. Discussion

We have seen that despite their long history of EFL learning, the subjects have low overall speech rates. There is an interesting similarity between the speech rates of EFL learners in their own speech production and the speech rates that they are comfortable with in perception. Foreign learners of English are more comfortable at speech rates of about 125-160 wpm, and an increase of speech rate above 200 wpm (considered to be of moderate speed) causes difficulties in listening comprehension [14, 18, 19].

An interesting turn in this experiment is the difference in overall speech rates and intra-tone-unit speech rates, which demonstrates that the speakers can actually articulate English words at high speeds, but the speech rates drops because of lengthy pauses. Very long pauses and hesitation pauses can be annoying to a conversation partner, and the teachers may consider making learners acquainted with other conversational strategies used to gain time. In terms of further developments for research, it may also be useful to investigate whether insertion of longer pauses into chunks of high-speed speech enhances L2 listening comprehension (as opposed to general lowering of speech rate). If so, it may be an alternative way of training FL learners to comprehend speech at higher speeds.

The abundance of pauses and low numbers of words per tone unit suggest that the learners need pronunciation exercises aimed at increasing the numbers of words linked together (like a 'snowball' word game, where one has to repeat everything said by the previous speakers plus add a new word to the expanding phrase).

Some of the prosodic errors in the speech of Chinese learners are similar to the typical errors of non-native speakers, such as the lack of vowel reduction, problems with liaison and other rhythmical problems, wrong chunking of speech across and against grammatical boundaries, assigning accents to grammar words, etc. [10]. However, some other errors connected to the assignment of tones appear to be specific to Chinese speakers. The preference for level tones, high rises probably takes its root in the L1 interference and the prosodic system of Chinese, which has high level, rising and falling tones. However, the use of rise-falls is more difficult to explain, since Mandarin and Cantonese Chinese have no risefall tone. The use of rise-falls needs further explanations, presumably related to the frequent combinations of Chinese tones (such as two falling tones together or a rise followed by fall which yield a pattern resembling a rise-fall) [20, 21]

5. Conclusion

Despite the fact that Chinese learners of English have been studying English for over 6 years and passed English proficiency tests, their speech exhibits a number of prosodic deficiencies, including low speech rates and long pauses, small number of words per tone unit, errors in rhythm, insufficient liaison, errors in tonality, tonicity and tone. It appears that the learners would benefit from more attention given to oral practice in class and from explicit pronunciation teaching, as recommended by modern EFL methodology [4, 22, 23]

References:

[1] Jian, H-L., 2005. Improving the communicative competence in spoken English of Taiwanese engineering students: A prosodic approach. *iCEER-2005 Proceedings*, Tainan, Taiwan, 1-5 March 2005.

http://www.iaalab.ncku.edu.tw/iceer2005/Form/PaperFile/14-002.pd

[2] Jenkins, J., 2000. *The Phonology of English as an International Language*. Oxford: OUP.

[3] Fraser, H., 1999. ESL pronunciation teaching: Could it be more effective? *Australian Language Matters*, 7(4), 7-8.

[4] Morley, J., 1991. The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25 (1), 51-74.

[5] Bell, M. A. 1996. Teaching pronunciation and intonation to EFL learners in Korea <u>http://www.geocities.com/Athens/</u> <u>Acropolis/9583/PRONUN.html</u>

[6] Makarova, V. & Rodgers, T. (Eds.), 2004. *English language teaching: The case of Japan*. Muenchen: Lincom.

[7] Kelly, L. G., 1969. *25 centuries of language teaching*. Rowley, MA: Newbury House.

[8]_Kjellin, O., 1999. Accent Addition: Prosody and Perception Facilitates Second Language Learning.

In O. Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings* of LP'98 (Linguistics and Phonetics Conference) at Ohio State University, Columbus, Ohio, September 1998, Vol. 2, 373-398. Prague: The Karolinum Press.

[9]_Chun, D. M., 1988. The Neglected Role of Intonation in Communicative Competence and Proficiency, *The Modern Language Journal*, 72 (3), 295-303.

[10] Celce-Murcia, M., Brinton, M., Goodwin, J.M., 1996. *Teaching Pronunciation: A Reference for Teachers of English* to Speakers of Other Languages Cambridge: Cambridge University Press

[11] Lin, Hsiang-Pao, 1995. *Teaching pronunciation in the learner-centered classroom*. Paper presented at the TEFL Conference, 20 May, Taichung, Taiwan.

[12]. Gong, J. 2002. Introducing English rhythm in Chinese EFL classrooms: A literature review. *Post-Script*, 3 (1), 26-42. [13] Cheng Fangzhi, 1998. The teaching of pronunciation to Chinese students of English. *Forum*, 36 (1). 37-41.

[14] Griffith, R., 1990. Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Applied Linguistics*, 40, 311-335.

[15] Hirai, A. 2002.*The relationship between Listening and reading rates*. Inst Modern Languages & Cultures, University of Tsukuba

Cruttenden, A., 1997. Intonation. Cambridge, CUP.

[16] Halliday, M. A. K., 1970. A course in spoken English: Intonation. London: OUP.

[17] Blau, E. K., 1990. The effect of syntax, speed and pauses on listening comprehension. *TESOL Quarterly*, 24, 746-752.

[18] Zhao, Y. The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics*, 18 (1), 49-68.

[19] Shih, C & Kochanski, G. P., 2000. Chinese tonemodeling with Stem-ML. *Proceedings of the International Conference on Spoken Language Processing*, Volume 2, (2000) pp. 67-70.. Beijing, China.

[20] Yuan, J, Shih, C and Kochanski, G P., 2002. Comparison of declarative and interrogative intonation in Cantonese. *Proceedings of the Speech Prosody 2002 Conference*, 711-714.. Aix-en-Provence, Laboratoire Parole et Langage

[21] Kramsch, C. J., 1983. Interaction in the classroom: Learning to negotiate roles and meaning. Unterrichtspraxis, 16, 175-90.

[22] Morley, J., 1996. Second language speech/ pronunciation: acquisition, instruction, standards, variation and accent. In: Altais, J. E. (Ed). *Georgetown University Round Table on Languages and Linguistics. Linguistics, language acquisition and language variation: current trends and future prospects:* 138-158. Georgetown University Press, Washington, DC.