

# The Influence of Breathing Function in Speech on Mastering English Pronunciation by Chinese Students

Irina Yu. Pavlovskaya

Doctor of Philology, St. Petersburg State University  
Saint Petersburg, Russia  
pavlovskayairina2@gmail.com

Lan Hao

B.A., St. Petersburg State University  
Saint Petersburg, Russia  
707481966@qq.com

**Abstract** – The relevance of the topic is due to the need to search for effective methods of teaching English pronunciation skills to speakers of ideographic languages. Thus, the Chinese language belongs to the group of «syllabic» (logographic, ideographic) languages, each syllable being represented with a certain hieroglyph and associated with a certain meaning. English is a phonetic (phonographic) language, and the meaning is built up with morphemes, usually consisting of several phonemes represented by letters. The research question is, how breathing habits in contact languages influence the success of learning. We claim that Chinese speakers tend to make a pause after each one or two syllables. English speakers pronounce longer breath groups with a big number of syllables without pauses. The purpose of the paper is to describe the general speech respiratory models for English and Chinese speakers and to measure the difference. It is done with the method of ergospirometry. In addition, the speech breath habits of a smoker and non-smoker are compared.

**Keywords:** *pronunciation, phonographic (phonetic) languages, syllabic (logographic) languages, speech breath, ergospirometry*

## I. INTRODUCTION

In China, as it is well known, the writing system is typically ideographic/logographic. Hieroglyphic writing has been existing in China for over a thousand years. Hieroglyphic symbols are associated directly with the meaning. Each hieroglyph has an individual meaning and corresponds to one syllable. Although, in Chinese there is no direct phonetic connection between pronunciation and spelling, as in phonographic languages [1]. At present there are no universal definitions towards «phonemic language» and «syllabic language» in English or Chinese. In Russia discussions of these two concepts have been continued for many years [2], but most linguists define the Chinese language as a “syllabic” one [3]. There are certain difficulties in acquiring the pronunciation of phonemic languages, such as Russian or English, by logographic learners. There have been some attempts done to analyze these difficulties in terms of phonetics and pedagogy [4-10]. To our knowledge these difficulties have not yet been researched from the point of view of breathing habits in speech.

## II. PRODUCTION OF SPEECH SOUNDS

There has been a wide spectrum of theoretical and experimental works written in the field of the physiology and acoustics of speech in English and Russian [11-15]. In modern phonetics the model of speech production is described as

follows.

The sound of speech is the result of the impact on a person’s auditory organ caused by the vibrational movements in the air. Sounds in general and sounds of speech in particular are considered from the two sides: firstly, to study the objective properties of oscillatory movements: their frequency, amplitude, spectral composition, and secondly, from the point of those sound sensations that somehow correspond to these properties: pitch, volume, timbre. The patterns of perception of sounds are studied in psychoacoustics.

Speech sounds are divided into noises and tones: tones in speech arise as a result of oscillations of vocal folds; noises are formed as a result of non-periodic oscillations of air emerging from the lungs. The tones are usually vowels; almost all voiceless consonants belong to noises. Voiced consonants are formed by merging noise and tones. Noises and tones are examined for their height, timbre, strength, and many other characteristics.

Breathing plays an important role in the physiology of pronunciation. Without breath no sound can be produced. Both noises and tones are the result of the influence of the air puff coming from lungs on the vocal cords and other barriers along the vocal tract. More than that, a person lives - it means that he or she has heart beat and blood circulation. In addition, breathing is the basic provision for the living. External breathing is understood as gas exchange between the body and the environment, including oxygen absorption and carbon dioxide emission, as well as the transport of these gases inside the body through the respiratory tubes (trachea, bronchi) and in the circulatory system. From this, it follows that the internal systems of the human body are always dynamic, not static, regardless of whether a person is moving or not. Therefore, inside human body certain sound vibrations occur, due to the heartbeat and breathing. Sometimes we cannot feel them, because they are infrasonic.

## III. THE ANALYSIS OF VOWELS AND CONSONANTS IN CHINESE AND ENGLISH LANGUAGES

In Tables I and II the number and quality of English and Chinese sounds are presented in IPA symbols and typical mistakes are analyzed. In Tables III and IV we have tried to find the causes of the errors made by Chinese learners.

One of the reasons of the mistakes is the historical development of the Chinese phonetic system. Descriptions of ancient Chinese phonetics are not preserved. Modern Chinese

linguists cannot accurately describe the actual pronunciation of the hieroglyphs in ancient China. In China there are two hieroglyphic systems, Wenyan (written language used in China until the beginning of the twentieth century) and Baihua (the official recording system of modern spoken Chinese). Between the two systems there is a big difference in pronunciation and writing. Modern Chinese linguists cannot fully understand the former pronunciation written in Chinese because one cannot immediately recognize the pronunciation of words only by the form of hieroglyphs.

TABLE I. COMPARISON OF CHINESE AND ENGLISH CONSONANTS

Consonants	English (26)	Chinese (22)
Voiceless consonants	<b>9</b> (/p/ /t/ /k/ /f/ /s/ /ʃ/ /tʃ/ /θ/ /h/)	<b>18</b> (/b/ /p/ /f/ /d/ /t/ /g/ /k/ /h/ /j/ /q/ /x/ /zh/ /ch/ /sh/ /z/ /c/ /s/ /ng/)
Voiced consonants	<b>17</b> (/b/ /d/ /g/ /v/ /z/ /dʒ/ /dz/ /ð/ /ʒ/ /m/ /n/ /ŋ/ /l/ /r/ /w/ /j/)	<b>4</b> (/m/ /n/ /l/ /r/)

TABLE II. COMPARISON OF CHINESE AND ENGLISH VOWELS

Vowels	English (20)	Chinese (39)
Monophthongs	<b>12</b> (/i:/ /ɪ/ /ɜ:/ /ə/ /ɔ:/ /ɒ/ /ʊ/ /u:/ /ʊ/ /a:/ /ʌ/ /æ/ /e/)	<b>10</b> (/a/ /o/ /e/ /è/ /i-/ /-i-/ /-i/ /u/ /ü/ /er/)
Diphthongs	<b>8</b> (/eɪ/ /aɪ/ /ɔɪ/ /aʊ/ /əʊ/ /ɪə/ /eə/ /ʊə/)	<b>13</b> (/ai/ /ei/ /ao/ /ou/ /ia/ /ie/ /ua/ /uo/ /üe/ /iao/ /iou/ /uai/ /uei)
Nasal vowels	<b>0</b>	<b>16</b> (/an/ /ian/ /uan/ /üan/ /en/ /in/ /un/ /ün/ /ang/ /iang/ /uang/ /eng/ /ing/ /ueng/ /ong/ /iong)

TABLE III. TYPICAL MISTAKES MADE BY NATIVE CHINESE SPEAKERS IN PRONOUNCING ENGLISH CONSONANTS

Typical problems	Reasons
Difficulties in pronouncing words which end with voiceless consonants	Historic reason
Difficulties in pronouncing words which end with the sound /m/	Historic reason
Underdifferentiation between /l/ and /n/	Special situation in the South Chinese dialects
Consonant clusters	Don't exist in Chinese language
Pronunciation of /θ/ and /ð/	Don't exist in Chinese language
Pronunciation /v/	Doesn't exist in Chinese language
Pronunciation /r/ and /l/	Underdifferentiation

Ancient four tones: level tone – Ping (平), rising tone – Shang (上), departing tone – Qu (去), “entering” tone – Ru (入), correspond to the pronunciation of the four words “天(Tian), 子(Zi), 圣(Sheng), 哲(Zhe)”, and make up the official

phonetic system of Medieval Chinese (the name of the language of the Southern and Northern dynasties, the Sui dynasty, the Tang dynasty and the Song dynasty, which spread from VI to X century). The first time this system has been mentioned in the book "Qieyun" (Author - Lu Fayuan, the book has already been lost). All of the above is written in the book "Guangyun" (rhyming dictionary), the first book of Chinese phonetics in accordance with the government requirement, 1008. The ancient four tones were officially introduced in the Tang dynasties (618 - 907) and the Song (960 - 1279). The system was created in accordance with the main principle of classification of tones: level tones and oblique tones (rising and falling).

TABLE IV. TYPICAL MISTAKES MADE BY NATIVE CHINESE SPEAKERS IN PRONOUNCING ENGLISH VOWELS

Typical problems	Reasons
Confusion between /i:/ and /ɪ/	Lack of duration and intensity of vowel /i:/
Confusion between /e/ and /æ/	Speakers don't differentiate the opening of the mouth
Confusion between /u:/ and /ʊ/	Speakers don't differentiate the opening of the mouth
Confusion between /oʊ/ and /ɔ:/	Lack of diphthongs
Confusion between /æ/ or /e/ and /ei/	Speakers don't differentiate the opening of the mouth and don't differentiate monophthongs and diphthongs

It is necessary to mention one important historical event in Chinese phonetics: the creation of the Scheme for the Chinese Phonetic Alphabet. First of all, let's look at the introduction of creating modern Chinese Phonetic Alphabet. In 1958 Chinese government decided to create a new phonetic standard which is based on Latin alphabet and which has simplified Chinese characters to unify intonation and improve literacy rate. In the Scheme for the Chinese Phonetic Alphabet we have four tones which correspond to ancient tones: 1) Yin Ping (阴平) – level tone, starts with a voiceless consonant; Yang Ping (阳平) – level tone, starts with a voiced consonant; 3) Shang (上) – rising tone, 4) Qu (去) – departing tone. So in the modern Chinese language the entering tone – Ru (入) is absent. It is absent in general modern Chinese Mandarin, but is preserved in many Southeastern dialects: Cantonese, Yue, Min, and Hakka. The most significant specificity of the tone Ru (入) is the pronunciation of the hieroglyph, ending with voiceless consonants like p, t, k. This sheds light on the reason for the inability of modern Chinese speakers to pronounce some English words correctly.

Modern Chinese Mandarin is based on the Beijing dialect, in which Ru (入) is completely absent. From the Yuan dynasty (1271–1368) to the last feudal Qing dynasty (1644–1912), the capital was moved from the southern city Chang'an (modern name: Xi'an) to Beijing, which means that the official Chinese pronunciation was gradually transferred to the base of the northern Chinese dialect, not the southern Chinese dialect. The

location of the capital of China in the North has a significant impact on the official Chinese pronunciation. There is a big difference between the Southern Chinese dialects and the Northern Chinese dialects [16 – 18].

Besides the historic reasons the difficulties in English pronunciation for Chinese learners can be connected with speech breathing models.

#### IV. DETERMINING SPEECH RESPIRATION MODELS OF ENGLISH AND CHINESE NATIVE SPEAKERS

Respiratory models were worked out empirically in an instrumental experiment.

The experiment was conducted at the Northwestern Federal Medical Research Center named after V.A. Almazov (Saint-Petersburg, Russia) by using ergospirometry techniques. The ergospirometry procedure is the only method of qualitative and quantitative assessment of reactions and interactions of the cardiovascular and respiratory systems, as well as the body's metabolic response during exercise, at the present stage of medicine development. Stress testing with gas analysis has been used since the 70s of the 20<sup>th</sup> century and is considered the reference method for assessing human performance [18].

In our case, ergospirometry was conducted in order to determine speech breathing patterns in native speakers of Chinese and English. We will consider speech as a load component. The ergospirometric recording was carried out simultaneously with the audio recording, which allowed us to synchronize the data and find the correlation of respiratory bursts with certain segments of speech (Fig. 1-9).

In this paper we present experimental data for the three typical representative subjects:

Participant No.1 (Luke) is a British teenager comes from Manchester (The UK), he is 21 years old (his birthday 10/22/1997), his weight is 72 kg, his height is 165 cm, he doesn't smoke and has correct bite, the subject has no diseases.

Participant No.2 (Zhang) is a Chinese girl comes from Xian (Northwest city of China), she is 24 years old (her birthday 10/01/1994), her weight is 62 kg, her weight is 168 cm. she doesn't smoke and has orthognathic bite, the subject struck 4 wisdom teeth.

Participant No.3 (Cui) is a Chinese teenager comes from Jinan (Northeastern city of China), he is 23 years old (his birthday 01/29/1996), his weight is 68 kg, his height is 172 cm, he smoked for 10 years and has orthognathic bite, the subject has a chronic runny nose and hit 3 wisdom teeth.

The biological data of the subjects are important in terms of the purity of the experiment and the comparability of the data.

All subjects read out the stimulus material, consisting of control phrases (11), and oppositions of words with strong and weak consonants at the beginning, middle and end of a word, as well as oppositions of words with tense and lax vowels:

##### **Phrases:**

1. Peter Piper picked a peck of pickled peppers.

2. A peck of pickled peppers Peter Piper picked.
3. If Peter Piper picked a peck of pickled peppers,
4. Where`s the peck of pickled peppers Peter Piper picked?
5. A tutor who tooted the flute, tried to tutor two tooters to toot.
6. A canner can can
7. Anything that he can
8. But he can't can a can,
9. Can he?
10. Betty Botter bought some butter,
11. "But - she said - "the butter`s bitter,
12. If I put it in my batter, it will make me batter bitter,
13. But a bit of better butter,
14. That will make my batter better".
15. She saw Sharif's shoes on the sofa.
16. But was she so sure those were Sharif's shoes she saw?
17. She sells sea shells on the sea shore. The shells she sells are sea shells, I'm sure.
18. How can a clam cram in a clean cream can?
19. How many cookies could a good cook cook, if a good cook could cook cookies?
20. A good cook could cook as much cookies
21. As a good cook who could cook cookies.
22. A skunk sat on a stump and thunk the stump stunk, but the stump thunk the skunk stunk.
23. Red lorry, yellow lorry, red lorry, yellow lorry.
24. Three thin thinkers thinking thick thoughtful thoughts.

##### **Word pairs:**

##### **Vowels**

1. Peach and pitch.
2. Beat and bit.
3. Beach and bitch.
4. Bought and box.
5. Bark and buck.
6. Boot and book.
7. Court and cox.
8. Sheet and shit.
9. Soot and look.
10. Sink and think.
11. Sank and thank.
12. Word and world.

13. Will and wheel and well.

**Consonants**

1. Lard and luck.
2. Pat and bat.
3. Patter and batter.
4. Stop and sob.
5. Pop and Bob.
6. Mop and mob.
7. Tar and dart.
8. Tea and deer.
9. Tat and Dad.
10. Kitty and giddy.
11. Sat and sad.
12. Lot and Lord.
13. Cat and gap.
14. Court and God.
15. Mack and mug.
16. Sock and bag.
17. Becky and Baggy.
18. Locker and jogging.

In the ergospirometric experiment the following variables were investigated:

1. BF - Respiratory rate in liters per minute (l/min)
2. ti / tot - The ratio of inspiratory time to total respiratory cycle time (%)
3. t-in — Inspiratory time in seconds (s)
4. V`E - Minute ventilation of the lungs in liters per minute (l / min)
5. VTin - Respiratory volume in liters (l)
6. t-ex - Expiration time in seconds (s)

These variables are presented in Tables V and VI for each subject. The data is obtained with the help of “Infrasonic wave Analysis inside human-body (Breathing channel)” [19]. In this paper we present a) each variable for all 3 subjects together (Fig. 1-6) four variables for each subject separately (Fig.7-9).

In all figures we observe that the curves fall into two parts: 1) (approximately 3 minutes) – breathing while pronouncing phrases; breathing while pronouncing pairs of words. For all the subjects pronouncing phrases requires more active breathing modulations than pronouncing words. In-breaths in the first part are longer and more frequent than in the second part.

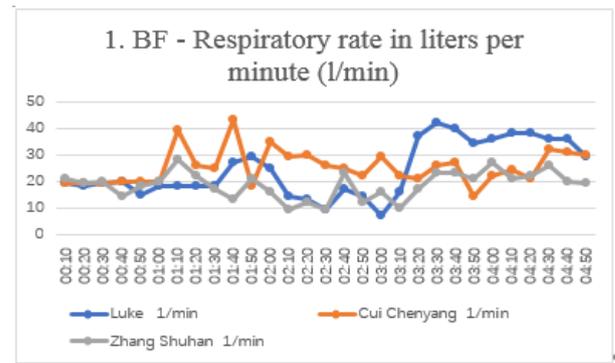


Fig. 1. BF - Respiratory rate in liters per minute (l/min).

Fig. 1 shows that the respiratory rate for Speaker 2 (Zhang, grey curve) is the lowest and varies little (three small peaks in the first part and very small fluctuations in the second). For Speaker 3 (Cui, red curve), who is a smoker, the function changes dramatically in the first part (three very big peaks). Speaker 1 (Luke, the Englishman, blue line), breathes regularly, with a high speed on words and low speed on phrases. Speech does not influence much the rate of his breathing. From Tab.7 it follows that in comparison with the words section, during the phrases section Luke’s respiratory rate increases (+0.03 l/min), while for Chinese subjects, vis a versa, it decreases significantly, (-2,16 for Speaker 2 and -0,19 for Speaker 3). It proves again that senseful signals (phrases) are not a problem for the native English speaker, while the Chinese speakers have to think intensely, they stop and breathe less regularly.

Similar observations can be made on the basis of the other functions.

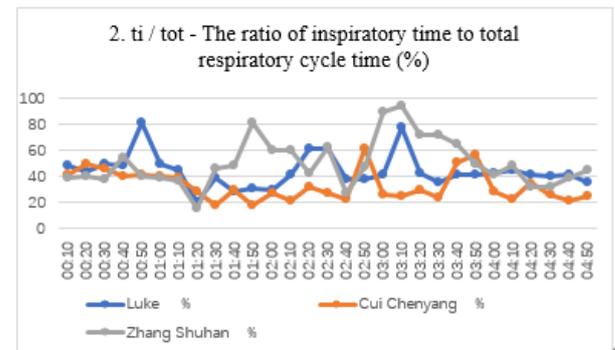


Fig. 2. ti / tot - The ratio of inspiratory time to total respiratory cycle time (%).

On Fig.2 we can observe the biggest peaks for Speaker 2 (nearly 100%). For Speaker 3 the curve doesn’t rise higher than 60%. The average ratio is the lowest for Speaker 1 and the highest for Speaker 2 (See Table VI).

From graph 3 it is clear that the native English speaker makes three long in-breaths, while Speaker 3, the smoker, only two little sighs. Speaker 2 made a very long inhale before the words and two shorter ones at the end of the phrases. May be, she was tired after producing the tongue twisters.

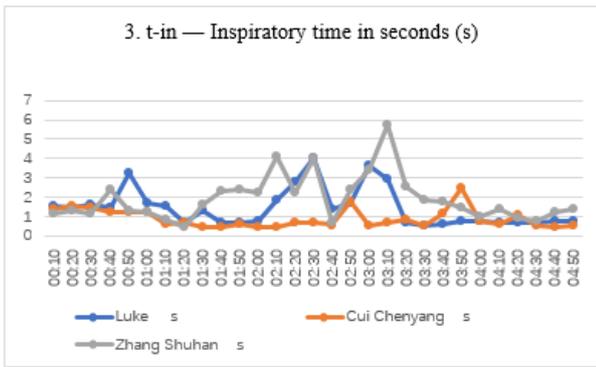


Fig. 3. t-in — Inspiratory time in seconds (s).

The other graphs confirm the observations from the previous figures.

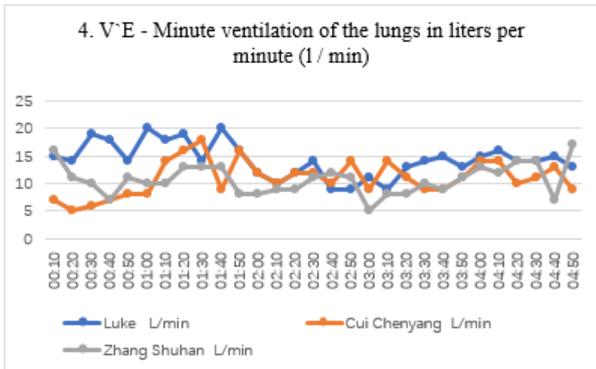


Fig. 4. V'E - Minute ventilation of the lungs in liters per minute (l / min).

Fig. 4 shows that minute ventilation of lungs at the beginning of the experiment (first 01.50 minutes) is much bigger than in further sections for all speakers. It proves that speech is an energy-consuming factor. But the average meanings for the speakers are different. Speaker 1 shows usually bigger figures (>3 l/min) than Speakers 2 and 3, but for the section of words it is the opposite: his ventilation is the smallest (1,93). His curve at this section is the steadiest.

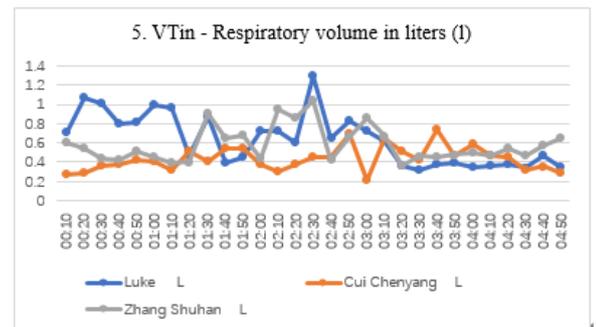


Fig. 5. VTin - Respiratory volume in liters (l).

Respiration volume during the production of phrases changes dramatically for all three speakers, but after 3,20 min it stabilizes. For Speaker 1 it turns to nearly a straight line at the level of volume  $\approx 0.3$  l. Table VI allows to conclude that Speaker 1 uses generally bigger volume of air, except for the section of words, where Speaker 3 shows bigger peaks.

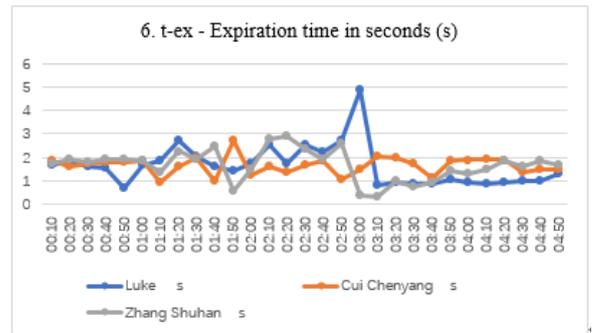


Fig. 6. t-ex - Expiration time in seconds (s).

Fig. 6 shows that at the moment of pronouncing / θ / (see Table V) at the end of the first section (tongue twisters) the native English speaker has a strong exhalation. This is the time of the last sound of the last phrase, after which individual words begin. During the pronunciation of words, the time of out-breaths for all 3 speakers does not exceed 2 sec. During the phrases section the length of out-breath is variable. It depends upon the length of the sentences. From Table V it is clear that the average time of expiration is steadily bigger for Speaker 1, than for the Chinese speakers during all the speaking time and during phrases. As for words, the Englishman shows much shorter exhales than the Chinese speakers.

To determine the conformity of respiratory bursts and dips to certain speech segments, an auditory analysis of the sound track of each speaker was conducted. The results are presented in Tables V and VI. The words that were spoken at a certain time and the key sounds that occur at critical points are shown in the Table.

TABLE V. THE AVERAGE CHARACTERS OF BREATHING FUNCTIONS FOR ALL SUBJECTS

MEAN – TOTAL TIME							
	BF	Ti/tot	t-in	V'E	VTin	t-tex	Period (min)
Speaker 1 Luke	10.36	13.04	0.98	3.13	0.27	0.87	00:10-04:50
Speaker 2 Zhang	5.14	18.38	1.19	2.77	0.18	0.67	00:10-04:50
Speaker 3 Cui	6.56	11.70	0.49	3.19	0.13	0.37	00:10-04:50
MEAN – PHRASES							
	BF	Ti/tot	t-in	V'E	VTin	t-tex	Period (min)
Speaker 1 Luke	7.12	12.80	1.20	3.77	0.25	0.96	01:20-03:00
Speaker 2 Zhang	4.81	23.07	1.43	2.51	0.23	0.92	01:20-03:20
Speaker 3 Cui	6.29	10.69	0.33	2.92	0.13	0.45	01:20-03:30
MEAN – WORD PAIRS							
	BF	Ti/tot	t-in	V'E	VTin	t-tex	Period (min)
Speaker 1 Luke	7.15	11.80	0.72	1.93	0.10	0.06	03:10-04:50
Speaker 2 Zhang	2.65	13.70	0.37	3.02	0.07	0.39	03:30-04:50
Speaker 3 Cui	6.10	13.37	0.66	2.07	0.15	0.31	03:40-04:50

TABLE VI. SOUND RECORDING TIMING OF THE THREE SPEAKERS

Participant №.1		Participant №.2		Participant №.3	
Time (min)	Word/  sound	Time (min)	Word/ sound	Time (min)	Word/ sound
00:10		00:10		00:10	
00:20		00:20		00:20	
00:30		00:30		00:30	
00:40		00:40		00:40	
00:50		00:50		00:50	
01:00		01:00		01:00	Peter /p/-
01:10	Peter /p/-	01:10		01:10	of-pickled
01:20	Peppers /p/-	01:20	Pickled (mistake)	01:20	Peppers -/s/-
01:30	to -tutor	01:30	Peter -/i:-/	01:30	The /ð/-
01:40	Can /k/-	01:40	Who -/u:/	01:40	Car -/(r)/
01:50	Butter /b/-	01:50	He /hi/	01:50	Some -/m/-
02:00	She /ʃ/-	02:00	Butter -/t/-	02:00	batter
02:10	She /ʃ/-	02:10	Will /w/-	02:10	Batter
02:20	“ ”	02:20	Those -/z/	02:20	those-were
02:30	Cook	02:30	She -/i:/	02:30	The
02:40	“ ”	02:40	Can -/n/	02:40	Cream /k/-
02:50	The	02:50	Much -/tʃ/	02:50	A
03:00	Thinking /θ/-	03:00	But -/t/	03:00	Cookies -/k/-
03:10	Peach -/tʃ/	03:10	Thinkers /θ/-	03:10	Stump (mistake)
03:20	Bought /b/-	03:20	Peach /p/-	03:20	Lorry /l/-
03:30	Court -/t/	03:30	Box -/ks/	03:30	“ ”
03:40	Think -/k/	03:40	Sheet -/t/	03:40	Beach /b/-
03:50	Well /w/-	03:50	World /w/-	03:50	Book -/k/
04:00	Batter /b/-	04:00	Pat /p/-	04:00	And /ən/
04:10	Mob /m/	04:10	Stop -/t/-	04:10	Lard /l/-
04:20	Dart /d/-	04:20	Deer -/i:/	04:20	Mop /m/-
04:30	Lord -/d/	04:30	Giddy -/di/	04:30	Lot -/t/
04:40	Mug /m/-	04:40	Sock /s/-	04:40	Mug -/g/
04:50	Jogging -/g/-	04:50	Jogging -/ŋ/	04:50	Jogging -/ŋ/

The soundtrack of each speaker was presented by a dynamic spectrum (oscillography) and synchronized with the ergospirometric data.

Fig. 7-9 show the dependence of physiological parameters on time (the x-axis).

The ordinate axis shows the numerical values of physiological parameters, the dimensions of which are indicated in the boxes on the pictures. All processes in each figure are synchronized in time. The logarithmic scale on the ordinate axis allows us to compare data of the three parameters in arbitrary units.

The time section of the first minute of the test was chosen as the "etalon" of comfortable physiological state. The speakers were asked to make a full inhale and exhale before they start to speak.

We have chosen to analyze four parameters. Two of them are the standard parameters of ergospirometry: 1) «respiratory rate», or «frequency of respiratory movements» further on graphs, and 2) «respiratory volume», that is the average volume of air per minute used by the subject; 3) the ratio of expiratory time to inspiratory time. This parameter is more important than the standard parameters in ergospirometry because in both languages under investigation the speech is connected with expiration (out-breaths), in-breaths are silent, so the expiration time may be associated with the sounding time; 4) amplitude values of the speech sound intensity for each of the subjects on a comparable scale. On top of each of Fig. 7-9 there are the oscillograms, which allow us to associate the periods of speech with the periods of respiratory variables.

The experiment at this stage of the study is aimed at testing the hypothesis that "meaningful sound" ("speaking") in the native language is consistent with the physiological functions of the body, while the violation of the physiologically conditioned native "rhythm" while speaking in a foreign language with different rhythmical rules leads to physiological "stress".

Evidence of that is shown on Fig. 7-9 and proved by the following data.

In the experiment, the three participants begin to speak from 1:00 min. To 1:20 min. Due to the fact that each participant begins to speak at a different time, it is impossible to show data for the three speakers in one graph. An individual analysis is made for each participant in an attempt to find similarities and differences between native speakers of English and Chinese.

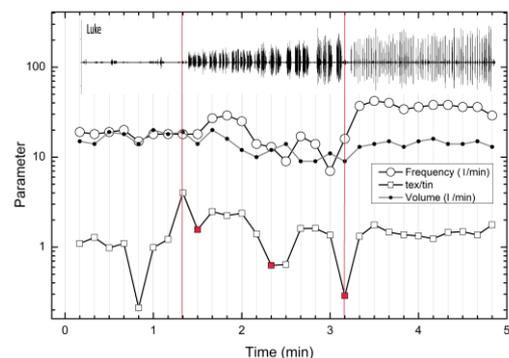


Fig. 7. Speech and physiological parameters of participant No. 1.

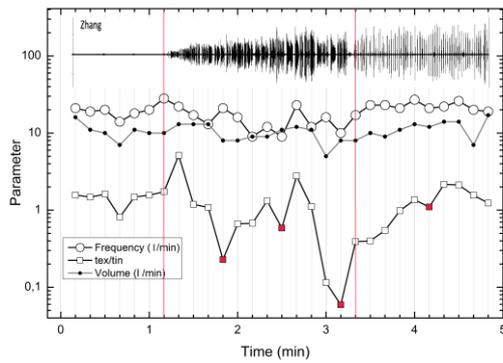


Fig. 8. Speech and physiological parameters of participant No. 2.

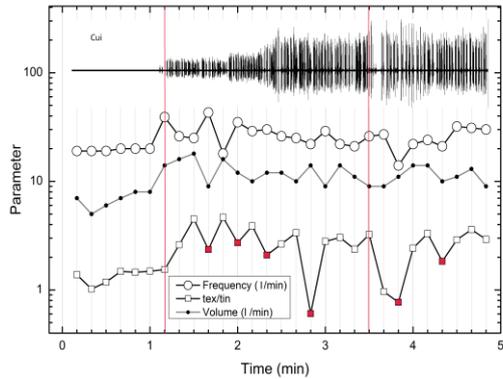


Fig. 9. Speech and physiological parameters of participant No. 3.

There are two red vertical lines in each figure which fix a) the beginning of reading phrases and b) the beginning of reading word pairs. Each low dot colored red fixes the turn of conversion from exhaling to inhaling.

The parameter "frequency of respiration movements" shows that:

Speaker 1 (Englishman) demonstrates an average rate of respiration at the beginning of phrases (about 12 l/min), that is slightly higher than the rate at rest (11 l/min). It drops down a bit towards the end of phrases (8 l/min) and rises high for the pronunciation of separate words (14 l/min) and is kept on. We may observe that the speech does not influence his breathing pattern dramatically. He is not making any special efforts to speak. The Chinese speakers, in contrast, demonstrate more dependence of this parameter on speech. The line is curvier for Speaker 2 and very curvy, especially at the beginning of the experiment, for Speaker 3. There is practically no difference for him between phrases and words.

The amount of air used at rest in all subjects is within the limit of one liter per minute ( $1 \pm 0.2$  liters per minute). For the parameter of respiratory volume, we may use a simple formula of dynamic range:  $(\text{max}-\text{min})/\text{min}$ . The meaning of it for Speaker 1 is  $\sim 0.5$  on average. It is only slightly higher on phrases, which require more effort, than on words. For Speaker 2 the volume for phrases is  $\sim (12-6)/6 = 1.0$ ; for words  $\sim (15-7)/7 = 1.1$ . There is a break of breath at  $t = 1:50$  and  $t = 3:00-3:10$ . For Speaker 3 for phrases the volume is  $\sim (20-8)/8 = 1.25$ ; for words  $\sim (15-9)/9 = 0.33$ . There is a break of breath at  $t = 2:50$  and  $t = 3:40-3:50$ . The data show that for Chinese

speakers the pronunciation of both phrases and words is stressful and they are tired at the end of each of the two sections.

The observations of the "time of expiration/ time of expiration ratio" prove the previous statement. The Englishman made one in-breath during the phrases section with no modulations during words (we do not count preliminary in-breath – outbreath section at the beginning and a sigh before the second section). Speaker 3 (Chinese smoker) demonstrates long and frequent in-breaths (6), notwithstanding, whether he is pronouncing phrases or words. They are equally stressful for him and he is making a lot of effort. Speaker 2 shows average meanings: two in-breaths on phrases and one on words.

## V. INTERPRETATION OF THE DATA OBTAINED

We describe the models of correlations between respiratory functions and speech production process of English and Chinese as follows:

- (1) in the oxygen-reducing process of a native English speaker 100% of oxygen is used, because breathing, pronouncing and thinking are synchronous processes;
- (2) in the oxygen-reducing process of a native Chinese speaker less than 100% of oxygen is used, because the speaker's thinking goes before the speech segment is produced. In the period of thinking the oxygen is continuing to reduce.

That means that for saying a sentence, usable amount of oxygen for a native Chinese speaker is smaller than for a native English speaker. It causes an absolutely different type of phonetic skills and even confusion of brain activity. Finally, it leads to misunderstanding due to the wrong pronunciation of English utterances by Chinese learners.

From the point of view of Chinese speakers, each syllable conveys an idea. Forming thoughts for the Chinese is like a permutation and a combination of numbers in a mathematical formula. Chinese speakers always think first, then say, first calculate what the result will be and finally come to this result. This may be caused by the influence of feudal society for thousand years in Chinese history. We have some famous idioms which are told to children by parents and by teachers at school such as «Be cautious with one's words and actions – 三思而后行», «All one's troubles were caused by his tongue – 祸从口出». Thus, the phonetic structure of the sentence in Chinese as "syllable (word) - pause - syllable (word) - pause". At school the teachers teach native Chinese speakers to add pauses between words. At the same time the word may consists of one, two or three syllables. Making pauses between syllables in connected English speech is impossible.

## VI. CONCLUSIONS

The experiment showed that the speech breathing models of Chinese speakers and English speakers tend to be quite different. Chinese speakers use shorter breath-groups with longer silent spans and frequent respiration cycles. More than that, the breath of a smoker can spoil his/her pronunciation and, consequently, the result of communication. English native

speakers have longer breath groups, consisting of many words, without pauses and they need more oxygen to do it. So Chinese learners have to train their breath and do special exercises for breathing to develop a new model for speaking English. They have to learn to breathe English.

It is too early to make more precise data-based conclusions, because the work is in progress both from the point of view of experimental phonetics and teaching methodology. By this paper we have just demonstrated the fact of the existence of different speech respiratory models for learners and teachers to be aware of. Our hypotheses have to be proved by further investigation. The number of subjects must be increased to give an opportunity for proper statistics. The methodology of the experiment at the next stages may include the measurements of heartbeat and blood pressure as well, which is possible within the ergospirometry method we used. The research of breathing functions in connection with certain sound (strongly aspirated English consonants and tense vowels) is perhaps the closest task, because the experimental language material was compiled with this perspective in mind.

Phonetics is attracting more and more attention in our information age. The urgency of the development of phonetic technology is important in order to improve the accuracy and reality of speech communication.

With the rapid development of computer technology, the importance of writing is gradually decreasing. On the smartphone now almost all chat programs are executed in a voice message format. This helps a person to save his/her time, but another problem may arise: if the speaker in the voice message speaks with incorrect phonetics due to various reasons (hurried breathing, feeling unwell, not paying special attention to the articulation, and so on) the signal will be spoiled and the communication broken.

With the development of 5G technology on the background of the fact that 4G-speed has already been popularized in most territories all over the world, we could transfer both text messages and voice messages to each other easily with our powerful smartphones. If we try to make phonetic analysis with the help of physics (acoustics) and physiology (articulation and breath), we will get different kinds of information about the speaker. On the other hand, we will get information on different aspects of a human being – psychological, physiological, articulatory, linguistic. If we want to successfully create and use these technologies which are related to phonetics, we have to pay much attention to the preparatory phase. That means not only concentrating on developing new apparatuses such as intelligent wearable devices, but also designing a method of teaching people's phonetic skills to match these high-tech devices in the future. To put it simply, we shouldn't just evolve machines and programs, but also ourselves.

#### ACKNOWLEDGMENTS

We would like to show our deepest gratitude to Dr. Aelita Berezina, Northwestern Federal Medical Research Center

named after V.A. Almazov, Dr. Vladislav Bogevoznov, Physics Department, St. Petersburg State University, for their consultancy and assistance, and to the three participants of the experiment.

#### REFERENCES

- [1] L. M. Surhone, M. T. Timpledon, and S. F. Marseken, "Phonetic Complement: Logogram, Morpheme, Phonogram (linguistics), Phoneme, Egyptian Hieroglyphs, Writing in Ancient Egypt, Transliteration of Ancient Egyptian, Egyptian Language, Cuneiform Script," Betascript Publishing, 72p, 2010.
- [2] Yu. K. Kuzmenko, "From phoneme to syllabomorpheme (about the development of Germanic languages in the last 100 years)," St. Petersburg Phonological Conference, 2019 [http://genling.spbu.ru/phonolog/conference/2019\\_Invitation\\_and\\_questionnaire.pdf](http://genling.spbu.ru/phonolog/conference/2019_Invitation_and_questionnaire.pdf).
- [3] N. A. Speshnev, "Phonetics of Chinese language: Introduction of Chinese language. Phonetics and spoken language," SPB.: KARO, pp. 8-14, 2003.
- [4] I. Yu. Pavlovskaya, E. K. Timofeeva, "Introductory Course of English Pronunciation: Phonosemantic approach," St. Petersburg: Hipocrates, 128p, 2009.
- [5] I. Yu. Pavlovskaya, "Experimental comparative study of English and Chinese intonation," Vestnik of Kostroma State University named after Nikolay Nekrasov № 5, ISSN 1998-0817, pp. 166-170, 2016.
- [6] Do Yunsha, "Methodological aspects of forming Russian pronunciation for students of the beginning stage," Ph.D. Russian Pedagogical University named after Alexander Herzen.- St. Petersburg, 242p, 2017.
- [7] Feifei Han, "Pronunciation Problems of Chinese Learners of English," ORTESOL, Volume 30, pp. 26-30, 2013
- [8] K. Ohata, "Phonological differences between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners," Asian EFL Journal, 2004. retrieved on 20th, July, 2012, from [http://www.asian-efl-journal.com/december\\_04\\_KO.php](http://www.asian-efl-journal.com/december_04_KO.php)
- [9] F. Zhang & P. Yin, "A study of pronunciation problems of English learners in China," Asian Social Science, 5, pp. 141-146, 2009.
- [10] I. Yu. Pavlovskaya, "Testing and Teaching English Intonation to Chinese Students," Научная статья TEA SIG IATEFL Newsletter, № 60., pp. 13-16, December, 2016. URL:<http://edition.pagesuite-professional.co.uk/Launch.aspx?EID=58c64536-02a9-4a82-aafe-0656958b85f8>
- [11] L. R. Zinder, "General Phonetics," Moscow, Vishaya Shkola, 312p, 1979.
- [12] S. V. Kodzasov, O. F. Krivnova, "General Phonetics," Moscow, Russian State Humanity University: 592p, 2001.
- [13] G. Funt, "Speech acoustics and phonetics: Selected Writings," Series: Text, Speech and Language Technology (Book 24). Springer, 320p, 2006.
- [14] P. Ladefoged, K. A. Jonson, "Course in Phonetics Sixth Edition," Wadsworth, Cengage Learning, 323p, 2010.
- [15] J. S. Perkell, "Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study," The MIT Press: 106p, 1969.
- [16] Wang Li, "Chinese phonetic history," Beijing.: Publishing, China social sciences press, 639p, 1985.
- [17] Zhang Zhongxing, "The written forms of classical Chinese and modern Chinese" Beijing.: Publishing, Zhonghua book company, 361p, 2012.
- [18] "Encyclopedia of China," Beijing.: Publishing, Encyclopedia of China Publishing House, 1119p, 1993.
- [19] Ergospirometry [Electronic Resource]. – Mode of code: <https://www.spbpmc.ru/departments/directions/ergospirometriya/> (Accessed: 20.05.2019).