

Prolongations as Hesitation Phenomena in Spoken Speech in First and Second Language

N.V. Bogdanova-Beglarian¹[0000-0002-7652-0358], K.D. Zaides^{1,2}[0000-0001-7528-0420],
D.A. Stoika^{1,3}[0000-0002-5254-0135], and X. Sun⁴[0000-0003-2319-7398]

¹ Saint Petersburg State University, St. Petersburg, Russia

² Tel Aviv University, Tel Aviv, Israel

³ Main Administration of the Internal Affairs Ministry of the City of St. Petersburg and the Leningrad Oblast, St. Petersburg, Russia

⁴ Southeast University, Nanjing, China

n.bogdanova@spbu.ru; zaides.kristina@gmail.com;
ruslangdaria@yandex.ru; sunxiaoli_smile@163.com

Abstract. The article considers sound prolongations as one of the most typical hesitation phenomena of spontaneous spoken speech in any language. The material for the analysis was 40 monologues-descriptions, partly taken from the corpus of Russian monologue speech “Balanced Annotated Text Library”, and partly recorded specifically for this study. The monologues were recorded from 4 groups of speakers: 10 Russians and 30 Chinese, of whom 10 described the comic strip in Russian and Chinese, and another 10 only in Chinese. The analysis showed that the appearance of prolongations correlates with the language factor: bilinguals use them more often than monolinguals. More prolongations were found in full-fledged words (not pragmatic markers), but most of them are function words and pronouns. In Russian speech, the percentage of prolongations is also high in pragmatic markers, especially in the word *tot*, which appears in spoken discourse most often as a hesitative boundary marker. Thus, it turned out that even language proficiency does not make the Russian colloquial speech of the Chinese as natural as one of the native speakers. Prolongations predominate in vowels and sonorants; they are frequent at the end of a word and in single-letter words. Often prolongations become a component of an extended hesitation chain that includes other hesitative units (physical hesitation pauses, breaks, repetitions, vocalizations, etc.). The data obtained can be useful in all aspects of applied linguistics: from linguodidactics and linguistic expertise to automatic speech processing systems and the creation of artificial intelligence.

Keywords: Spontaneous Monologue, Monologue-Description, Hesitation Phenomenon, Prolongation, Monolingual, Bilingual, Hesitation Chain.

1 Introduction

Spontaneous speech generation, which accompanies all our everyday spoken communication, inevitably forces speakers to act in conditions of time deficit, when they have to think and speak at the same time, cf.: “in real communication conditions, during

natural spontaneous dialogue [the same applies, presumably, to a spontaneous monologue as well.—Authors] the production of a text (utterance) essentially occurs in ‘extreme conditions’—with a lack of time and the absence of the opportunity to carefully plan the strategy” [13: 162]. Such “extreme conditions” pose many problems for the speaker related to the formulation of a thought, the selection of words that would best express this thought, and the correction of inevitable inaccuracies, stipulations, and even mistakes. This complex speech and thought activity often leads to a “gap between thoughts and language”, which is expressed “in speech hesitations, false starts, and reformulations that abound in everyday speech” [5: 64].

Hesitation phenomena, or *signs of spontaneity*, that accompany speech production are characteristic of spontaneous speech in any language, and the speech of any speakers, regardless of their social and/or psychological characteristics. One of these hesitation phenomena are *prolongations of sounds* [1, 9, 18], both vowels (*vo-o-ot*, *nu-u-u* (‘well’)) and consonants (*s-skazal* (‘said’), *khodim-m-m* (‘we walk’)). Such prolongations on monosyllabic (and one-phonemic) words (*a-a* (‘but’), *i-i* (‘and’), *v-v* (‘in’)) should be distinguished from insertions of non-speech sounds (*vocalizations*). The criterion for their differentiation can be the presence or absence of grammatical meaning of these units within the text.

This kind of material is of interest in various aspects. In this paper, it is considered from the point of view of general linguistics (speech in the native and non-native language, the speech of monolinguals or bilinguals, as well as the characteristics of prolongations), but it can also be considered in terms of sociolinguistics (men and women speech, the influence of the level of Chinese proficiency in Russian on their speech, including the appearance of prolongations as a hesitating phenomenon) and psycholinguistics (texts produced by speakers with different psychological characteristics). This approach allows applying the study results in linguodidactics, for the creation of speech portraits of both individual speakers and various social groups. The results can be useful for linguistic expertise and linguocriminalistics, as well as for automatic natural language processing systems and the creation of artificial intelligence.

2 Material and Methodology

To obtain the material, 4 groups of texts were recorded from:

- 1) Russian monolinguals speaking Russian as their native language (R)—10 texts;
- 2) Chinese monolinguals who speak Chinese as their native language and do not know Russian (Ch-Ch)—10 texts;
- 3) bilingual Chinese speaking Chinese as their native language and know Russian (Ch-Ch1)—10 texts;
- 4) the same bilingual Chinese speaking Russian as their second language (Ch-R)—10 texts.

Calling speakers *monolinguals* or *bilinguals*, we take into account only their attitudes towards Russian and Chinese. Moreover, in this paper, we use the term ‘bilingual’ in its wider meaning to describe a person who can speak two languages or frequently uses them both in a relevant language environment while studying one of them, e.g.,

Chinese and Russian. Thus, Russian monolinguals speak Russian as their native language and do not know Chinese; Chinese monolinguals speak Chinese as their native language and do not know Russian. Accordingly, Chinese bilinguals, in addition to their native Chinese, also speak Russian (they study it). At the same time, both Russians and Chinese may be proficient in some other foreign languages (third, fourth, etc.), which in this case is not taken into account in any way.

All speakers are from the same age group: they are students and graduate students 20-29 years old. The groups of speakers are balanced by gender (5 men and 5 women in each group); the group of Russian monolinguals is balanced in terms of psychotype (5 extroverts and 5 introverts); for the group of Chinese, it was not possible to achieve such a balance. For the Chinese speaking Russian, the level of Russian as a foreign language proficiency (TORFL) is also taken into account (5 speakers with B2 level and 5 with C1 level). To define the psychotype, the speakers were psychologically tested (test by H. Eysenck, which all speakers passed in their native language). The level of TORFL for the Chinese was determined at the place of their study: a lower level B2 (TORFL-2) or higher C1 (TORFL-3), according to the Russian State Testing System for Foreign Citizens. Some of the texts were borrowed from the corpus of monologues “Balanced Annotated Text Library” [2], other texts, primarily in Chinese, were recorded specifically for this study.

The initial stimulus for the production of the monologue in all cases was the image with the plot— H. Bidstrup’s comic strip “Hair loss treatment”. The speakers were asked to describe this comic while looking at it, i.e., to compose a story based on the picture. All speakers are familiar with this genre of spoken monologue: for native speakers, this was an exercise for speech development in elementary school; the Chinese often performed such exercises during Russian language lessons at the university.

All recorded material was transcribed based on auditory analysis and annotated, including prolongations of vowels and consonants. As a result of the annotation, a large pilot subcorpus was created, where two types of data were added: (1) metadata of speakers (native language, group, gender, psychotype, and level of TORFL for the Chinese) and (2) characteristics of the prolongations (P). In this case, the following points were taken into account: a content or a functional word contains the P; a part of speech; a full-fledged word or a pragmatic marker (PM) contains the P (for more on pragmatic markers, see: [3]); a position of the P in the word: at the beginning, within the word, at the end; a separate group consisted of extensions of a single-letter word; the P as an isolated hesitating phenomenon or as part of a hesitating chain; the length of such a chain; the P of a vowel or a consonant; the duration of the P in ms; the proportion of Ps depending on the volume in tokens/words and on the duration of the monologue.

2.1 Prolongations of Vowels and Consonants in Russian

When determining the duration of a hesitation prolongation of a particular sound, the following factors were taken into account:

- change (difference) in the amplitude of the oscillation of the vocal folds, with fixation of the appearance of two-peak vowels (oscillographic analysis);

- change in sound quality; for example, when a vowel is represented by a non-basic or reduced allophone, the main allophone may become clearer (*'ne'* [n'éc]) during lengthening, and at the absolute end of the word the vowel may turn into a reduced one (*'na'* [náъ]) (spectral analysis);
- the average duration of sound in speech.

The average duration of vowels and consonants in Russian speech was determined based on the experimental data from some of the main phonetic studies [4, 24, 8, 19, 7, 11, 17], as well as on the extended expert experience of the authors in the analysis of spontaneous speech. Table 1 shows the values of the average duration for those Russian sounds, the hesitation lengthening of which was found in the material.

Table 1. The average duration of Russian sounds with prolongation in the study material

Type of sound	Sound	Average duration, ms
Vowels	/i ы u e o a/	130
Occlusive plosive sonorants	/m n n'/	100
Obstruent fricative consonants	/s s' z f/	100
Obstruent occlusive plosive consonants	/k t t'/	80

The duration of sounds in speech is very variable and depends on many factors: the type of a sound (for example, occlusive consonants are shorter than fricative ones; closed vowels are shorter than mid and open vowels), the position of sound in the word and the place of stress, the number of sounds in a word, the type of a syllable (closed or open), the quality of an adjacent consonant, the position of a phonetic word in a phrase, the length of a phrase itself and the communicative type of a phrase [21: 184], the speech rate. In [7], for speech synthesis, the following average values of the duration of sounds are indicated: for vowels—180 ms, for consonants—95 ms; the range of vowel duration is 30-300 ms [7: 216]. In other papers, the range of vowel duration is 60-200 ms [4, 24, 19].

Considering the above experimental data, to describe Russian hesitation prolongations, in this paper it was decided to analyze only those sounds whose duration exceeds 180-200 ms (for vowels) and 95-100 ms (for consonants).

2.2 Prolongations of Vowels and Consonants in Chinese

Standard Chinese (Mandarin) phonology is based on the Beijing dialect. A Mandarin syllable consists of three components: an initial (consonant), a final (vowel), and a tone [10]. There are 22 initials in Chinese and three types of final sounds: simple finals, compound finals, and nasal finals. The 6 simple finals *a*, *o*, *e*, *i*, *u*, *ü* are the most fundamental elements of Chinese vowels. There are 13 compound finals and 14 nasal finals in Chinese. In addition, two semi-vowels *y* and *w* can be distinguished in Chinese. The tone is the pitch contour of the syllable. Mandarin has four contour tones and a neutral tone used on weak syllables. All syllables in Chinese must have a final, but many syllables exist with no initial, for example, 爱 *ai* and 五 *wu*.

Each Chinese character is spoken as one syllable. Chinese words are made up of one, two, or more characters (syllables). In Chinese, there are whole syllables that are not divided into an initial consonant sound and a final. There are 16 in total.

The prolongation is presented in Chinese native speech in both content words and function words, but “due to the particularity of Chinese syllables (most Chinese syllables are open syllables and closed syllables ending in nasal consonants), only word-ending vowels and nasal consonants appear prolongation” [22: 119]. There are both single vowels and compound vowels in the vowels where prolonged sounds occur. Nasal consonants that tend to be elongated are *n* and *ng* [id.], and the results of this study also confirm this.

The data obtained in this study shows that the prolongations in Chinese may be related to the tone. The soft tone is the easiest to prolong, followed by the first and fourth tones, and the least elongated sounds are the second and third tones. In addition to being influenced by the phonemic rules, the frequency of prolongations is also influenced by personal characteristics, such as expressiveness and psychological factors.

Prolongation in Chinese is the duration of a sound that usually exceeds the average duration of the same sound. The average duration of Chinese vowels is presented in Table 2.

Table 2. The average duration of Chinese vowels

Vowel	Average duration, ms
Simple final	104-118 ms
Compound final	154-158 ms
Nasal final	160 ms

Feng Long’s research data on the sound duration of Beijing dialect shows that the average duration of the simple final in the front syllable is 104 ms, and the average duration of the simple final in the following syllable is 118 ms, and the simple final in the back syllable is 14 ms longer than the simple final in the front syllable [6]. This study did not directly count the average duration of compound finals, but we can infer that the average duration of compound finals is about 40 ms longer than that of simple finals by comparing the average duration of syllables with simple finals and syllables with compound finals. Furthermore, it can be inferred that the average duration of the compound finals in the front syllable is about 154 ms, and the average duration of the compound finals in the back syllable is 158 ms. Because the prolongation in this study only occurs in the following syllables, here we only need to consider the average duration of the finals in the back syllables. Since no literature was found about the average duration of nasal finals, we measured and calculated the average duration of nasal finals in the material. The results show that the average length of Chinese nasal finals is 160 ms.

Qi Shiqian and Zhang Jialu analyzed 7 male and 6 female speech excerpts, and obtained the average length of the 22 initials in Mandarin [16], as shown in Table 3.

Table 3. The average duration of Chinese initials

Consonant	Aver. dur. (ms)	Consonant	Aver. dur. (ms)	Consonant	Aver. dur. (ms)
b	14.2	g	23.2	sh	138.6
p	98.2	k	94.4	r	56.3
m	71.3	h	106.6	z	53.9
f	106.6	j	55	c	126.9
d	9.6	q	143.7	s	140.3
t	106.3	x	129.1	ng	134.2
l	63.9	zh	21.1		
n	61.5	ch	108.9		

The pronunciation of the whole syllable will remain the same as the consonant even after adding a vowel behind the consonant (*zhi, chi, shi, ri, zi, ci, si*) or as the vowel even after adding a semi-vowel *y* behind the vowel (*ye, yi, yin, ying, wu, yu, yue, yun, yuan*). Therefore, the overall sound length of the whole syllables can refer to the average length of the actual pronunciation of initials or finals.

2.3 Problems of Analysis and Decisions Made

In the process of analyzing sound prolongations by instrumental methods (oscillographic and spectral analysis), some disagreements arose with the results of the initial auditory analysis carried out by the transcribers: some of the Ps identified by ear were not confirmed instrumentally. Thus, the following inconsistencies were identified (for the signs of the orthographic representation of corpus materials, see: [23]).

1. When a vowel appeared at the junction with another vowel (*u nego volosy rastut prosto s neimovernoj skorost' yu i: v kontse kontsov *B on ves' stanovitsya () takoi () mokhnatyi* ('his hair grow in fact with incredible speed and after all *B he became () such () furry'); *chelovek pokupayet v magazine: eliksir dlya rosta volos* ('a man buys in a shop a hair loss treatment')), auditory analysis determined the sound prolongation due to the merging of neighboring vowels, although instrumental analysis showed that there was no prolongation in these cases, the duration of the vowels falls within the average sound range: the vowel of the conjunction *i* ('and') [i]—137 ms, the final vowel in the word *magazine* ('shop') [i]—119 ms.

2. In two cases, a vowel prolongation was noted after long fortis fricative [s], [s']: *volosy*: ('hair'), *eliksi:r* ('treatment'). The duration of the fricatives in these words was 334 and 216 ms, respectively, while the duration of the vowels following them fell into the average values.

3. Quite often, in the course of auditory analysis, a prolongation was noted in the words *vot* and *nu* ('well'), which, as a rule, appear in speech not as particles, but as pragmatic markers of one or another functional type [3], which could affect the subjective perception of duration, not supported by instrumental analysis, cf.:

(1) *to yest' snachala podumal chto eto kapli // **vo:t*** (final or navigational boundary marker)

so he first thought that these were drops // well;

(2) *a tut on kak by () z:alez e-e () nu: e nu pri etom volosy u nego vyrosli yeshchyо bol'she* (polyfunctional boundary and hesitation marker)

and here he is like () cli:mbed e-e () well e well meanwhile his hair grew even more.

4. In several cases, the speaker's hesitation was mistaken for the sound prolongation in the course of auditory analysis, cf.:

(3) *ne m:og* [n`im/mók] ('he couldn't') (a slash denotes a hesitation pause),

(4) *neo:bkhodimo* [n`eb/a`aphad`íma] ('is necessary') (from the speech of a Chinese speaker).

In all these cases, the data obtained after the instrumental analysis and correction were used for further investigation.

In the monologue of the speaker 8Ch-R, prolongations of vowels were annotated, which were not hesitant, but emphatic. It is interesting to note that this phenomenon occurred in the speech of only one informant, a 26-year-old Chinese woman, who very emotionally described the comic strip offered to her. Emotionality is manifested in a wide range of changes in the basic tone, vocabulary choice, and phonetic features of the implementation of particular sounds, cf.:

(5) *Gospodi / kakoi krasota: !*

God / what a beauty !

(6) *na sleduyushchii den' / on yeshchyо: rastyot ! dlinne:ye ! u menya uzhe sre:dnii volOs*

next day / it grows more ! longer ! I have now middle hair.

This kind of elongation of a sound was also not considered hesitation prolongation.

2.4 Phonetic Properties of Prolongations

Before proceeding to the description of the results obtained during the study, it is interesting to note several phonetic features of prolongations in the Russian speech of Russian and Chinese speakers.

Thus, Russians often have prolongations with a change in sound quality, the transition of a vowel into vocalizations *m-m*, *e-e*: *dlya: + m-m* [dl`á:m:] ('for *m-m*'), *no: + m-m* [nó:m:] ('but *m-m*'), *on: + e-e* [ón:e:] ('he *e-e*'), *i: + e-e* [í:e:] ('and *e-e*'), *no: + e-e* [nó:e:] ('but *e-e*'). In prolongations produced by the Chinese speakers, the retention of the sound quality or its transition to a reduced one is more often observed.

Let us illustrate such an elongation using the example of the two longest prolongations: in Russian speech, it is the prolongation in the conjunction *i* ('and') with the duration of 1104 ms (see Fig. 1), in Chinese speech, it is the prolongation of the final vowel [a] in the word *zavtra* ('tomorrow') with a duration of 1314 ms (see Fig. 2).

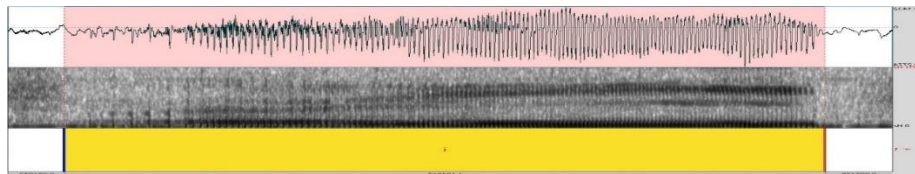


Fig. 1. Oscillogram and spectrogram of the conjunction 'i' ('and') from the speech of a Russian monolingual

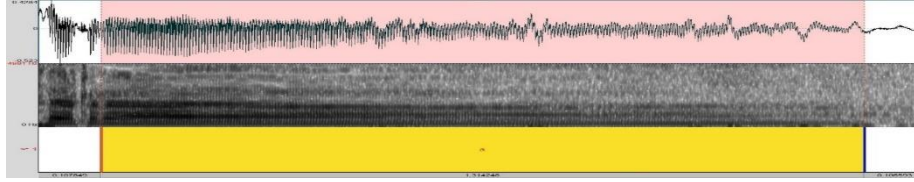


Fig. 2. Oscillogram and spectrogram of the final vowel [a] in the word ‘zavtra’ (‘tomorrow’) from the Russian speech of a bilingual Chinese

The spectrogram of the vowel [i] shows a change in its quality—the frequencies F1 and F2 change: F1 rises from 222 Hz to 510 Hz, and F2 drops from 2123 Hz to 1614 Hz. Thus, [i] becomes more open and moved back, i.e., approaching the sound [ɛ]. The spectrum of the vowel [a] is more stable: being at the absolute end of the word, the vowel becomes reduced (there is a slight decrease in F2).

Another feature noted in the Russian speech of bilingual Chinese is the hesitative lengthening of final plosives:

- *otrezat'*: (‘to cut off’): [tʰ] 351 ms—strengthening of the affrication of the soft [tʰ] with the transition to the affricate [tsʰ];
- *tak*: (‘so’): [k] 428 ms—h-sounded fricativization [kʰ];
- *budet*: (‘will be’): [t] 183 ms—long-term plosive and enhanced explosion phase;
- *raschosyvat'* (‘to comb’): [tʰ] 611 ms—strengthening of the affrication of the soft [tʰ] with the transition to the affricate [tsʰ] with a vowel overtone after the consonant.

All these features are taken into account in the further analysis of prolongations.

3 Results

3.1 Quantitative Overview

The monologues-descriptions that have become the object of analysis differ primarily in their volume in words/tokens: the longest monologues were in Russian, both for Russian monolinguals (2216 tokens) and for bilingual Chinese (1984 tokens); significantly (2-3 times) shorter are monologues in Chinese: 974 in the Ch-Ch1 group (bilingual Chinese) and 694 in the Ch-Ch group (monolingual Chinese). In other words, in their native language, within the same type of text, speakers of different languages produced monologues of fundamentally different volumes: Russian monolinguals produced very large monologues, and monolingual Chinese—more than 3 times shorter. The reason for such a significant difference should be determined, apparently, by the structural features of the languages used, since the bilingual Chinese also produced larger texts in Russian than Chinese. However, the monologue length can be also influenced by various factors, such as the speaker’s gender, age, speech rate, etc., which are not considered in this study.

Table 4 shows the general quantitative data on prolongations in the monologues of all groups of speakers.

Table 4. General quantitative data on prolongations in monologues of all speakers

Group	R	Ch-R	Ch-Ch1	Ch-Ch
Total	69	171	49	7
% of all prolongations	23.3	57.8	16.6	2.4
Proportion of words (per 100 words)	3.1	8.6	5.0	1.0
Proportion of duration (per 100 s)	6.9	9.7	4.5	1.2
Mean (per text)	6.9	17.1	4.8	0.8
Median (per text)	7.0	16.0	3.5	1.0
Standard deviation	4.5	9.4	3.0	0.9

Table 4 shows that bilingual Chinese (Ch-R) use Ps as a hesitation phenomenon most often: their monologues-descriptions in non-native Russian contain more than half (57.8%) of all identified P, the maximum is reached also in all other respects: 8,6 prolongations per 100 words, average 17.1 (± 9.4) and median 16.0 per one monologue.

Chinese monolinguals (Ch-Ch), who speak their native language, use the least number of prolongations. Interestingly, Russian monolinguals use almost a quarter (23.3%) of all identified prolongations. It can be assumed, therefore, that the factor of spontaneity “outweighs” the factor of the native or non-native language of speech production since the native speakers of two languages use different number of prolongations in their spontaneous monologues.

The Kruskal–Wallis test [12] was used to compare medians, which showed that all median values differ significantly: statistic=25.554183644461276; pvalue=1.18234602351863e-05.

3.2 Frequency List of Words with Prolongations in Russian and Chinese

Tables 5 and 6 show the lists of the most frequently elongated words from the analyzed monologues for both languages.

Table 5. TOP-5 elongated words in Russian speech

#	Word	Total	Number R	% of all R	Number Ch-R	% of all Ch-R
1	i (‘and’)	91	24	34.8	67	39.2
2	on/nego (‘he/his’)	22	5	7.2	17	9.9
3	volosy (‘hair’)	9	2	2.9	7	4.1
4	na (‘on’)	8	2	2.9	6	3.5
5	vot (‘well’)	8	8	11.6	0	0

It can be seen that in the Russian material (see Table 5) the maximum of prolongations (which mark, as it was said, speech disfluencies) can be found in the conjunction *i* (‘and’) (the absolute maximum: more than a third of all prolongations in both cases, i.e., in a speech in a native and a non-native language), as well as in forms of the personal pronoun *on/nego* (‘he/his’), replacing the name of the main character of the comic

strip in monologues-descriptions. The appearance in this top of the only content (not a pronoun) word *volosy* ('hair') is explained by the topic of the comic and, consequently, the main topic of all monologues-descriptions.

Also noteworthy is a high percentage of prolongations in the word *vot* ('well') in the speech of Russian monolinguals—11.6%, which exceeds the percentage in *on/nego* ('he/his') (7.2%). This can be easily explained by the fact that *vot* is high-frequent in Russian speech in the role of primarily a pragmatic marker, including the boundary-hesitative [15: 91-109], which is within the list of hesitation phenomena. Interestingly, in the Russian speech of Chinese bilinguals, there are no prolongations in this word, i.e., the Chinese, even those who speak Russian well, do not fully master the arsenal of hesitation phenomena that native speakers have. It can be assumed that the ability to use such "signs of colloquialism" can serve as one of the criteria for good (natural) colloquial Russian speech for foreigners.

Table 6. TOP-5 and TOP-2 elongated words in Chinese speech (*These words are untranslatable particles)

Ch-Ch1				Ch-Ch			
#	Word	Number	% of all	#	Word	Number	% of all
		Ch-Ch1	Ch-Ch1			Ch-Ch	Ch-Ch
1	yi ge ('one')	9	18.4	1	de*	3	42.8
2	de*	7	14.3	2	you yu ('because')	1	14.3
3	zhe ge ('this')	5	10.2				
4	ne*	4	8.3	2	qu ('go')	1	14.3
5	dan shi ('but')	3	6.2	2	le*	1	14.3
5	ran hou ('and then')	3	6.2	2	dao ('arrive')	1	14.3
5	ta ('he')	3	6.2				

Table 6 shows that in Chinese speech, too, most of the hesitation prolongations (i.e., the speaker's disfluencies) are not in content words, but in auxiliary particles, discourse or pragmatic markers. Thus, the auxiliary particle *de* (的/得/地) can be used after definition and indicate a qualitative, possessive attribute (*tu tou de: nan zi*); is used as a suffix for a state verb (*gao xing de: mo shang qu*); is used as a suffix of a verb that accompanies an adverbial mode of action (*zhang de: yue lai yue kuai*). *Yi ge* ('one') is a counter word, after which a noun is usually searched, and *zhe ge* ('this') is a demonstrative pronoun or a hesitation marker, an analog of the Russian markers *eto* and *eto samoye* [15: 432-458] (about the method of searching for analogs of Russian verbal hesitative markers using parallel text analysis, see [20]).

3.3 Position of Prolongation

Table 7 shows data on the prolongation position within a word.

Table 7. Number of prolongations in different positions in a word (abs./relat. (%))

Position in word	R	Ch-R	Ch-Ch1	Ch-Ch	Total
Single-letter words	25/36.2	73/42.7	0/0	0/0	98/33.1
Word beginning	4/5.8	6/3.5	0/0	0/0	10/3.4
Within a word	13/18.8	19/11.1	0/0	0/0	32/10.8
Word end	27/39.1	73/42.7	49/100.0	7/100.0	156/52.7

Table 7 suggests that a higher number of prolongations as hesitation phenomena occurs at the end of a word (52.7%)—in the speech of all groups of speakers. At the same time, in the Chinese speech of all Chinese speakers (Ch-Ch1 and Ch-Ch), prolongations appear only in this position.

The second most common type is the prolongations of single-letter words (33.1%), and in the Russian speech of bilingual Chinese (Ch-R), their number coincides with such in the position at the end of the word (42.7% each).

Prolongations occur least often at the beginning of a word (3.4%). Apparently, the process of hesitation begins towards the end of the pronounced word, when the speaker begins to think about the next word or portion of the text.

Let us further consider the results of quantitative and qualitative analysis of particular sounds prolongations in Russian and Chinese spontaneous speech.

3.4 Prolongations of Vowels and Consonants

In general, the number of vowel Ps in a speech in both languages predominates consonant Ps: 198 (83%) vs. 42 (17%) in Russian and 48 (86%) vs. 8 (14%) in Chinese.

Data on prolongations of particular sounds are presented in Tables 8-11.

Table 8. Number of vowel prolongations in Russian speech of Russians and Chinese

#	Vowel	% of all R	Vowel	% of all Ch-R
1	i	45,6	i	42,1
2	o	19,3	a	15,8
3	a	8,8	o	11,6
4	ы	5,3	u	4,3
5	e	3,5	ы	3,6
6	u	1,7	e	3,1

Most often in Russian speech, speakers of both groups (R and Ch-R) elongate the sound [i], the range of the prolongation duration of this vowel for Russian monolinguals is 205-1104 ms, for bilingual Chinese—212-919 ms. The sound [o] (208-625 ms) is the second most frequent sound with prolongations among Russian speakers, and the sound [a] (218-1314 ms)—among the Chinese. The Chinese speech is characterized by prolongations of the final [a] in content words (*prodavtsa*: ('seller'), *segodnya*: ('today'), *rasteniya*: ('plants')), while Russians elongate [a] more often in functional words (*na*:

(*'on'*), *dlya:* (*'for'*), *ta:k* (*'thus'*)). As noted above, in the speech of bilingual Chinese, the sound [a] in the word *zavtra* (*'tomorrow'*) is the most often elongated (see Fig. 2).

Table 9. Number of consonant prolongations in Russian speech of Russians and Chinese

#	Consonant	% of all R	% of all Ch-R
1	n	10,5	7,9
2	n'	3,5	0,6
3	z	1,7	0
4	s	0	4,9
5	s'	0	1,2
6	m	0	1,8
7	t	0	0,6
8	t'	0	1,2
9	k	0	0,6
10	f	0	0,6

Of the consonants in Russian speech, speakers of both groups most often elongate [n], the range of the duration of this prolongation for Russian monolinguals is 180-412 ms, for bilingual Chinese—211-619 ms. Frequently, such a prolongation is placed in the pronoun *on:* (*'he'*). The remaining consonants are elongated mainly in the Russian speech of the Chinese and mainly in content words (*volos:y* (*'hair'*), *s:mog* (*'could'*), *s:nom* (*'by dream'*), *vecherom:* (*'in the evening'*), *budet:* (*'will'*), etc.), which is not typical for Russian speech.

Table 10. Number of vowel prolongations in Chinese speech of Chinese monolinguals and bilinguals

#	Vowel	% of all Ch-Ch1	#	Vowel	% of all Ch-Ch
1	e	63,4	1	e	57,1
2	a	10,0	2	u	28,6
2	i	10,0	3	ao	14,3
2	ou	10,0			
3	iu	1,2			
3	u	1,2			
3	ü	1,2			

Most often, in Chinese speech, speakers from both groups (Ch-Ch and Ch-Ch1) elongate the vowel [e], the range of the duration of this prolongation is 263-586 ms for Chinese monolinguals, and 169-789 ms for Chinese bilinguals. For Chinese monolinguals, the prolongation is most often found at the end of the auxiliary particle *de* (3 out of 7 cases; 42.9%). Chinese bilinguals' speech also is characterized by prolongations at the end of the auxiliary particle *de* (12 out of 41 cases; 29.2%), the counter word *yi*

ge ('one') (7 uses; 17.1%), the pragmatic marker or demonstrative pronoun *zhe ge* ('this') (5 ; 12.2%). In second place by frequency among Chinese bilinguals, there are three elongated vowels [a], [i], [ou], and their prolongation duration range is 257-445 ms, 433-655 ms, and 188-385 ms, respectively. Other vowels are rarely elongated.

Table 11. Number of consonant prolongations in Chinese speech of Chinese monolinguals and bilinguals

#	Consonant	% of all Ch-Ch1
1	sh	62,5
2	z	12,5
2	ng	12,5
2	n	12,5

It should be noted that in the Chinese speech of Chinese monolinguals there are no consonant prolongations. Of the consonants in the Chinese speech of bilingual Chinese, *shi* is most often elongated in conjunctions *dan shi* ('but') and *ke shi* ('however'), which is an inseparable syllable in Chinese and is pronounced the same as the consonant sound [sh]. The range of duration of this prolongation in Chinese bilinguals' speech is 345-869 ms. In contrast to the result of the study by Teng Hai and Li Yifang [22], the prolongations of the nasal consonants [ng] and [n] in the present study are unique cases and take second place, behind the consonant [sh].

Table 12 summarizes the data by sound types.

Table 12. Number of consonant prolongations of different types of sounds

Sound types	R	Ch-R	Ch-Ch1	Ch-Ch	Total
Fricative obstruent consonants	1/1.4	15/8.8	6/12.2	0/0	22/7.4
Sonorant	10/14.5	16/9.4	2/4.1	0/0	28/9.5
Vowels	58/84.1	140/81.9	41/83.7	7/100.0	246/83.1

Table 12 confirms the data presented at the beginning of this section that most prolongations were found in vowels (83.1%), the difference between consonants of different types is insignificant.

3.5 Isolated Prolongations and Prolongations in Hesitation Chain

The standalone prolongations are two times rarer (30 %) than the prolongations within a hesitation chain (70%). It can be explained by the fact that the single prolongation does not usually provide enough time for word search and speech planning. However, the length of a hesitation chain produced by the speakers can be influenced by their language knowledge: the second language speakers use longer hesitation chains than the first language speakers. A hesitation chain can consist of different items: hesitation pauses (filled and silent), word repetitions, word breaks, sighs, and other types of verbalized speech disfluencies. The average length of the chain among the first language

speakers is two items, while the average length of the chain among the second language speakers is four items, and the difference is statistically significant, according to the Mann–Whitney U test (statistic = 3966.0, pvalue = 5.01748542859133e-05) [14].

The most common sequence in a hesitation chain that includes Ps is the prolongation and one or two hesitation pauses (36%). Among them, the sequences with a short unfilled hesitation pause are taken the first place (22%). The ‘prolongation + long unfilled hesitation pause’ is the second most common type (9%). At last, the ‘prolongation + filled hesitation pause’ are 3%. 2% of cases are the ‘prolongation + filled hesitation pause + short unfilled hesitation pause’. Other types of hesitation chains form 64%.

Within the language groups, the ratio remains relatively similar.

Note that almost all cases with the prolongation and short hesitation pause are such that the pause follows the P. Thus, the prolongation signals that there is a disfluency that is happening now, and a pause is needed to find a way to resolve this disfluency.

The length of a hesitation chain lies in the range 2-13 units. Figure 3 shows the number of hesitation chains with prolongations among all the speakers.

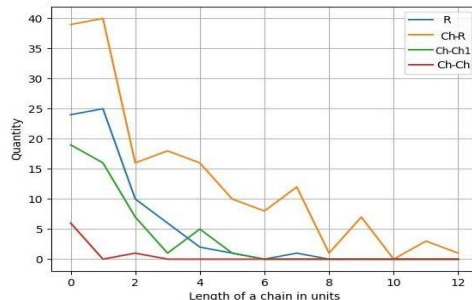


Fig. 3. Number of hesitation chains with prolongations in Russian and Chinese speech

The trend in usages of hesitation chains with prolongations is that the longer the chain is, the rarer it appears in spoken speech. 39% of all hesitation chains with prolongations are sequences of two elements.

The longest hesitation chain in the material consists of 13 elements, including filled and unfilled long and short hesitation pauses, a break, and a repetition, as well as there are three cases of a 12-elements chain:

- (7) *ran'she o... () odin muzhchina / e-y () u nego () u nego net () volos:y u... volosOv*
// n-n (...) e-y eto: (...) ochen' (...) y-yn (...) g... e-y e-y (...) yemu ochen'
grustno (1Ch-R, w., 23)
before o... one man / e-y () he () he doesn't have () hair h... hair // n-n (...) e-y
this (...) very (...) y-yn (...) s... e-y e-y (...) he is very sad.

However, all such cases with the longest hesitation chains belong to the speaker 1Ch-R, which has II level of Russian as a foreign language. So, the length of the chain can be potentially affected by the level of Russian as a foreign language knowledge in this way that the speakers with lower Russian level use longer hesitation chains. We tested this hypothesis, and it was confirmed: the average length of the chain among the II-

level speakers was 4.6 units, the average length of the chain among the III-level speakers was 3.2. The difference between the means was statistically significant, according to the Mann–Whitney U test (statistic = 4507.0, pvalue = 0.006).

4 Conclusion

The study allowed to draw several preliminary conclusions on prolongations based on the analysis of the material.

The monologue-description has various signs of spontaneity, among which a significant place in speech of all groups of speakers is taken by prolongations of sounds, both vowels and consonants.

Most of the prolongations were found in the Russian speech of bilingual Chinese (Ch-R)—more than half of the total number of prolongations. The least number of prolongations was found in the monologues of monolingual Chinese (Ch-Ch). Quite many prolongations were also observed in Russian speech—almost a quarter of the total. Apparently, the factor of spontaneity of speech in this respect is much stronger than the factor of the native language.

In the Russian speech of both Russians and Chinese, the maximum prolongations in the material fall on the conjunction *i* ('and') and forms of the pronoun *on/nego* ('he/his'); in Chinese speech—on the structural particle *de* ('of'), the counter word *yi ge* ('one') and the demonstrative pronoun/discourse-pragmatic marker *zhe ge* ('this') (analogs of the Russian PM *eto* and *eto samoye*).

Of the content parts of speech, only the word *volosy* ('hair') is in the TOP-5 most frequent Russian words with prolongations, which is easily explained by the topic of the comic strip being described.

In general, in both languages, there were more prolongations in full-fledged words than in markers, although in the Russian speech of Russians, the percentage is also high for PM, especially in the word *vot* ('well').

Prolongations at the end of a word and in single-letter words clearly predominate.

There are much more prolongations of vowels than of consonants, and of the latter—of sonorants. In the Russian speech of Russians, vowel extensions [i] and [o] predominate, in the Russian speech of the Chinese—[i] and [a]. Of the Russian consonants, this type of hesitation most often occurs in the sound [n] in the speech of Russians and the sounds [n] and [s] in the speech of the Chinese. In the Chinese speech of Chinese bilinguals and monolinguals, the stretching of the vowel [e] predominates. In the Chinese speech of Chinese monolinguals, there is no stretching of consonants. Of the consonants in the Chinese speech of Chinese bilinguals, [sh] is most often elongated.

Prolongations can more often be found within the hesitation chain, and the second language speakers (Chinese speakers who know Russian), especially those with a lower level of language proficiency, use longer hesitation chains than the first language speakers. 39% of all hesitation chains with prolongations are sequences of two elements, one of them, in addition to the prolongation, tends to be an unfilled hesitation pause.

The study allowed us to obtain a lot of specific data on prolongations as a type of hesitation phenomena that is characteristic of the speech of any speaker and in any

language. This is a universal phonetic and discourse feature of spoken speech, which must be taken into account in all applied aspects of linguistics: from linguodidactics and linguistic expertise to automatic speech processing systems and the creation of AI.

Acknowledgment

The research was conducted was partially supported by St. Petersburg State University, project #94033528 “Modeling of Russian Megalopolis Citizens’ Communicative Behavior in Social, Speech and Pragmatic Aspects Using Artificial Intelligence Methods”.

References

1. Betz, S., Eklund, R., Wagner, P.: Prolongation in German. In: Eklund, R. (ed.), Proceedings of DiSS 2017, Royal Institute of Technology, Stockholm, Sweden, 5-8 (2017).
2. Bogdanova-Beglarian, N.V., Blinova, O.V., Zaides, K.D., Sherstinova, T.Yu.: Korpus «Sbalansirovannaya annotirovannaya tekstoteka» (SAT): izucheniye spetsifiki russkoy monologicheskoy rechi [Corpus “Balanced Annotated Text Library” (SAT): studying the specificity of Russian monological speech]. In: Proceeding of the V.V. Vinogradov Institute of the Russian language RAS. Iss. 21. Natsional’nyy korpus russkogo yazyka: issledovaniya i razrabotki [Russian National Corpus: research and development]. Moscow, 111-126 (2019).
3. Bogdanova-Beglarian, N.V.: Predisloviye redaktora [Editor’s preface]. In: Pragmaticheskiye markery russkoy povsednevnoy rechi: slovar’-monografiya [Pragmatic markers of Russian everyday speech: dictionary-monograph]. St. Petersburg, 5-52 (2021).
4. Bondarko, L.V.: Zvukovoy stroy sovremennogo russkogo yazyka [Sound system of the modern Russian language]. Moscow, 175 p. (1977).
5. Cheyf, U.: Na puti k lingvistike, osnovannoy na myshlenii [Toward a thought-based linguistics]. In: Yazyk i mysl’: sovremennaya kognitivnaya lingvistika [Language and Thought: Modern Cognitive Linguistics]. Moscow, 60-88 (2015).
6. Feng, L.: The duration of rhyme and tone in Beijing discourse flow. In: Lin Tao, Wang Lijia, etc. Beijing phonetic experiment record. Peking University Press, 131-159 (1985). 冯隆. 北京话语流中声韵调的时长. 选自林焘、王理嘉等著. 北京语音试验录. 北京大学出版社.
7. Fonetika spontannoy rechi [Phonetics of spontaneous speech]. Leningrad, 243 p. (1988).
8. Geyl’man, N.I. Foneticheskiye kharakteristiki spontannoy rechi (eksperimental’no-foneticheskoye issledovaniye na materiale soglasnykh) [Phonetic characteristics of spontaneous speech (experimental phonetic consonant research)] PhD thesis abstract, Leningrad, 16 p. (1983).
9. Gósy, M., Eklund, R.: Segment prolongation in Hungarian. In: Proceedings of DiSS 2017, Royal Institute of Technology, Stockholm, Sweden, 29-32 (2017).
10. Huang, B., Liao, X.: Modern Chinese. Beijing Higher Education Press, 643 p. (2002). 黄伯荣, 廖序东. 现代汉语(增订三版). 北京高等教育出版社.
11. Krivnova, O.F.: Dlitel’nost’ kak sredstvo realizatsii slovesnogo udareniya v tekste (sopostavitel’nyy analiz raznykh sposobov otsenki vyrazhennosti udareniya v slove) [Duration as a means of realizing verbal stress in a text (comparative analysis of different methods for assessing the severity of a word stress)]. In: Yazyk i rech’: problemy i resheniya: Sb. trudov k yubileyu L.V. Zlatoustovoy [Language and speech: problems and solutions: Coll. of scientific papers for the anniversary of L.V. Zlatoustova]. Moscow, 77-99 (2004).

12. Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis. In: *Journal of the American Statistical Association*, 47, 260, 583-621 (1952).
13. Levitskiy, Yu.A.: K voprosu o vtoroy kommunikatsii [To the question of the second communication]. In: *Problemy sotsio- i psikholingvistiki* [Problems of socio- and psycholinguistics]. Iss. 15. Perm socio-psycholinguistic school: ideas of three generations, to the 70th anniversary of Alla Solomonovna Stern, Coll. articles, Perm, 159-167 (2011).
14. Mann, H.B., Whitney, D.R.: On a test of whether one of two random variables is stochastically larger than the other. In: *Annals of Mathematical Statistics*, 18, 50-60 (1947).
15. *Pragmatischekiye markery russkoy povsednevnoy rechi: slovar'-monografiya* [Pragmatic markers of Russian everyday speech: dictionary-monograph]. St. Petersburg, 520 p. (2021).
16. Qi, S., Zhang, J.: Analysis of consonant length in Mandarin Chinese. In: *Journal of Acoustics*, 7 (1), 8-13 (1982). 齐士钤, 张家驩. 汉语普通话辅音音长分析. *声学学报*.
17. *Russkaya fonetika v razvitii: foneticheskiye «Ottsy» i «Deti» nachala XXI v.* [Russian phonetics in development: phonetic "Fathers" and "Children" of the beginning of the XXI century]. V.V. Vinogradov Institute of the Russian language RAS, Moscow, 464 p. (2013).
18. Seifart, F., Strunk, J., Danielsen, S., Hartmann, I., Pakendorf, B., Wichmann, S., Witzlack-Makarevich, A., Himmelmann, N.P., Bickel, B.: The extent and degree of utterance-final word lengthening in spontaneous speech from 10 languages. *Linguistics Vanguard*, 7 (2021).
19. Shcherba, L.V.: *Russkiye glasnyye v kachestvennom i kolichestvennom otnoshenii* [Russian vowels in qualitative and quantitative respects]. Leningrad, 155 p. (1983).
20. Sun, X.: The ways of translating pragmatic marker ETO SAMOE (based on the material of parallel Russian and Chinese literary texts). In: *Communication Studies*, 8 (2), 323-332 (2021).
21. Svetozarova, N.D.: *Prosodicheskaya organizatsiya vyskazyvaniya i intonatsionnaya sistema yazyka* [Utterance prosodic organization and language intonation system] Doctoral dissertation. Leningrad, 514 p. (1983).
22. Teng, H., Li, Y.: The commonality and characteristics of the hesitant dragging phenomenon in natural spoken Russian and Chinese. In: *Journal of Shanxi Datong University (Social Science Edition)*, 3, 116-120 (2022). 滕海, 李毅芳. 俄语和汉语自然口语犹豫型拖音现象的共性及特性. *山西大同大学学报(社会科学版)*.
23. Zaides, K.D.: Ob unifikatsii razmetki korpusa "Sbalansirovannaya annotirovannaya tekstoteka" [On the standardization of the corpus "Balanced Annotated Text Library"]. In: *Trudy mezhdunarodnoy konferentsii «Korpusnaya lingvistika-2019»* [Proceedings of the international conference "Corpus linguistics-2019"]. St. Petersburg, 332-339 (2019).
24. Zlatoustova, L.V.: *Foneticheskiye yedinitsey russkoy rechi* [Phonetic Russian speech units] Moscow, 105 p. (1981).