RUSSIAN SPONTANEOUS SPEECH RATE — BASED ON THE SPEECH CORPUS OF RUSSIAN EVERYDAY INTERACTION

Svetlana Stepanova

Philological Department, St.Petersburg State University, Russia
stsvet_2002@mail.ru

ABSTRACT

The paper presents the results of the analysis of Russian spontaneous speech rate made on the basis of recordings of 40 speakers and their communicants. The impact of gender, age and place of residence on the speech rate is illustrated by the findings.

Keywords: speech rate, Speech Corpus of Russian Everyday Interaction, ORD

1. INTRODUCTION

Temporal speech characteristics may be studied from various perspectives. On the one hand, researchers focus on changes in the pronunciation tempo of spoken fragments (sounds, syllables, words) depending on the position of the fragment within a text, phrase, syntagm, or word, or depending on the functional yield of a given speech fragment. In this case, researchers register changes in the fragment length: e.g., the pre-pause lengthening of final words in a syntagm, or longer duration of vowels in the absolute final position in a word, or the higher tempo in the pronunciation of semantically less significant fragments, etc. On the other hand, tempo is considered as an integral characteristic of a certain person's speech: one person speaks fast, "jabbers", another speaks slowly, with lengthy pauses. What does it mean? According to E. Sapir, "when it is said that we pronounce an utterance fast or slowly, the implication might only be that we surpass a certain tempo inherent to a given society, or fall behind it" [9].It is assumed that there is an average generally accepted tempo in each linguistic community, and deviations from this rate may be attributed to numerous factors: the speaker's age, gender, education, and so on. These hypotheses are currently being verified with the help of spontaneous speech recordings of Russian native speakers, collected in Speech Corpus of Russian Everyday Interaction "One Day of Speech" (ORD Corpus). A characteristic feature of the corpus is that speech is recorded not in a laboratory or in the presence of the researcher or interviewer, but with the help of a technique designed by the developers of the British National Corpus [1]: on an average day, volunteer informants carry a speech recorder on the neck all day long and thus record their own speech and the speech of their conversation partners.

Currently, due to the efforts of a research group, 350 hours of Russian spontaneous speech have been recorded in this manner. The recordings include the speech of 46 informants (23 men and 23 women) of different age (varying from 16 to 70 years old) and different professions, as well as the speech of their 650 communicants. About 40 hours of the recording have been transcribed (on the average, one hour per first 40 informants) with multi-level annotation using the professional tool ELAN [3].

While making the recordings for the Corpus, we asked the informants to fill in a form and indicate their sociological data: gender, age, place of most lengthy residence, education, and profession. With the help of the recordings, linguists-experts assessed the speakers' level of verbal competence [7]. The acquired information allows establishing the dependence of the speakers' speech rate on these personal characteristics. Let us consider in detail the first results.

2. GENERAL CHARAKTERISTICS OF SPEECH RATE IN THE ORD CORPUS

It is well-known that the perception of the speech rate is affected by both the tempo of articulating sounds and by the presence of pauses of certain duration. At the first stage of the project, we decided to consider the "articulation" tempo, regardless of pauses.

In the ORD Corpus, we selected the utterances without pauses, sighs, laughter, and without overlapping of several speakers' words. The total number of selected utterances is 13,459, their total length is 19,298 seconds, which equals 5.36 hours. Rate was calculated as the number of syllables per

second. The number of syllables was automatically set as the number of vowels. The material was processed using the STATISTIKA program [10].

All material considered, the average rate equaled 5.31 syllables per second (syll/s), see Table 1. In the Table, "m" is for male speakers, "f" is for female speakers, "k" is for children.

The result we received, as regards average values, does not contradict the data available in publications on Russian experimental phonetics. Thus, for example, O.F. Krivnova claims that the average duration of a syllable in the Russian language is 150-200 ms, which corresponds to 4.76 - 6.67 syll/s [5].

Table 1: Average speech rate.

Descriptive Statistics (Tempo.sta)

|	Valid N	Mean	Minimum	Maximum	Std. Dev.
f+m+k	13459	**5,31**	0,63	13,19	1,93
m	5067	5,46	0,64	13,19	2,02
k	471	3,86	0,64	9,16	1,67
f	7921	5,23	0,63	12,69	1,84

3. THE INFLUENCE OF THE GENDER FACTOR

The analysis of the material on the basis of Student's t-criterion allowed outlining the statistically valid difference between men's (m) and women's (f) speech rate: men speak substantially (from the statistical perspective) faster than women. Children's rate (k) is the slowest (see Table 1).

Similar results as regards the impact of the gender factor on the speech rate in other languages have been acquired in the research on the Chinese language, American English [4, 12] and Dutch [8, 11] Characteristically, all the authors observe that the difference between men's and women's average speech rate is not substantial but statistically considerable. H. Quene explains men's faster speech rate by their attempts to express their social dominance: «male speakers may also express their social dominance by speaking somewhat faster than female speakers ([8], p.1112).

However, all the researchers recognize that the speech rate is to a great extent determined by the length of the generated utterance. It is possible, as M. Liberman writes, that all other factors affecting the tempo are derivatives of the utterance length: «Maybe region, sex and age don't really influence speaking rate after all, except indirectly via their influence on phrase length ([6] p.1).
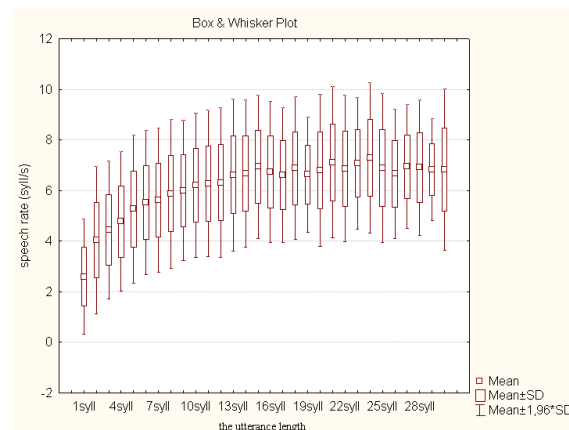
Indeed, it is obvious in fig.1 that the longer the phrase, the faster the rate: all material considered, the average rate in single-syllable utterances is 2.59 syll/s, while in ten-syllable ones it is 6.25 syll/s. However, in even lengthier utterances, the speech rate does not change substantially and remains within the range of average values of 6.2-7.2 syll/s. It may be explained by purely physiological reasons: a syllable, which consists of at least one vowel, has to be articulated by the speech organs, and the movement rate of the latter cannot grow infinitely. Besides, speech perception also sets certain limitations subconsciously recognized by the speaker: "when the rate exceed maximum, syllable recognition may be impeded or even become impossible. These limitations may be explained by the fact that the process of interpretation requires time and that duration is one of the formal indicators of a stream element necessary for defining the latter (a section shorter than a certain value cannot be an element)" [2] p.14.

In order to take into account this phenomenon, which in western linguistics is referred to as «anticipatory shortening», we checked the average syllable length of the analyzed fragments. Our material showed that the average number of syllables in men's utterances is slightly less than that in women's utterances (8.3 vs. 8.5), which means that Russian men on the average pronounce shorter phrases but at a higher rate.

We also made a more detailed examination of the material, comparing the average speech rate of men and women in utterances ranging from one to 20 syllables.

Statistical analysis showed that in single-syllable utterances the speech rate does not depend on the speaker's gender and is almost twice as

slow as the average rate in the whole body of the material - 2.6 syll/s (see Fig.2). Two-syllable utterances were pronounced by women a little faster than by men (4.09 vs. 3.97 syll/s), however, the difference is not statistically significant. In all other 18 cases (utterances of more than two syllables), men's speech rate was higher than women's, and in 5-, 7-, 10-, 11-, 14- and 20-syllable utterances the difference was statistically significant.

Figure 2: Average speech rate depending on the utterance length (men and women separately).
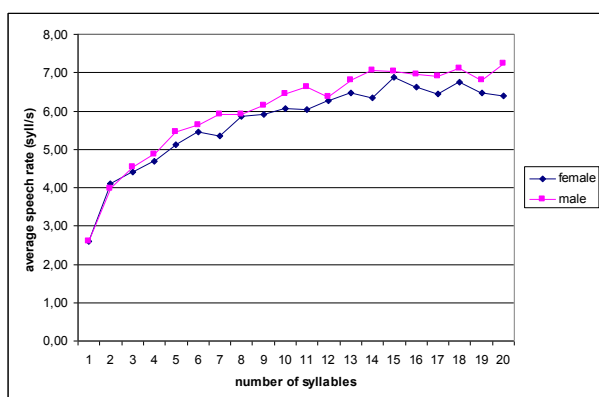
4. AGE INFLUENCE ON THE SPEECH RATE

Next factor which, as the above mentioned research showed, may considerably affect the speech rate is age. We cannot say that the speech rate is directly connected with age. However all our informants may be divided into two groups: below and above 40 years old. H. Quene calls the age of 35 for the Flemish and 45 for the Dutch a "turning point", a point when the gradually growing speech rates starts decreasing [8] p.1110. For our pilot research, we chose the average value of this turning point. The rate difference between these two age groups turned to be quite substantial and confirmed the trend described in all articles on this subject: with age we start speaking more slowly (see Fig. 3), although not reaching the slowness level of children's rate (cf. Table 1).

5. PLACE OF RESIDENCE

Thirty out of 40 informants (75.0%) named St. Petersburg as the place of their lengthiest residence.

Other 10 informants spent most part of their lives in various regions of Russia and former USSR: Karelia (2), Kazan (3), Murmansk (2), Tumen (1), Kiev (1), Belorussia (1). All of them

consider Russian as their native language but their speech shows regional phonetic variations. Obviously, with such an unbalanced corpus of informants, it is not possible to conduct an accurate analysis of the dependence of speech rate on the place of residence. However, we found it interesting to compare the speech rate of residents and non-residents of St. Petersburg.

Figure 3: Substantiality of average speech rate differences (ages below and above 40).
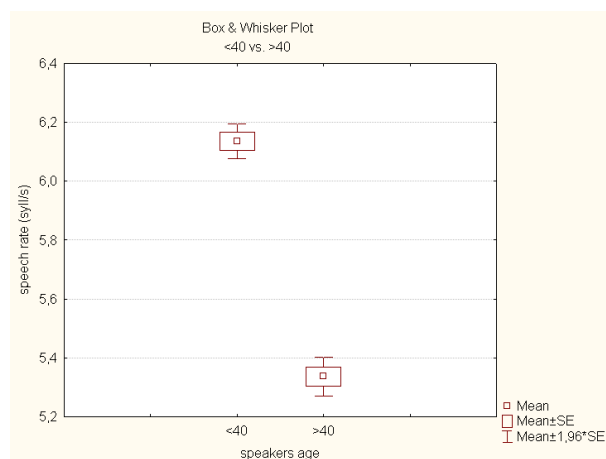
Having divided our informants into two respective groups, we received the data presented in Fig. 4. As we can see, the difference between these two groups is statistically considerable: residents of St. Petersburg speak more slowly.

Figure 4: Substantiality of differences in average speech rate values: St. Petersburg residents and non-residents.


6. OTHER FACTORS

The factor of "education" and "profession" did not prove to yield substantial differences in statistical analysis of groups with higher, incomplete higher, and secondary education. It may probably be explained by the lack of recordings of people

without higher education in the corpus and the high level of dispersion of average values of their speech rate due to other factors.

However, the level of verbal competence of the informants obviously affected the difference in the average values of their speech rate (see Fig.5): it turned out that the informants whose level of verbal competence was assessed by experts as high speak more slowly, while the articulation rate higher than the average is typical of speakers with lower level of verbal competence. Possibly, a slow articulation rate is a sign for experts to assess the informants as verbally competent. Probably, the analysis of "full rate" with consideration of pauses and their length will yield different results.

Figure 5: Substantiality of differences in average speech rate values: high – average – low verbal competence level.
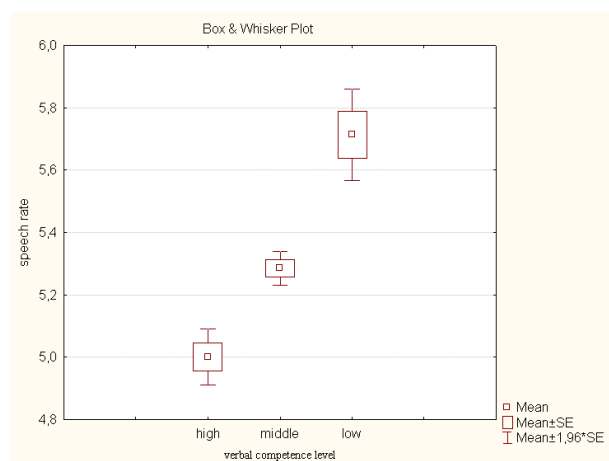
7. PROSPECTIVE RESEARCH

The paper presents the first results of the analysis of temporal characteristics of the recordings in the Speech Corpus of Russian Everyday Communication (ORD-Corpus). They provide general average values of the speech rate of Russian speakers and show the variation of these average values due to gender, age, and verbal competence. Further research may show how the speech rate is affected by the psychological characteristics of a person and to what extent the rate may vary as an integral characteristic of a person's speech depending on the communicative situation.

8. ACKNOWLEDGEMENTS

9. REFERENCES

[1] British National Corpus*: http://www.natcorp.ox.ac.uk*
[2] Chistovich, L., Ventsov, A., Granstrem, M., et al. 1976. Speech Physiology. Human Perception of Speech. *Manual of Physiology*. Leningrad: Nauka.
[3] ELAN. Program available for free downloading from the site of Max Planck Institute of Psycholinguistics: *http://www.lat-mpi.eu/tools/elan/.*
[4] Jacewicz, E., Fox, R.A., Wei, L. 2010. Between-speaker and within-speaker variation in speech tempo of American English. *J Acoust Soc Am.* 128(2), 839-850.
[5] Krivnova, O. 2007. *Rythmization and Intonation Articulation in the "Speech-Thought Process".* Dissertation abstract, Moscow.
[6] Liberman, M. 2006. Guys are a bit gabbier in Dutch, too. LanguageLOG: *http://itre.cis.upenn.edu/~myl/languagelog/archives/003682.html*
[7] ORD-Corpus. In Asinovsky, A., Bogdanova, N., Rusakova, M., Ryko, A., Stepanova, S., Sherstinova, T. 2009. The ORD speech corpus of Russian everyday communication "one speaker's day": Creation principles and annotation. *In LNCS/LNAI Series. Text, Speech and Dialogue* TSD-2009. Springer-Verlag Berlin Heidelberg, 250-257.
[8] Quene, H. 2008. Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *J Acoust Soc Am.* 123(2), 1104-1113.
[9] Sapir, E. 1993 Speech as a personality characteristic. *Selected Works in Linguistics and Culture Studies.* Moscow: 285-297.
[10] STATISTICA-8. *http://95.211.120.162/96128-soft-statsoft-statistica-8-0-550-enterprise-single-user.html*
[11] Verhoeven, J., De Pauw, G., Kloots, H. 2004. Speech rate in a pluicentric language. A comparison between Dutch in Belgium and in Netherlands. *Language and Speech* 47(3), 297-308.
[12] Yuan, J., Liberman, M., Cieri, C. 2006. Towards an integrated understanding of speaking rate in conversation. *ICSLP International Conference on Spoken Language Processing* Pittsburgh, Pennsylvania.