

Ščerba's Leningrad Phonological School in the XXI Century

Pavel SKRELIN

I. INTRODUCTION

It has been commonly accepted that Lev Ščerba's concept of the phoneme as the smallest unit of the sound structure that can serve to differentiate words (Ščerba 1974: 156-158) has had the strongest influence on the development of general phoneme theory, as it provided a new and widely accepted link between sounds and meaning. At the same time, Ščerba did not regard phonetics and phonology as independent from one another. Rather he emphasized the importance of experimental phonetic studies since the phonetic material obtained from the analysis of real speech events provides the basis for phonological generalizations. Ščerba's (The Leningrad / St. Petersburg) phonological school has been known for the combination of theoretical postulates based on the analysis of the language system with careful experimental verification of the features of its sound manifestation.

Lev Ščerba had studied phonetics in Paris with Passy and Rousselot before coming to St. Petersburg to head the experimental phonetics lab in 1909. Actually, experimental phonetic research at our Department started with the acquisition of special equipment available at that time: Lev Ščerba purchased a unique collection of tuning forks from a well-known German firm Zimmerman, resonators from Leppin & Masche, and resonator tubes from K.L. Schaefer. They now form part of the museum collection of our Department.

In the 1970-1980s, experimental phonetic research was based on analog technology. The transition to the new – digital – level of technology was associated with Prof. Dr. Christian Sappok from Ruhr University (Bochum, Germany). In 1988 he presented the Department of Phonetics with its first personal computer along with special software "SONA" and organized a seminar at which he trained our researchers.

The combination of our traditional methods of experimental-phonetic research with the potential of new digital technology stimulated a wide range of scientific projects in different areas of speech investigation.

Mastering digital technology took place in parallel with the creation of The Phonetic Fund of the Russian language (a part of the “Mašinnyj Fond Russkogo Jazyka”), the first national speech corpus. A method of phonetic (“acoustic”) transcription based on the perceptive analysis of delexicalized speech fragments of 1-2 syllables long taking into account their spectral characteristics was developed under the guidance of Prof. L. Bondarko. The method was first used by V. Kuznecov in his study of Russian vowels (Kuznecov 1997). Later it was applied for the creation of the corpora of Russian spontaneous and read-aloud speech (De Silva, Ullakonoja 2009) and of Russian professional read speech (Skrelin, Vol'skaja, Košarov, Evgrafova, Glotova, Evdokimova, 2010). At the same time a method of speech signal segmentation into fragments equal to physical realizations of allophones of the Russian phonemes was developed (Skrelin 1999). This method was used for labeling the indicated corpora and for the segmentation of the sets of diphones (Bondarko, Kuznecov, Skrelin, Svetožarova, Talanov, Vol'skaja, Žarkov, 1996), allophones and sub-allophones (Skrelin 2000) for Russian text-to-speech synthesis systems.

Different database technologies were elaborated to facilitate access to digitized sound archives (several genres of folklore recordings, not only Russian), dialectal and accented speech (Skrelin 2004). This work is still going on, because new aspects of speech material description emerge: the data provided by the articulograph, electroglottograph, video, etc. (Košarov, Skrelin 2014).

Special scripts and software to retrieve and evaluate phonetic parameters of speech signal together with methods of formal description and interpretation of sound units were required. In 2005 a section of Formal methods of Russian speech analysis was organized as part of the International Philological Conference of our faculty. In the framework of this section problems of automatic speech analysis and interpretation had been discussed until 2015 when it was integrated with the section of Phonetics. In 2008 a Special interest group of ISCA (International Speech Communication Association) for formal methods of Russian speech analysis (SIGRU) was organized to promote interest in such methods and to provide access for group members to the developed speech corpora as well as tools for the automatic processing of speech signal. At the same time the annual workshop for the Analysis of Russian conversational speech hosted by the St. Petersburg Institute of Informatics of the Russian Academy of Sciences provided the opportunity to discuss similar issues with mathematicians, engineers and business representatives.

As a result, at the beginning of the XXI century, we have at our disposal different kinds of carefully prepared speech material (language material – Ščerba' 3-rd aspect of language phenomena), tools for its automatic processing, new methods for analysis and description of the sound units functioning at different language levels (language system – Ščerba' 2-nd aspect of language phenomena) during speech activity

(Ščerba's 1-st aspect of language phenomena) in speech production and speech perception (Ščerba 1931). A short presentation of the most interesting research results obtained at the Department of Phonetics follows.

2. PHONETICS OF VARIOUS SPEECH TYPES

2.1. PHONETICS OF SPONTANEOUS VS READ-ALOUD-SPEECH

This study began in 2001. Its goal was to reveal both common and language-specific phonetic properties of read and spontaneous speech in three typologically unrelated languages – Russian, Finnish, and Dutch. These languages differ in prosody, sound systems, and means for conveying intonational meaning. Spontaneous speech was recorded from 8 to 10 speakers in each language. Transliterated extracts from spontaneous speech recordings were read aloud by the same speakers. The two types of speech in the three languages studied provided data for comparing their F0 statistics, segmental duration, and the number of some consonant elisions. (For more detail see: De Silva, Ullakonoja 2009; De Silva, Iivonen, Bondarko, Pols 2003; Bondarko, Vol'skaja, Tananajko, Vasil'jeva 2003). In my opinion, the most interesting results obtained for the Russian material are as follows:

a) Individual strategies of speakers when differentiating spontaneous speech from read speech rather than some general tendencies.

b) The difference in phoneme realization between spontaneous speech and text reading is not regular and in some cases is absent (Skrelin 2004).

a) At the prosodic level we observe considerable deviation of real speech data from what was theoretically expected. There is no single prosodic parameter which any speaker consistently employs to demonstrate "spontaneous versus read speech" differences. Prosodic parameters variability signals a change in speech style; by randomly varying a set of features speakers help listeners to perceive this change. In this respect Standard Deviation seems to be the most reliable parameter with which to describe the degree of such variability (Skrelin, Vol'skaja 2006).

2.2. READ-ALOUD SPEECH: CORPRES

CORPRES is a fully annotated **COR**pus of **R**ussian **P**rofessionally **R**ead **S**peech developed at the Department of Phonetics, St. Petersburg State University, as a result of a three-year project. The corpus was originally intended for the use in unit-selection TTS synthesis. However, it appeared expedient to create a corpus that would be an adequate and reliable

representation of Standard Russian speech suitable as a basis for a wider range of research, e.g. variation and change in Standard Russian, among other research subjects. Manual expert segmentation of 40 % of the corpus and expert annotation and transcription also make the corpus an excellent database for phonetic research on contemporary Russian.

Table 1 shows general corpus statistics. The number of phonemes in the part of the corpus which was not annotated on phonetic transcription levels has not been provided, therefore two cells in the table remain unfilled. For more details see Skrelin, Vol'skaja, Košarov, Evgrafova, Glotova, Evdokimova 2010.

For the last 5 years, the CORPRES has been used as a basis for different experimental-phonetic studies the presentation of which follows.

Table 1. General corpus statistics

	Fully annotated data	Partly annotated data	Total amount
Phonemes	1,048,867	–	–
Words	211,437	317,021	528,458
Tone units	64,055	86,546	150,601
Hours	24	36	60

2.3. SPEECH VS SINGING (2012 – 2015)

The study is concerned with issues of low vowel intelligibility in singing especially at high pitch levels. In a first step, the Russian vowels were produced by two professional opera singers (female singers aged 28 and 32) currently employed at the Mariinsky Theater in St. Petersburg.

The perceptive tests showed that the high-pitched vowels [i] and [u] sung in isolation have a relatively low intelligibility and tend to be perceived as [a]. Their formants obtained by the acoustic analysis are repositioned significantly in comparison with those typical of normal speech. The intelligibility of high-pitched vowels decreases to such an extent that their phonological status can be challenged (Evgrafova, Evdokimova 2012).

The next study employs method of electromagnetic articulography (EMA) to obtain exact data on articulatory settings in singing. Two types of recording experiments with the use of EMA have been conducted involving four trained female singers. In the first experiment they were instructed to sing one of the Russian classical romances with the AG500 sensors attached to main articulators. The second experiment involved reading aloud the text of the same romance.

The comparing of kinematic data in singing and reading showed that in general the amplitude and patterns of articulatory movements in singing differs considerably from those in reading. The vertical and

horizontal articulatory movements are more prominent than those in reading. The analysis of difference in kinematic characteristics provided reasons for the acoustic distortion in the quality of sung vowels (Evgrafova, Evdokimova, Skrelin, Čukaeva 2015).

2.3. VOCAL FATIGUE

Vocal fatigue is a voice disorder which particularly concerns professional voice users and can lead to serious pathological conditions. Vocal fatigue may result in some acoustic characteristics like changes in pitch, loudness, pauses number and duration, voice quality etc..

A pilot study was performed a few years ago on the basis of recordings of the phonetically representative text, made by 5 pronunciation teachers in their normal state and in a tired state, after 6 or 8 hours of intensive work. The first results were very promising: the acoustic analysis showed a consistent dependency between acoustic parameters and vocal fatigue. After an intensive working day, F0 values were higher, the duration of vowels and consonants increased; pitch and loudness range values increased. The differences in the acoustic parameters after a vocally loaded working day mainly seem to reflect increased muscle activity as a consequence of excessive vocal loading (Evgrafova, Evdokimova 2015).

In our next experiment we studied vocal fatigue in lecturers, professional speakers and tour guides (20 male and 20 female subjects). After their working day, the subjects reported symptoms of a high degree of vocal fatigue such as a high level of muscular tension / discomfort, hoarse voice quality, breathy voice quality, unsteady voice, inability to maintain typical pitch, dry throat etc. The decrease of precision in the determination of the speaker's physiological state in extensive research material demonstrates that similar fatigue states and speaker self-assessment may display different acoustic manifestations due to the speaking behavior in different professional groups (Skrelin, in print).

3. LANGUAGE SYSTEM: SOUND UNITS FUNCTIONING

3.1. INTONATION MODEL CLASSIFICATION

Appropriate prosody should include models for conveying linguistic functions of intonation – sentence delimiting, sentence forming, distinctive and expressive or attitudinal. More sophisticated models regard the possibility of covering pragmatic information conveyed by intonation contours.

For the CORPRES prosodic annotation, a new classification has been developed as an extension of the well-known intonation system of

E. Bryzgunova (Bryzgunova 1980). The proposed classification consists of 13 intonation models with their variants. Its main goal is to fill the gaps in Bryzgunova's system, which proved insufficient for the analysis of spontaneous speech intonation, and to ensure close connection between functional differences between intonation models and variation of their formal acoustic features.

Most theories of intonation make distinctions between falling and rising tones to fulfill these functions. Each prosodic model necessarily comprises these two categories – falling and rising, each type being linked to a linguistic function. At present moment our model includes contours which end with a final fall (used in declaratives, wh-questions, imperatives and most exclamations), a final rise (such as used in non-final units and certain types of questions), and level tones, both final and non-final. The labeling scheme consists of 13 basic contour types with up to 4 subtypes for each and 6 break levels. For more detail see Vol'skaja, Skrelin 2009; Skrelin, Vol'skaja 2009.

A set of the acoustic features used for automatic interpretation of sentence prosody is described in Skrelin, Košarov 2009.

3.2. DECLINATION STUDY

The study was conducted using statistical data derived from the CORPRES. We concentrated on two prosodic constituents: the intonational phrase (IP) and the prosodic word (PW) (other terms for it are “metrical group”, “rhythmic group”, “accent group”, “phonological phrase” etc.).

Selected IPs differ in size (from 3 to 6 PWs) and represent different types of utterances. The total material analyzed includes 13321 IPs.

F0 declination was estimated by the top-line of F0 contour. The top-line was calculated using F0 data (in semitones) for each successive pitch accent in the IP as the difference (in semitones) of the F0 maximum of the accented vowel in the PW and the F0 maximum of the accented vowel of the first PW in the IP. Scaling of pitch accents for various intonation contours was calculated by averaging F0 values for all the intonational phrases of a corresponding type (final, non-final declarative and interrogative).

The results of the study confirm the tendency for F0 to decline to the end of final declaratives in Russian. At the same time our approach to calculating declination revealed the various strategies a speaker follows to fulfill the task of completing the downstepping trend within the intonational unit: “classical” (Fig. 1), in which the length of the intonational phrase “sets” the end frequency level for the speaker (constant slope); “proportional”, within a fixed range: when pitch accents are almost regularly spaced out and the pitch level of each successive accent is defined by the length of the intonational unit in prosodic words (Fig. 2); “obligatory” (induced), when downstepping involves the last three pitch accents in the intonational phrase (Fig. 3).

For non-final intonational phrases, declination line is less steep, though it also demonstrates dependency on the length of the intonational phrase – for all speakers it increases with the number of the PWs, but does not exceed 4 semitones (Fig. 4).

F0 trend in Russian general questions shows that declination is under the control of the speaker and that pre-planning of the whole contour does take place, for in this type of utterance there is no F0 declination regardless of the length of the intonational phrase (Fig. 5).

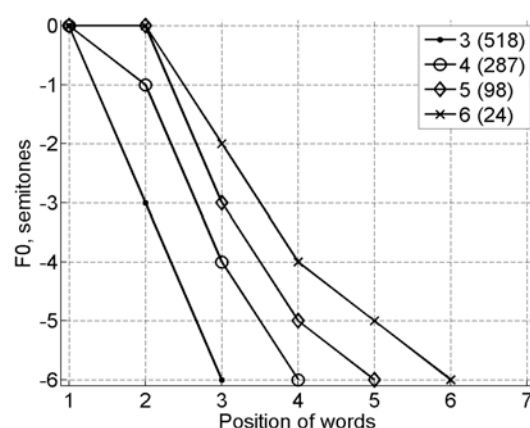


Fig. 1. "Classical" declination line

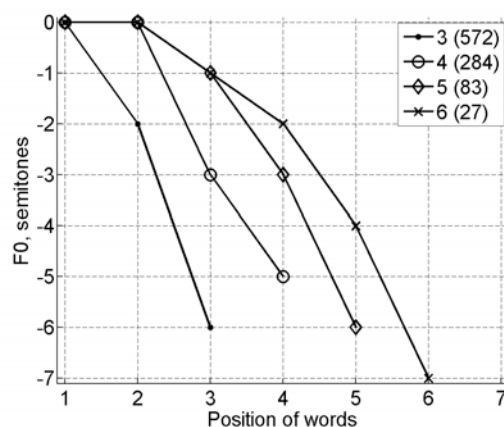


Fig. 2. "Proportional" declination line

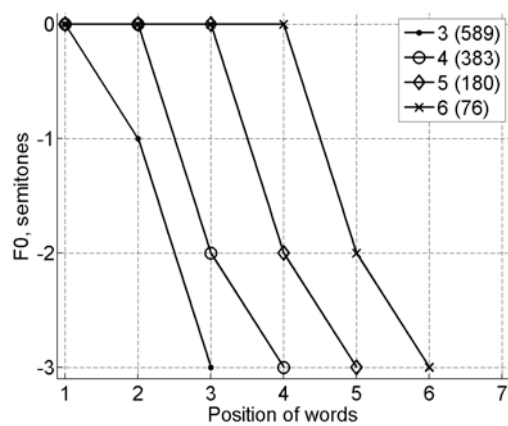


Fig. 3. "Obligatory" declination

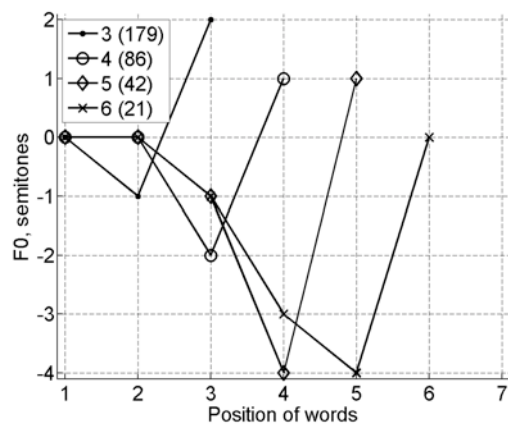


Fig. 4. Declination line in non-final phrase

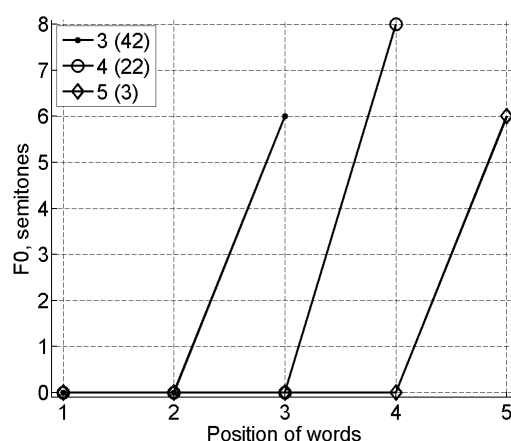


Fig. 5. No declination in general questions

The relationship between F0 slope and the utterance length – the shorter the utterance the steeper the slope – confirms the possibility of planning the declination and the existence of the look-ahead strategy of the speaker (Levelt 1989: 400). At the same time, there is clear evidence of individual strategies in scaling the pitch level of successive accents particularly in intonational phrases containing more than 5 words: for speakers with a narrow individual range, the pitch level at the beginning of the intonational phrase is sustained until the last two pitch accents before the nucleus when downstepping of the pitch accents is resumed to reach the pre-planned F0 target end level. For other speakers, increased length of the intonational phrase results in lowering the pitch level of the final accent. For more detail see Košarov, Skrelin, Vol'skaja 2014; Vol'skaja, Vorob'jeva 2010.

The description of declination would not be complete without considering its behavior in case of prominence. Figures 6 and 7 illustrate prominence in final declaratives containing 4 PWs pronounced by the same speaker. They show F0 trends in units with and without an emphasized word. A solid line shows F0 change within the accented vowel of the prominent word and the arrow indicates direction of the tonal movement: falling or rising.

Fig. 6 shows that the emphatic accent is manifested by the declination reset (upstep) and F0 rise within the prominent word. The declination returns to its 'original' trend afterwards; note that scaling of pitch accents between the third and the fourth word is independent of prominence. Fig. 7 shows that shifting the accent from the fourth word to the third does not influence the declination at the beginning of the phrase; the declination slope resumes its trend after the prominent syllable.

Our research shows that in both cases prominence influences the declination locally, but not globally: prominence affects only the F0 of the prominent word. For more detail see Vol'skaja, Vorob'jeva 2010; Košarov, Vol'skaja, Skrelin 2015.

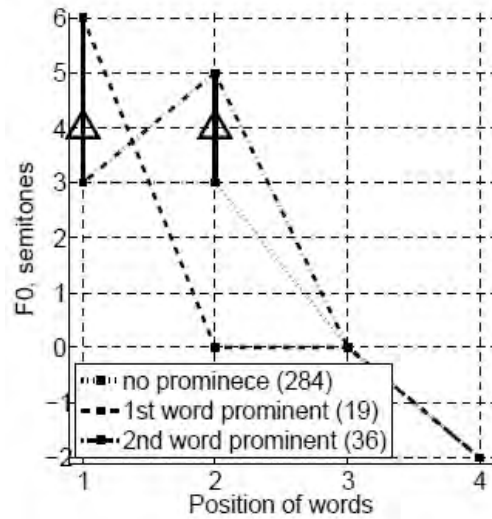


Fig. 6. F0 declination in case of emphatic accent

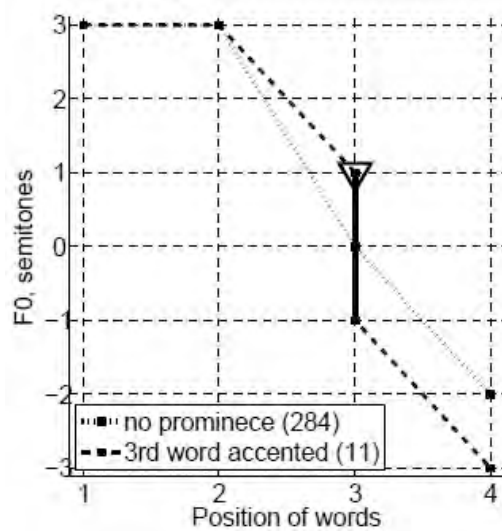


Fig. 7. F0 declination in case of accent shift

3.3. VOWEL REDUCTION

Previous studies of the unstressed vowel reduction in Russian show lack of agreement between researchers concerning the number of its degrees and the quality of unstressed vowels (Košarov, Kaškovskaja, Skrelin 2015).

According to our data (CORPRES), for pre-stressed vowels there are two clear degrees of vowel reduction:

- 1-st degree: for vowels in the 1-st pre-stressed syllable and absolute-initial position;
- 2-nd degree (stronger): for vowels in other pre-stressed syllables.

Most post-stressed vowels are reduced in the same way as pre-stressed vowels of the 2-nd degree. However, our data seem to provide evidence for an additional, stronger degree of post-stressed vowel reduction, which is observed for vowels in long words in the 2-nd post-stressed syllable. Thus:

- the 3-rd (equal to the 2-nd) degree : for vowels in most post-stressed syllables;
- the 4-th (stronger) degree: for vowels in 2-nd post-stressed syllables of long words.

The existence of the additional 4-th degree of vowel reduction supports the ideas put forward earlier by L.V. Bondarko et al. (Bondarko, Verbickaja, Gordina 1991) for a post-stressed [a].

Our results enable us to treat vowels in final open syllables as having the 3-rd degree of reduction. Therefore, these results seem to support the ideas suggested by R.I. Avanesov (Avanesov 1984), confronting those of L.V. Bondarko et al. (Bondarko, Verbickaja, Gordina 1991).

In general, the hypothesis that quality and quantity changes of vowels are correlated is not supported by our data. The correlation is observed only for the pre-stressed part of the word: the longer the vowel the lower the number of omissions or replacements. The absence of similar correlation for post-stressed vowels may be explained by the purely grammatical meanings of word endings that are often not pronounced properly. For more detail see Košarov, Kaškovskaja, Skrelin 2015.

3.4. PRE-BOUNDARY LENGTHENING

The study of pre-boundary lengthening in Russian was also based on a large speech corpus (CORPRES) that allowed us to obtain statistically reliable results.

The first question in this study addressed the cause of such lengthening: the presence of a boundary or the presence of a pause. To answer it the analysis of the duration of stressed and post-stressed vowels in words occurring at the beginning/middle vs. end of the intonational phrase with vs. without a following pause was performed. The results show:

– In words followed by a pause, stressed vowels in penultimate syllables are longer in IP-final position than in IP-initial/medial position. Therefore, here the lengthening is caused not only by the presence of a pause, but also by the position of the word within the phrase.

– For post-stressed vowels in final open syllables the opposite is observed: absolute final vowels are much longer in non-phrase final position before a pause than in phrase-final position before a pause.

– Post-stressed vowels in final closed syllables do not show any lengthening in either of the cases (Kaškovskaja 2014).

For Russian our data provide evidence that the deeper the boundary, the weaker the lengthening effect. In other words, the speaker marks the end of non utterance-final intonational phrase better than the end of the utterance-final intonational phrase — thus showing that the utterance is not finished and the listener is expected to wait for its ending. However, this relation might be speaker-specific, since not all the speakers show a statistically significant difference between these two boundary types (Kaškovskaja 2015a).

Our data show that stressed vowels in penultimate syllables play a greater role in phrase-final lengthening than post-stressed vowels. However, it is worth noting here that final rhyme may not only consist of a vowel, but also include following consonants. Previous studies (Kaškovskaja, Vol'skaja 2013), however, have shown that absolute-final consonants do play a significant role in phrase final lengthening.

The CORPRES was also used here to study the interaction between the type of pitch movement and the position of the word within the phrase in their influence on segments' duration. The results can be found in Kaškovskaja 2015b.

4. SPEECH ACTIVITY: SPEECH PRODUCTION AND SPEECH PERCEPTION

4.1. ACOUSTIC THEORY OF SPEECH PRODUCTION

A few years ago, we proposed a new method of recording a voice source signal. It allows the voice source to be registered by means of a special miniature microphone which is placed in the proximity of the vocal folds (Evgrafova, Evdokimova, Skrelin, Čukaeva, Švalev 2015). Thus an opportunity was presented to record the voice source signal and the output speech signal synchronously. The comparison of the recorded signals allowed the structure of the speech signal at different stages of its generation to be analyzed.

The comparison of the recorded signals made it possible to calculate different vowel transfer function, to generate artificial sounds with set

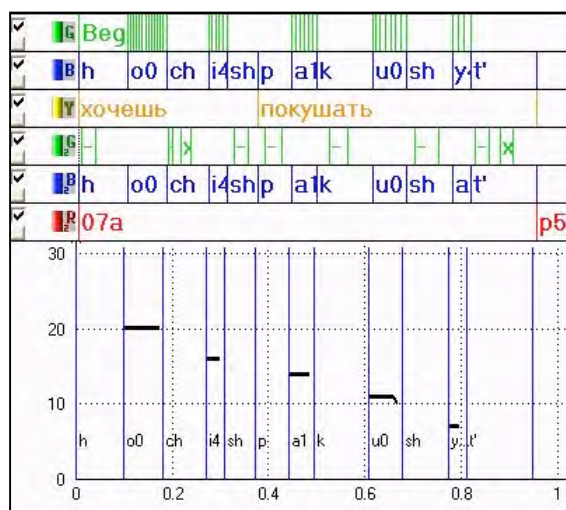
parameters and to realize different perceptive tests on vowel recognition. The results of our studies show:

1. The low frequency formants are formed near the vocal chords and do not change significantly in the vocal tract.
2. The high frequency formants are absent near the vocal chords and are formed by the vocal tract.
3. The low frequency formants formed by the vocal chords carry sufficient information for intelligibility of those vowel phonemes that do not depend essentially on high frequency formants. For more detail see Evdokimova, Evgrafova, Skrelin 2015; Barabanov, Evdokimova, Skrelin 2015.

4.2. PERCEPTION OF RUSSIAN INTONATION

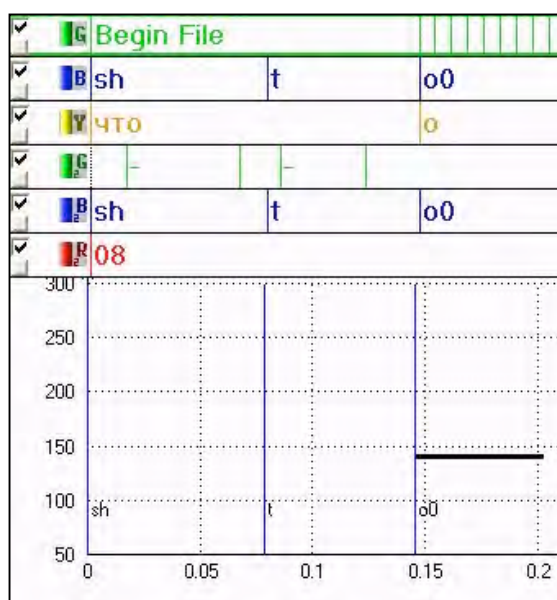
The use of the previously (3.1) described classification of intonation models showed the discrepancy between the acoustic properties and perceptual evaluation of the rising-falling contour #07 that is proper to general question. Due to “phonological hearing”, Russian listeners interpret the acoustically falling melody of a question like ‘*Хочешь покушать?*’ (Fig. 8) – it begins with a high tone on a first stressed vowel following a voiceless consonant – as a rising one (Skrelin 2012). Such interpretation is based on the fact that the intonation center of this kind of question is usually preceded by a few words or at least by one or more pre-stressed syllables carrying the rising tone. In this case, the foreign listeners hear a falling tone.

Fig. 8. High-level tone in focus.



A high-level tone in a question like ‘*Чмо?*’ (Fig.9) Russian listeners interpret this as a rising melody and foreign listeners, as a high-level tone. But the problem is, “how does any listener know that the tone is high if the speaker is unfamiliar to him?” (Skrelin 2011).

Fig. 9. High-level tone.



Russian has a rich inventory of rising tones. Rising-falling intonation, which is typical for general questions and non-finality in standard spoken Russian, has always been considered specifically Russian. This creates problems for speakers of those languages whose intonation systems either lack this type of contour altogether, such as Finnish f.ex., or use it for other purposes, such as English or German.

The study of the perception and interpretation of Russian intonation by foreign listeners was implemented in the set of perception tests, with the participation of speakers of Finnish, German, English, etc. (Skrelin, Vol'skaja, Evgrafova, Ullakonoja 2014).

Observations of the speech behavior of foreign learners of Russian allow us to admit that Russian question intonation is often misinterpreted both phonologically and pragmatically: this leads to misunderstanding and miscommunication.

Perception experiments revealed that for a German listener, Russian rising-falling tone is not associated either with a rise or a question (most answers regarding the type of the tone were "falling"). It is least of all

associated with a polite request for information (most answers were "statement or exclamation"). Second, a rise-fall was often perceived as conveying a negative overtone. Neither of these have been intended by the Russian speaker! (Skrelin, in print).

4.3. PERCEPTION OF CHILDREN'S EMOTIONS IN SPEECH

Our study was based on the speech material of two corpora. The FAU Aibo Emotion Corpus had been collected in the Pattern Recognition Lab of Friedrich-Alexander University and the Corpus of Russian Children's Emotional Speech was recorded at the Phonetics Department of St. Petersburg State University. Our interest in children's emotional speech is based on the assumption that emotions are universal, but the forms of their manifestation or masking may be language and culture dependent. The emotional reactions of children are expected to be more physiological and less culture- and language-specific.

Three experiments were carried out to investigate differences and similarities in the assessment of emotions by German and Russian adult listeners. The corpora of German and Russian emotional children's speech were employed in the first and second experiments. In the third experiment German and Russian 'delexicalised' utterances were used. They were selected from both corpora and then white noise was added to them. Thus the semantic content was removed while the prosodic features stayed intact. The experiment was aimed at analyzing recognition strategies when listeners rely only on prosody after segmental information had been removed. The experiments revealed similar and different patterns of assessing emotions in children's speech in German and Russian. For more detail see Evgrafova, Skrelin, Šatalova 2015.

5. CONCLUSION

As can be seen from the above, the Ščerba's school is still alive and working on diverse aspects of contemporary research in phonetics and phonology. We follow Ščerba's ideas in language teaching as well: thus, a learner-centered teaching model, one of the trends in teaching methods whereby the teacher gives authority to the student has always been favored at the Department of Phonetics. We encourage students to discover facts about language rather than just remain recipients of information. The Department of Phonetics of St. Petersburg University seems to be the only place in Europe (or even the world?) which provides specialization in phonetics and phonology as well as speech communication and speech technology for BA, MA and PhD students. Our students take part in research related to various aspects of native and foreign language pronunciation, cross-language phonetic and prosodic interference, cross-

language comparison between the foreign and the native language, speech technology, new methods of speech signal analysis, interpretation and modeling. We hope that in future more Russian names appear on the list of IPA certified students of phonetics.

© Pavel Skrelin

Acknowledgments.

The research presented in the Chapter 4.1 was supported by Saint Petersburg State University projects 6.37.349.2015 and 31.37.353.2015; the researches presented in Chapters 3.2 - 3.4 are supported by the Russian Science Foundation (research grant # 14-18-01352).

REFERENCES

- AVANESOV Ruben, 1984: *Russkoe literaturnoe proiznošenie*, Moksva: Prosveščenie, p. 76 – 105.
- BARABANOV Andrej, EVDOKIMOVA Vera, SKRELIN Pavel, 2015: "Estimation of Vowel Spectra Near Vocal Chords with Restoration of a Clipped Speech Signal", in: A. Ronžin, R. Potapova, N. Fakotakis (Eds.), *Speech and Computer. 17th International Conference, SPECOM 2015, Athens, Greece, September 20–24, Proceedings, LNAI*, vol. 9319, p. 209-216.
- BONDARKO Lija, KUZNECOV Vladimir, SKRELIN Pavel, SVETOZAROVA Natalia, TALANOV Andrej, VOL'SKAJA Nina, ŽARKOV Igor', 1996: "RUSVOX — the Concatenation Speech Synthesis System for Russian", in: *Proceedings of SPECOM'96*, Sankt-Peterburg, p. 165-166
- BONDARKO Lija, VOL'SKAJA Nina, TANANAJKO Svetlana, VASIL'JEVA Ljudmila, 2003: "Phonetic Properties of Russian Spontaneous Speech", in: *Proceeding of 15th ICPHS*, Barcelona, p. 2973-2976.
- BONDARKO Lija, VERBICKAJA Ljudmila, GORDINA Mirra, 1991: *Osnovy obščej fonetiki*, Sankt-Peterburg: SPbGU, p. 114.
- BRYZGUNOVA Elena A., 1980: "Intonacija", in: *Russkaja grammatika*, Vol. I. Moskva: Nauka, p. 99-120.
- DE SILVA Viola, ULLAKONOJA, Riikka (Eds), 2009: *Phonetics of Russian and Finnish. General Description of Phonetic Systems. Experimental Studies on Spontaneous and Read-aloud Speech*, Frankfurt am Main: Peter Lang.
- DE SILVA Viola, IIVONEN Antti, BONDARKO Lija, POLS Louis C.W., 2003: "Common and Language Dependent Phonetic Differences between Read and Spontaneous Speech in Russian, Finnish and Dutch", In: *Proceeding of 15th ICPHS*, Barcelona, p. 2977-2980.
- EVDOKIMOVA Vera, EVGRAFOVA Karina, SKRELIN Pavel, 2015: "Investigating Source-Filter Interaction to Specify Classic Speech Production Theory", In: *The Scottish Consortium for ICPHS 2015, Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow.
- EVGRAFOVA Karina, EVDOKIMOVA Vera, 2012: "Perception of Russian Vowels in Singing", in: *Proceedings of the Fifth International Conference Baltic HLT 2012*, A. Tavast et al. (Eds.), Tartu, p. 42-49.
- , 2015: "Acoustic analysis of vocal fatigue in professional voice users", in: *Models and analysis of vocal emissions for biomedical applications: 7th international workshop : August 25-27, 2011*, C. Manfredi (ed.), Firenze : Firenze University Press, p. 153-156.
- EVGRAFOVA Karina, EVDOKIMOVA Vera, SKRELIN Pavel, ČUKAEVA Tat'jana, 2015: "The Study of Acoustic-Articulatory Relations in Producing Singing Vowels with the Use of EMA", In: *Models and analysis of vocal emissions for biomedical applications: 9th*

- international workshop: September 2-4, 2015*, C. Manfredi (ed.), Firenze: Firenze University Press, p. 95-98.
- EVGRAFOVA Karina, EVDOKIMOVA Vera, SKRELIN Pavel, ČUKAEVA Tatjana, ŠVALEV Nikolaj, 2013: "A New Technique to Record a Voice Source Signal", in: Cl. Manfredi (Ed.), *Models and analysis of vocal emissions for biomedical applications: 8th international workshop : December 16-18, 2013*, p. 181-182.
- EVGRAFOVA Karina, SKRELIN Pavel, D. ŠATALOVA, 2015: "Cross-Language Perception of Emotional Children's Speech in German and Russian", in: *The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, Paper number 0801.1-9.
- KAŠKOVSKAJA Tat'jana, 2014: "Phrase-Final Lengthening in Russian: Pre-boundary or Pre-pausal?", In: A. Ronžin et al. (eds.): *SPECOM 2014, LNAI 8773*, p. 353-359.
- , 2015 (in print): "The Influence of Boundary Depth on Phrase-Final Lengthening in Russian", in: A.-H. Dediu et al. (Eds.): *SLSP 2015, LNAI 9449*.
- , 2015b: "Temporal Organization of Phrase-final Words as a Function of Pitch Movement Type", in: A. Ronžin et al. (Eds.): *SPECOM 2015, LNAI 9319*, p. 422-428.
- KAŠKOVSKAJA Tat'jana, VOL'SKAJA Nina, 2013: "Phrase-final segment lengthening in Russian: preliminary results of a corpus-based study", In: Železný M., Habernal I., Ronžin A. (eds.), *SPECOM 2013, LNCS*, vol. 8113, p. 257-263.
- KOŠAROV Daniil, Kaškovskaja Tat'jana, SKRELIN Pavel, 2015: "Position-dependent Vowel Reduction in Russian", In: *The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK.
- KOŠAROV Daniil, SKRELIN Pavel, 2014: "Russian Speech Corpora Framework for Linguistic Purposes", In: *Best Practices in Spoken Corpora in Linguistic Research*, M. Haugh, S. Ruhi, T. Schmidt, K. Wörner (Eds.), Newcastle upon Tyne, p. 118-127.
- KOŠAROV Daniil, VOL'SKAJA Nina, SKRELIN Pavel, 2015: "F0 Declination in Russian Revisited", in: *The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK.
- KUZNECOV Vladimir, 1997: *Vokalizm svjaznoj reči*, Sankt-Peterburg: SPbGU
- LEVELT Willem, 1989: *Speaking: from Intention to Articulation*, Cambridge University Press, Cambridge.
- SKRELIN Pavel, 1999: *Segmentacija i transkripcija*, Sankt-Peterburg: Sankt-Peterburgskij gosudarstvennyj universitet.
- , 2000: "Allophone-and suballophone-based speech synthesis system for Russian", in: P. Sojka, I. Kopeček, L. Pala (Eds.), *Text, Speech, and Dialogue, Proceedings LNAI*, v. 1902, p. 271-276.

- , 2004a: "Resursy dlja opisaniya i issledovaniya zvukovoy formy russkoj reči", in: *Sovremennaja russkaja reč': sostojanie i funkcionirovanie*, Sankt-Peterburg, p. 127-130. ['Resources for Description and Study of the Sound Form of Russian Speech']
- , 2004b: "Segment Features in Different Speech Styles", in: *Proceeding of SPECOM2004*, Sankt-Peterburg, p. 11-16.
- , 2011: "Russian Intonation: New Aspects of Acoustic and Perceptual Features", in: *Proceedings of the 1-st International Conference "Phonetics without Borders", June 13-15, 2011*, Blagoveščensk, p. 110-115.
- , 2012: "Avtomatičeskaja obrabotka prosodičeskix dannyx v issledovatel'skix celjax", in: *Čelovek govorjaščij: issledovaniya XXI veka. K 80-letiju so dnja roždenija L.V. Bondarko*, Ivanovo, p. 89-95. ['Automatic Processing of the Prosodic Data for Scientific Purposes']
- , (in print): "German-Russian Language Contact: Is it in our power to foresee the flight of word we uttered?", in: *Deutsch als Bindeglied zwischen inlands- und auslandsgermanistik, Beiträge zu den 23. GeSuS-Linguistiktagen*, Sankt Petersburg, 22-24 Juni 2015, Verlag Dr.
- SKRELIN Pavel, KOŠAROV Daniil, 2009: "Avtomatičeskaja obrabotka prosodičeskogo oformleniya vyskazyvanija: relevantnye prosodičeskie priznaki dlja avtomatičeskoj interpretacii intonacionnoj modeli", In: *Trudy tret'jego meždisciplinarnogo seminaru "Analiz razgovornoj russkoj reči (AR-3 - 2009)"*, Sankt-Peterburg, p. 41-46. ['Automatic Processing of the sentence prosody: relevant prosodic features for automatic interpretation of the intonation model']
- SKRELIN Pavel, VOL'SKAJA Nina, 2006: "Russian read and spontaneous speech: prosodic data analysis", In: *Nordic Prosody IX*, Frankfurt-am-Main, p. 235-244.
- , 2009: "Sistema intonatsionnyx modelej dlja avtomatičeskoj interpretacii intonacionnogo oformlenija vyskazyvanija: funkcional'nye i perceptivnye xarakteristiki", In: *Trudy tret'jego meždisciplinarnogo seminaru "Analiz razgovornoj russkoj reči (AR-3 - 2009)"*, Sankt-Peterburg, p. 28-40. ['The system of intonational models for automatic interpretation of sentence prosody: functional and perceptive features']
- SKRELIN Pavel, VOL'SKAJA Nina, EVGRAFOVA Karina, ULLAKONOJA Riikka, 2014: "The Development of New Corpora for Under-resourced Languages Using Data Available for Well-resourced Ones", In: *Proceedings of the 4th International Workshop on Spoken Language Technologies for Under-resourced Languages SLTU*, p. 243-246.
- SKRELIN Pavel, VOL'SKAJA Nina, KOŠAROV Daniil, EVGRAFOVA Karina, GLOTOVA Olga, EVDOKIMOVA Vera, 2010: "Corpus of Russian Professionally Read Speech" In: P. Sojka, A. Horák, I. Kopecek, K. Pala (Eds.), *TSD 2010, LNAI 6231*, p. 386-393.
- ŠČERBA Lev, 1931: "O trojakom aspekte jazykovyx javlenij i ob eksperimente v jazykoznanii", *Izvestia Akademii Nauk SSSR*, p. 113-

129. [’Du triple aspect des phénomènes langagiers et de l’expérimentation en linguistique’]
———, 1974: “O foneme i ee ottenkax”, In: *Jazykovaja sistema i rečevaja dejatel’nost’*, Leningrad: Nauka, p. 156-158.
- VOL’SKAJA Nina, SKRELIN Pavel, 2009: “Prosodic model for Russian”, in: *Nordic Prosody X*, R. Aulanko (ed.), Frankfurt am Main, p. 249–260.
- VOL’SKAJA Nina, VOROB’JOVA Sof’ja, 2010: “Narušenie linii deklinacii kak sposob realizacii akcentnogo vydelenija”, in: *Trudy četvertogo meždisciplinarnogo seminara “Analiz razgovornoj russkoj reči (AR-3 - 2010)”*, Sankt-Peterburg, p. 64–71. [Deviation from the declination line as a way to prominence realization’]