Russian Phonetic Variability and Connected Speech Transcription

Vladimir I. Kuznetsov and Tatiana Y. Sherstinova

Department of Phonetics, St.Petersburg State University,
Universitetskaya nab. 11, Saint-Petersburg, Russia
e-mail:paul@phonet.lang.pu.ru; tanya@ts4306.spb.edu

Abstract. The study of pronunciation variability is an important phonetic task, which has many applications (in speech synthesis and recognition systems, language teaching, forensic phonetics, etc.), and well-known variability of Russian speech deserves special investigations. Generally, one may distinguish geographical, functional, social, national variability of pronunciation. At the same time any type of pronunciation is characterized by its own internal variability. It concerns as well the standard speech. Database of contemporary Russian speech variability is currently being created in the Laboratory of the Experimental Phonetics of Saint-Petersburg University in order to investigate pronunciation variance of different speech types. Database sound material is formed by recordings of the Phonetically Representative Texts pronounced by 1) standard Russian speakers, 2) regional Russian speakers, 3) speakers from the former USSR republics to which Russian is a second language, and 4) foreigners from European, Asian and American countries. Along with the sound database an expert linguistic system is being created, which allows analysis and investigation of particular speech samples, types and corresponding parameters. An important role here plays the system of speech transcription, which should adequately reflect the actual variability.

1 Introduction

Natural continuous speech of any language is known to be a complex phenomenon depending on great amount of factors. "Weak" articulation of Russian (in contrast to many other languages) and the abundance of Russian grammar structure usually cause considerable modifications and simplifications in continuous speech sound sequences as compared with that of the ideal (full-style) speech. The variety of phonemes pronouncing turns out to be extremely high therefore complicating (sometimes even making impossible) their identification. The problem becomes more complicated when we analyze not-standard speech: that of regional Russian speakers, speakers from the former USSR republics (to whom Russian is a second language), and foreigners. The study of pronunciation variability is an important phonetic task, which has many applications (in speech synthesis and recognition systems, language teaching, forensic phonetics, etc.).

2 Research Methodology

The "distance" from any natural pronunciation and it's "ideal" variant in any phonetic aspect (both segmental and prosodic) may be statistically evaluated. The smallest distance may be observed in standard speech, the distance will increase for colloquial standard speech, regional and dialect speech, and Russian speech by foreigners. We see our task to elaborate the procedures of speech estimation based on its comparison with the ideal pronunciation rules; first - on the material of standard speech, and then - regional and dialect speech.

The following method of speech estimation (which is now realizing in the segmental (transcriptional) aspect) is proposed. It is planned to analyze the real transcription of database segments (allophones, groups of allophones and syllables) together with conditions of their realization (e.g., stress, phonemic identity, syntagmatic position, context). All sound realizations may be divided then into the followings groups: 1) regular, which regularly coincide with ideal transcription (always or in determined cases); 2) not-regular, which may vary from ideal; 3) "inadequate", which practically never coincide with ideal; 4) personally specific, characterizing the particular speaker. For the each group the sets of rules (conditions) should be defined, and the variation series for each parameter may be built. Generalization of the obtained material will produce the empirical model of Russian speech.

Analyzing not-standard (regional, dialect, etc.) speech variety makes it possible to determine the boundary values of estimation parameters that divide the pronunciation into "right" (standard) and "wrong" (not-standard). Only on this stage of research language pronunciation norm may be actually estimated and the adequate model of Russian speech may be created.

The described method may be used as well for evaluation of speech in other aspects, including prosody and temporal dynamics.

3 Standard speech variability

3.1 Sound material

Research material is represented by recordings of the Phonetically Representative Text (PRT), which is the one of the central components of the Phonetic Fund of Russian Language. PRT was specially composed so that to comprise 200 the most frequent syllables of Russian. Moreover, 95

The speakers with standard pronunciation have been selected for the recordings of the PRT. The material has been recorded in specially equipped sound-treated chamber. The speakers read out the text, with which they had been beforehand acquainted.

3.2 Data analysis

The pronouncing of phoneme segments has been analyzed using auditory and acoustic methods: researchers-auditors, experts of the phonetic analysis, have been listening to the vowels both in context (syllable) and separated off the context. At their disposal there were the sonograms and oscillograms, and the broad context was known. In the result of complex analysis of each sound (phoneme), experts have determined their so-called "real" phonetic transcription [3]. The information on the actual pronouncing in the recordings of the PRT was presented in the relational database [4], which allowed to evaluate quality variation of the given speech samples. The detailed description is being made for each segment of the database. It comprises in particular the "real" transcription, "ideal" phonemic and phonetic transcriptions, attributes of the sound, duration of the signal segments, acoustic features, phonetic and paralinguistic comments.

3.3 Conclusions

The processing of connected speech revealed high variety of Russian phonemes (especially vowels) pronouncing [3, 4, 6]. In fact, standard speakers may pronounce instead of theoretically predicted (or "near") vowels a variety of different sounds, even surprisingly unlike to the prescribed ones. On the other hand individual speakers may show an inclination to the use in speech of some or other particular sounds either generally or in concrete cases. All these differences are not usually noticeable by the Russians (all analyzed recordings are being perceived as just "standard", without any distinctive features) and may be revealed only by means of computer analysis [7]. Generally, strong phrase position (under syntagmatic, logical or emphatic stress) assists pronouncing striving to an ideal one (both phonetically and phonemically). Weak phrase position on the contrary causes considerable modifications of the phonemes up to their complete falling out. In the result a number of vowels loses their own relevant features and even gains the characteristics of some other phonemes.

4 Connected speech transcription

While processing and analyzing standard speech data we have noticed that there exist essential differences between actual data and traditional transcribing tools. We have discussed this phenomenon in our recent works [4, 6]. As vowels turned out to be much more variable as compared with consonants, most "problems" of transcription related to them. First, each vowel phoneme has individual number of allophones which may not coincide with the theoretically predicted ones. Then one and the same segment (from both acoustic and perceptive points of view) may represent allophones of different phonemes, whereas quite unlike segments may represent the same phoneme. Another important factor which may modify their pronouncing is prosodic position of sounds. It is very important, that the development of transcribing tools destined for connected speech processing, speech synthesis and recognition systems should take into account these results of experimental phonetic speech analysis.

5 Regional speech variability

5.1 Sound material

Russian speech has been a subject of a many-years research program of the Department of Phonetics of Saint-Petersburg State University. A special phonetically representative text has been created, which allows analysis of regional pronunciation peculiarities. In the result of a serial of expeditions a big collection of audio recordings of this text from the different points of Russia and the former USSR republics has been formed [8]. Detailed phonetic transcription was made for all recordings. Summary regional peculiarities are presented in the special tables of phonetic parameters [9].

5.2 Data analysis

From all the multitude of recordings the most "representative" ones reflecting 20 regional variants of Russian speech are selected. For the primary analysis it is planned to investigate both "good"-Russian and "poor"-Russian recordings for each regional/national region. The selected recording are segmented into syllables, words, syntagmes, and phrases. This segmentation will be further used for the access to the regional peculiarities in the database. Tables of phonetic parameters reflecting summary regional peculiarities are being organized as a relational database, which should allow to conduct parameter search on request. Each pronunciation peculiarity (phonetic parameter) will have references to the correspondent sound samples. The expert database should allow the following: 1) on request of particular regional pronunciation the user may see the orthographic and transcriptional texts, to listen to it, and see the tables of the relevant parameters and linguistic comments; 2) on request of phonetic parameter, to find regional variants with such peculiarities and to listen and review the other database samples. The reduced version of this database is planned to be presented in the Internet.

6 Conclusion

Variability models of real speech should be necessarily taken into account for optimization of many speech technologies and products. The obtained algorithms and statistical data may be interesting as well for theoretical phonetics and linguistic studies, for Russian teaching, and especially for phonetic interference studies and forensic phonetics.

7 Acknowledgments

This work is supported by the Russian Foundation for Humanitarian Research: project N 99-04-12015b "Regional Variants of the Russian Speech in the Internet".

References

1. Bondarko L.V., Stepanova S.B. (1992) Text Module "Phonetically Representative Text". In: Bulletin of the Phonetic Fund of Russian Language. Bochum-Leningrad. N 4. 132-134.
2. Bondarko L.V., Zinder L.R., Stern A.S. (1977) Statistic Characteristics of Russian Speech. In: Hearing and Speech in Norm and Pathology. Leningrad.
3. Kuznetsov V.I. (1997) Vowels in Russian Speech. St.Petersburg State University.
4. Kuznetsov V.I. and Sherstinova T.Y. (1998) Phonetic Transcription for the Systems of Speech Synthesis and Recognition. In: Proceedings of the International Workshop "Text, Speech and Dialog" TSD'98. 22-26 September, 1998. Brno - Czech Republic. 263-267.
5. Sherstinova T.Y. (1997) Database FON2 of Russian Normative Speech. In: Proceedings of the Second International Workshop "Speech and computer" SPECOM'97. 27-30 October 1997, Cluj-Napoka - Romania. 81-84.
6. Sherstinova T.Y. (1998) Russian Phonemes and Allophones in Continuous Speech: for Speech Synthesis and Recognition Systems. In: Proceedings of the Second International Workshop "Speech and computer" SPECOM'98. 26-29 October 1998, Saint-Petersburg. 269-274.
7. Sherstinova T.Y. (1998) Speech Evaluation in Russian Phonetic Database. In: Proceedings of the First International Conference on Language resources and Evaluation. Granada, Spain, 28-30 May 1998.
8. Interference of the Sound Systems (1984). Saint-Petersburg State University.
9. Scientific Reports "Peculiarities of Russian Language Functioning in the Dialect and National Regions of the USSR (sound aspect)". Leningrad State University. Department of Phonetics. 1975-1986.