

**AN INTELLEAGENT SYSTEM OF ANALYSIS  
OF INTONATION STRUCTURES: APPLICATION  
IN TEACHING THE RUSSIAN LANGUAGE  
TO THE CHINESE LANGUAGE NATIVE SPEAKERS**

© 2018 Y.L. Beresovskaya, T.D. Isupova, D.A. Katsay, O.I. Sharafutdinova,  
L.I. Shestakova, O.B. Elagina

*South Ural State University*

*(pr. Lenina 76, Chelyabinsk, 454080 Russia)*

*E-mail: berezovskaiaail@susu.ru, isupovtd@list.ru, katcaida@susu.ru,  
sharafutdinovaoi@susu.ru, shestakovali@susu.ru, elaginaob@susu.ru*

Received: 24.10.2018

This paper describes the development of a program for analysis of intoning of verbal pieces in the Russian language. The goal is to measure the differences between the intoning of verbal pieces by both native and international Russian language speakers. The research methodology is based on the application of neural network analysis for solving the task of identification of speech samples, obtained by recording inophones' speech. The experiment was carried out with the participation of 12 people: native speakers of the Russian language and the Chinese language, both male and female, aged from 20 to 35. A total number of speech samples amounted to 4800 items. Overall, 10 speech items in declarative and interrogative intonation were analyzed. A neural network that provides an assessment of correspondence of a speech sample to the standard variant of intoning was formed and trained. The results of experimental research are presented in the form of statistical assessments of pronouncing the verbal pieces with various intonations. These results are recommended to be applied in the process of learning Russian as a foreign language: the obtained data are considered as the confidence threshold of intoning identification, which complies with the standard or deviates from it. The results can also be applied for the individualized automated compilation of recommendations on correction of mistakes.

*Keywords: cepstral analysis, neural network, intoning component.*

#### **FOR CITATION**

Beresovskaya Y.L., Isupova T.D., Katsay D.A., Sharafutdinova O.I., Shestakova L.I., Elagina O.B. An Intelligent System of Analysis of Intonation Structures: Application in Teaching the Russian Language to the Chinese Language Native Speakers. *Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering*. 2018. vol. 7, no. 4. pp. 105–121. (in Russian) DOI: 10.14529/cmse180407.

#### **Introduction**

Modern linguistics and lingoudidactics set out a number of tasks to researchers and methodologists: communicative and pragmatic approach in teaching, individualization and national orientation of teaching, optimization of balance between classroom and extracurricular (independent) work of students. The formation of communicative competence and then its control in the oral types of speech activities are connected with the pronouncing skills. We consider the sound, phonetic aspect of speech to be the basis for the communicative competence realization in oral types of speech activities since errors in pronunciation and general phonetic illegibility of speech lead to communicative errors which, in turn, can cause communication failures.

Analysis of speech is performed by various methods among which there can be distinguished algorithmic methods and methods of neural networks. Both approaches are based on

the transformation from the time domain to the frequency domain with the help of Fourier transform. The next stage of analysis is to study properties of the enveloping frequency characteristic of the studied speech sample. Popular methods of formant and cepstral analysis are used for these purposes. A great influence on the quality of decomposition of the voice signal and assessment of the speaker's voice path is made by the procedure of splitting the time window, which occupies the speech sample [1]. Improving the efficiency of the cepstral analysis method can be achieved by optimizing the parameters of the cepstral vector by the criterion of minimum cross-connections between speech samples [2]. However, algorithmic methods are limited in their capabilities due to the high variability of the parameters of speech samples models caused by the peculiarities of the human articulatory organs.

Intonation is an important aspect of practical teaching phonetics. In the context of globalization and optimization of learning processes, there is needed a system considering specific character of the native language intonation (ethnomethodological approach) and individual characteristics of a student. In order to develop such a system and increase the use of dialogues in the process of teaching a language, it is necessary to measure the differences between the intoning of individual speech segments by the Russian language native speakers and by the Chinese language native speakers studying the Russian language; determine the influence of the intonation system of the native language on the studied language and systemize the pronouncing variants interpreted as close as possible to the standard invariant of pronunciation.

Analysis of intoning in the speech of inophones and account of specific characteristics of students' native language allows avoiding violations of rhythmic pattern, incorrect emphasis of the intonation structure center, replacement of the connotative coloring of a word or a phrase while teaching Russian as a foreign language to Chinese language native speakers.

This paper is organized as follows. Section 1 is devoted to the method of analysis of intonation structures using a neural network. In section 2, we present the results of learning of the neural network and its testing by the Chinese language native speakers. Conclusion summarizes the study and points directions for further work.

## **1. Methodology**

Methods of teaching phonetics are based on fundamental achievements in the field of phonetics description as a subsystem of language. For the first time, the question of phoneme functional understanding and psychophysical basis of phoneme perception was considered in the works by I.A. Baudouin de Courtenay [3], F. de Saussure [4], V.A. Bogoroditsky [5].

Modern phonology uses a phonemic-cluster method to solve practical tasks of the Russian language phonetics teaching [6]. In this aspect the phonemic-cluster approach meets the principles of clarity and simplicity which allow developing programs for processing speech samples for the speakers of different languages. The current research presents the experience of developing an electronic training system based on experimental data obtained during the analysis of speech samples of Chinese language native speakers. In addition, different aspects of phonetics are taken into account such as articulatory, acoustic, perceptive and functional aspects [7].

In the modern science of language, there are distinguished several concepts in intonation learning. E.A. Bryzgunova [8] having proposed to describe the intonation of the Russian language with the help of "intonation structure" notion, provided methodologists with an important tool for teaching phonetics (cf., the notion of "intoneme" by V.A. Artemov [9]). Ac-

According to the researcher, intonation is a combination of tone movement, sound strength, timbre, and duration [10]. Intonation is one of the most important parameters of speech abilities, while studying a language the multifunctionality of intonation in the Russian language makes mastering the laws and rules of intonation a necessary condition for mastering speech as a tool for expressing different shades of a thought and feelings. N.N. Rogoznaya defines intonation as a functional-semantic of macrounit of an utterance which includes pitch frequency, intensity, duration, timbre, and so on; intonation is being described in terms of acoustics and responds to objective (instrumental) analysis [11]. The work on intoning becomes especially important during the study of non-native speech which is primarily due to the difference in the intonation pattern of various structures of the studied languages – native and foreign languages, which, by definition, generates various kinds of errors of speech coding and decoding by inophones while speaking a non-native language. Being a distinctive means of language intonation performs a pragmatic function. Moreover, as it is proved experimentally by L.V. Bondarko, N.B. Volskaya and their colleagues [12], the intonation of spontaneous speech differs from the reproduction of a ready text, what should be taken into account when teaching the generation and perception of oral speech.

Modern linguodidactics is developing national-oriented approaches to education. Taking into account the facts of the native language that affect the mastery of a foreign language is very productive, as noted in the works of researchers and methodologists dealing with the problems of teaching languages of different types (see, for example, N.N. Rogoznaya [13], V.A.V. Shafiro, Kharkhurin [14], S.T. Best [15], L. Wade-Woolley [16], R.S. Panova [17], Zhao Zhe [18]). A number of works are devoted to the problem of teaching the intonation of the Russian language in relation and correlation with the system of the native language or another studied language (for example, I.I. Trubchaninova [19], L.Z. Mazina [20]). This ethnomethodological approach is taken into account in our research in work with the Chinese audience in terms of reflection of Chinese phonetics peculiarities in inophones' Russian speech. The problem of recognition of intonation structures in the Russian speech of Chinese speakers is connected with the issue of differentiation of intonation and tone in the Chinese language. M.K. Rumyantsev distinguishes such a semantically distinguishing function of tone for identification of words and morphemes, while intonation, according to the scientist, represents a change in a pitch over a large segment of an utterance [21].

There are known technologies for the use of neural networks (NN) for speech recognition. In work [22] it is noted an expediency of usage of feedforward neural network the input layer of which contains such a number of neurons which corresponds to the number of analyzed features. Cepstral vectors as the most informative descriptions of speech samples in a frequency domain are used as input for the NN. Fig. 1 shows a diagram of an intelligent system for intonation sample analysis which includes NN as an integral part.

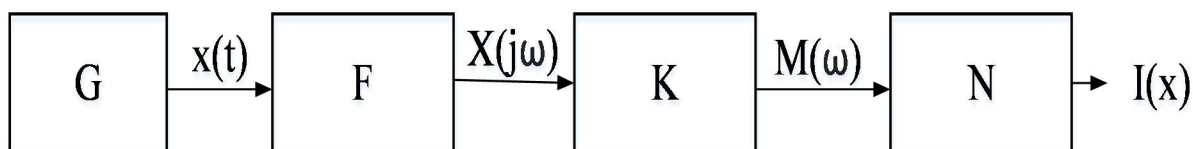


Fig. 1. Diagram of intellegent system of analysis

The speech sample  $x(t)$  is formed by a recording device G. In block F the speech sample is transformed into a frequency domain, which in the form of  $X(j\omega)$  enters block K of  $M(\omega)$  the composite cepstral vector formation. For each speech sample a composite cepstral vector of  $m$  dimension is formed of successively connected cepstral vectors calculated for a single frame,

into which a speech sample is divided. A number of parameters of a cepstral vector of a single frame can be selected by the criterion of the maximum correlation of an intonation image which is formed in it together with all speech samples which make up their general totality. A preliminary analysis of intonation samples showed that the minimum number of  $n_k$  parameters of a spectral vector of a single frame should be at least 30. The connection of dimension of a composite cepstral vector with the dimension of a cepstral vector of  $m = n_t \cdot n_k$  the individual frame, where  $n_t$  is a number of frames. The singularity of the research is that such separate words as “дом” (house), “кот” (cat), “магазин” (shop), “мама” (mother), “семья” (family), “сок” (juice), “документ” (document), “комната” (room), “собака” (dog), “сумка” (bag) were selected as intonation samples. Any speech samples are limited in duration of their pronunciation. This makes it possible to perform their normalization by a number of  $n_t$  time intervals, on which they will be divided before conversion to a frequency domain. The operation of normalization of cepstral vectors provides with the formatting of input for NN what is a necessary condition for its operation.

$M(\omega)$  the composite cepstral vector enters N feedforward neural network, the diagram of which is shown on Fig.2.

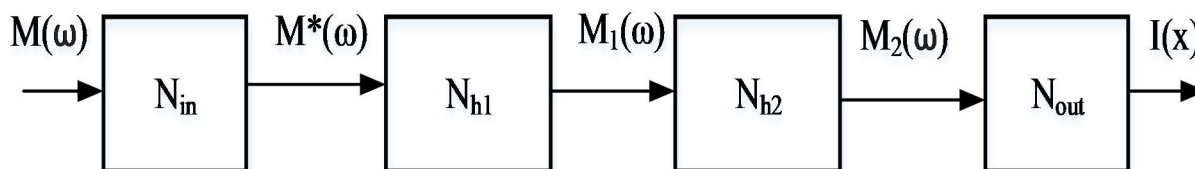


Fig. 2. Diagram of neural network

The NN parameters are presented in Tab. 1. There are 240 neurons in the input layer. Their number is determined by the number of cepstral coefficients. The NN contains two internal layers: in the first there are 50 neurons, in the second – 2 neurons. Output of network forms  $I(x)$  the two-component vector. The first component speaks for declarative intonation of a speech sample; the second component speaks for interrogative intonation.

Table 1

NN Characteristic

Designation of a layer	Number of neurons in a layer	Characteristics of a layer
$N_{in}$ the input layer	240 neurons	Complies with the dimensions of input vector
$N_{h1}$ the autoencoder	50 neurons	Learns with a decoder without a teacher by minimum mean-square error with the L2 metric and adjustable sparsity of NN
$N_{h2}$ the classifier	2 neurons	Learns with a vector-teacher after network assembly using method of scalable conjugate gradients
$N_{out}$ the output	2 neurons	Type of intonation: [1 0] – for declarative intonation and [0 1] – for interrogative intonation

## 2. Results

### 2.1. Errors related to the identification of speech samples

Specific feature of the Chinese language native speakers consists in continuous changing of tone frequency during the whole utterance as well as a single lexical unit that is confirmed by the results of experiment of recording inophones' speech, both male and female (cf, for example, changing the pitch, a shift of lexical units intonation centers in interrogative sentences in such words as “документ” (document), “собака” (dog), “мама” (mother)). The errors connected with the emphasis of the last syllable in a lexical unit are due to the fact that the Chinese interrogative sentence contains \*ma particle at the end of the sentence differentiating a question. It is the interrogative particle that most often accounts for the melodic peak of the utterance in the Chinese language.

While identifying declarative and interrogative constructions the difficulty consists in the pitch including such components as the height of the starting point, interval between the starting and the ending points, pauses, a stress (cf., for example, in experimental base an interrogative intonation if lexical unit “дом” (house) pronounced by the Chinese language native speakers and compared to the third tone in the Chinese language; the vowel is being stretched and, consequently, there are a few melodic peaks).

### 2.2. The role of technical means in teaching russian phonetics

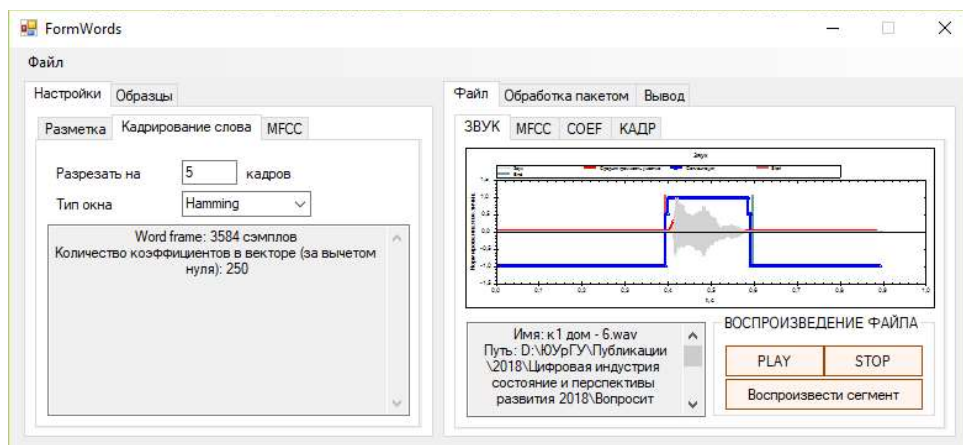
Experts associate the optimization of the formation of pronunciation skills with the remote and electronic educational technologies. The electronic educational environment provides many opportunities for modern linguodidactics, therefore it is the subject of consideration both in terms of technical capabilities [23] and in terms of difficulties and problems of language teaching. In particular, the issue of implementing a competence-based approach to language teaching is being developed (see, for example, [24]).

The popularity of electronic educational resources and a large number of them do not solve the problems caused by the specific features of teaching phonetics. The main problem is the dialogization of learning: the majority of educational resources give the opportunity to listen to the studied sound or intonation structure and record your own version of pronunciation. However, at this stage, there is a problem of measuring the degree of compliance of a particular pronunciation with the variant which is included in the set of standard variants making the phonemic invariant.

### 2.3. Results of the experimental research

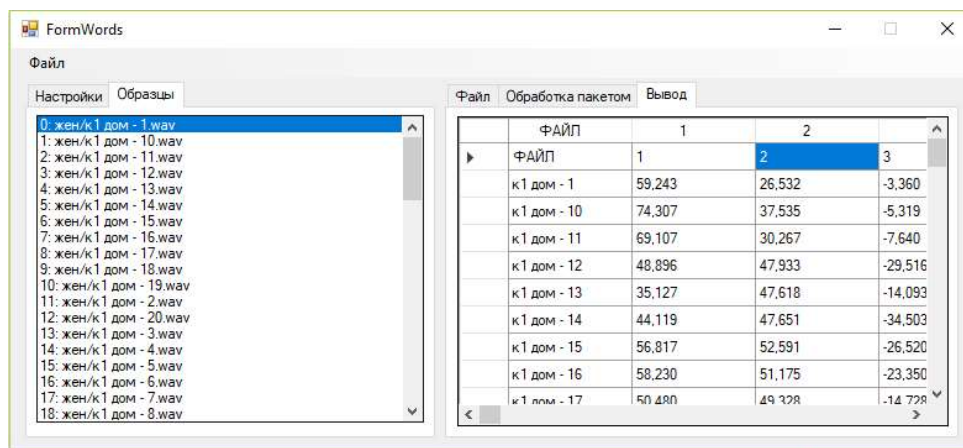
Preparation of cepstral vectors was carried out in a program that allows adjusting the parameters of the cepstral vectors under formation. Some of the program dialog windows are shown in Fig. 3–6.

Fig. 3 shows the setting the number of frames of the speech sample. In the presented analysis it is 5 frames. To reduce the errors of the Fourier transform on the borders of the windows applied Hamming window. The right side of Fig. 3 shows the graph of the speech pattern in the time domain. The informative part of the speech sample is automatically highlighted in blue, according to which the analysis will be performed. With the help of the built-in player, it is possible to play the entire file or its informative part. It is possible to go to the window “Mel Frequency Cepstral Coefficients” (MFCC).



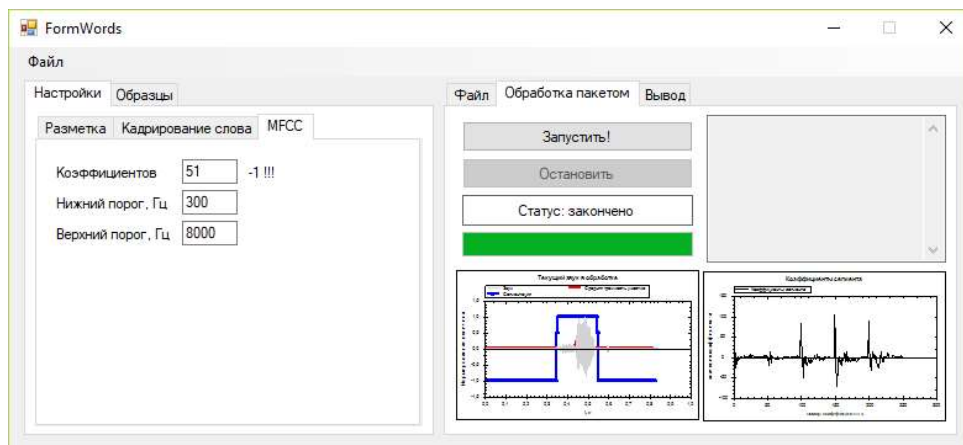
**Fig. 3.** MFCC Calculation Program's Dialog Window: setting the number of frames of the speech sample

Fig. 4. shows MFCC options window. The program performs batch processing of the speech samples files. MFCC vector is formed for each file as a sequence of MFCC vectors extracted from successive frames of the speech signal.



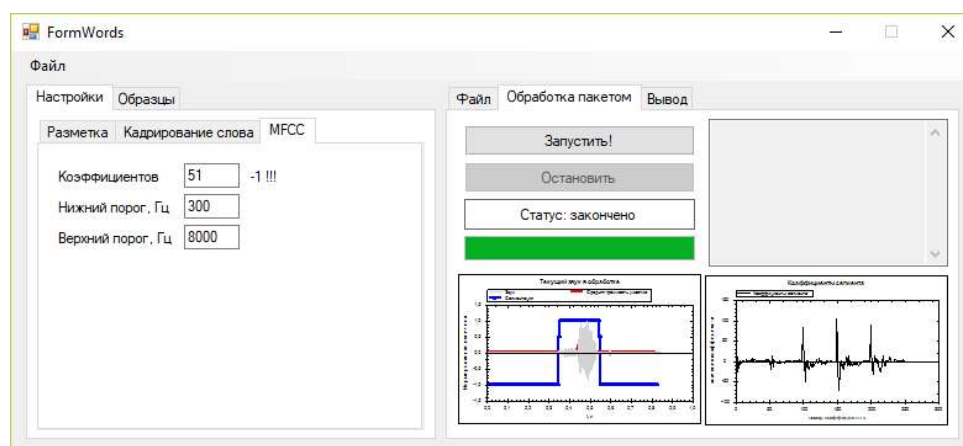
**Fig. 4.** MFCC Calculation Program's Dialog Window: MFCC vectors extraction

Fig. 5. shows the adjustment of MFCC extraction parameters. Here one can set up the parameters of MFCC extraction such as coefficients number or the herz-scale size.



**Fig. 5.** MFCC Calculation Program's Dialog Window: MFCC options and a batch mode panel

Fig. 6 shows the results of cepstral vectors calculation in a batch mode.



**Fig. 6.** MFCC Calculation Program's Dialog Window:  
results of calculation of cepstral vectors in a batch mode

The neural network was trained on an array of speech samples the characteristic of which is presented in Tab. 2.

**Table 2**

Distribution of the speech samples by type of intonation and native speakers

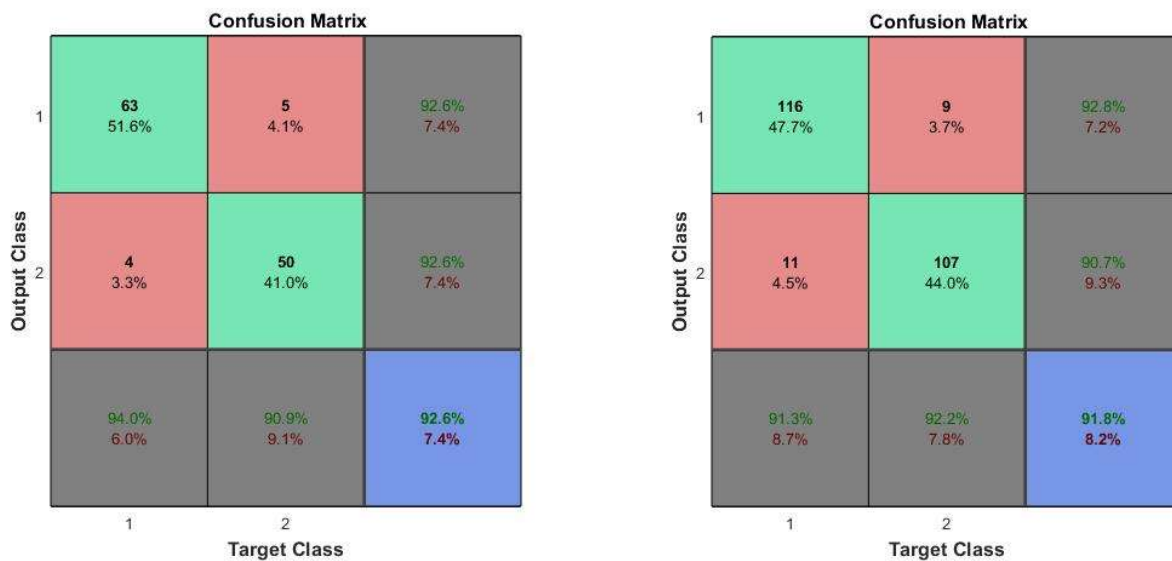
Language speakers	Intonation	Number of speech samples	
		<i>by type of intonation</i>	<i>by type of speakers</i>
Russian (teaching)	Interrogative	1229	2308
	Declarative	1079	
Chinese	Interrogative	509	903
	Declarative	394	
Russian (5% test)	Interrogative	60	120
	Declarative	60	

The total number of samples in the table makes up 3331 units out of the 4800 initial speech samples. The decrease of the samples quantity is caused by their sifting in the course of inspection for their compliance with the required intonation.

Fig. 7 shows the confusion matrix for a neural network with a single-layer autoencoder (240-50-2): a – calculated by 5% test samples; b – calculated by 10% test samples.

Fig. 8 shows the confusion matrix of a neural network with a single-layer autoencoder (240-50-2) for the Chinese language native speakers. The samples amount made up 903 units.

Fig. 9 shows the confusion matrix for a neural network with a two-layer autoencoder containing 10 neurons in the second layer (240-50-10-2): a – calculated by 5% test samples; b – calculated by 10% test samples.



a) Calculated by 5% test samples

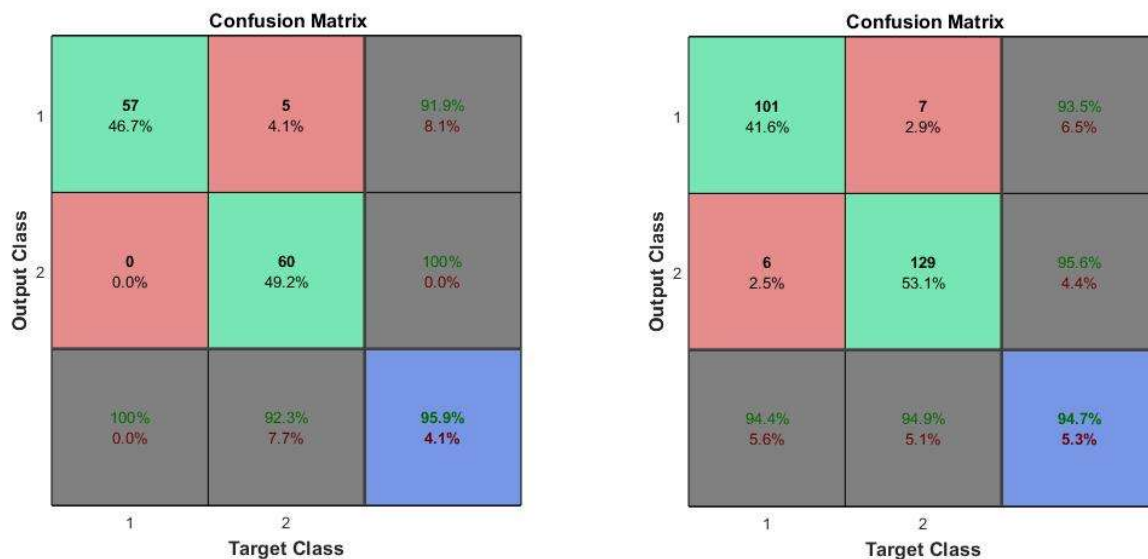
b) Calculated by 10% test samples

**Fig. 7.** Confusion matrices by test samples for the neural network with a single-layer autoencoder (240-50-2)



**Fig. 8.** Confusion matrix of the neural network with a single-layer autoencoder for the Chinese language native speakers (240-50-2)



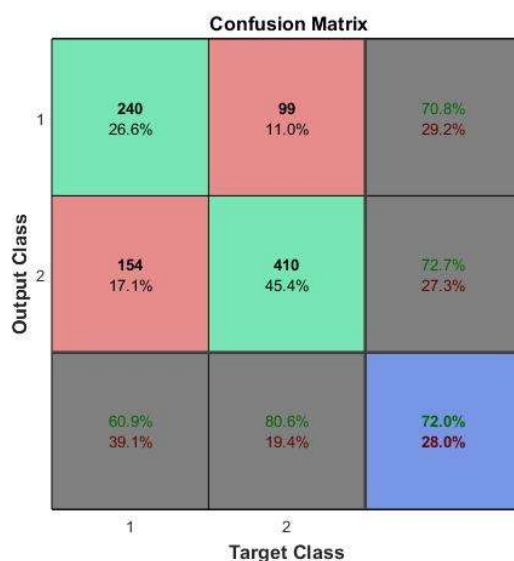


a) Calculated by 5% test samples

b) Calculated by 10% test samples

**Fig. 9.** Confusion matrices for the neural network with a two-layer autoencoder (240-50-10-2)

Fig. 10 shows the confusion matrix of a neural network with a two-layer autoencoder (240-50-10-2) for the Chinese language native speakers.



**Fig. 10.** Confusion matrix for the neural network with a two-layer autoencoder (240-50-10-2) for the Chinese language native speakers

Fig. 11 shows the diagram “Minimum Mean-Square Error with the L2 metric and adjustable sparsity of NN” depending on the quantity of epochs of learning for the neural network with a two-layer autoencoder (240-50-10-2).

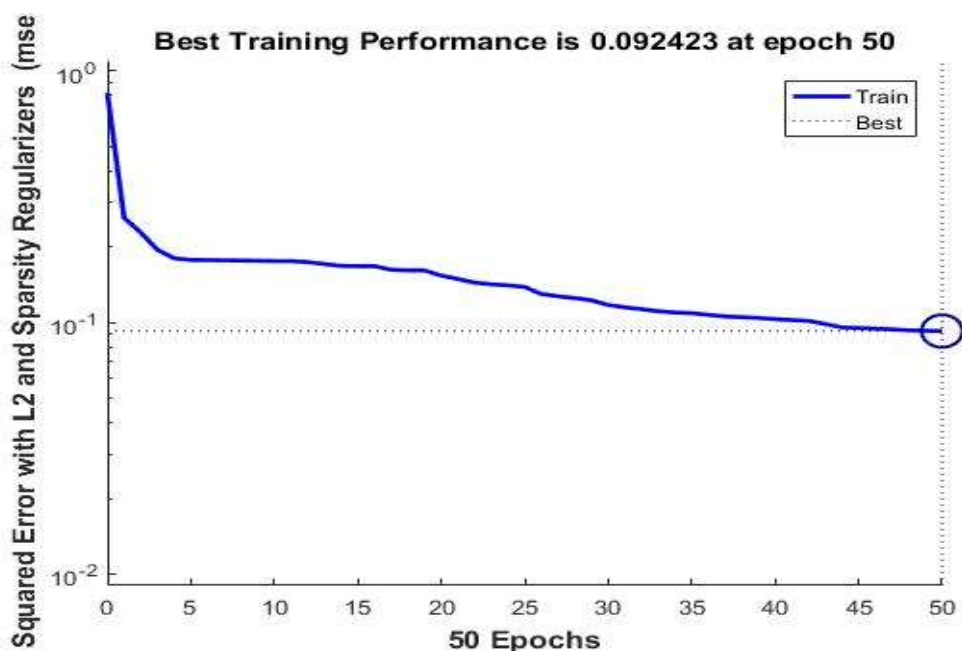
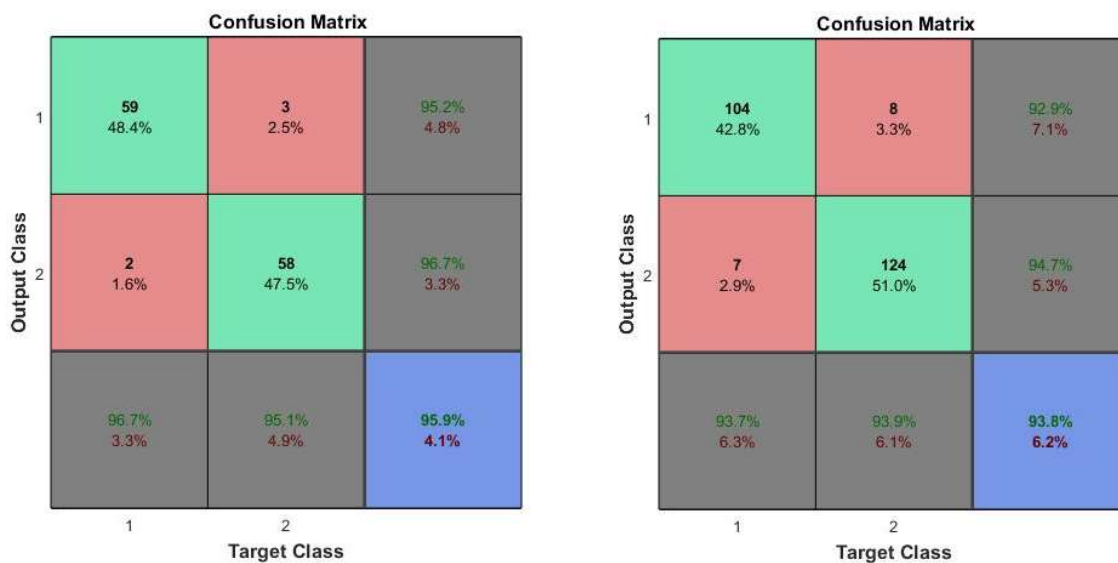


Fig. 11. Diagram of learning of the neural network with a two-layer autoencoder (240-50-10-2)

Fig. 12 shows the confusion matrix for a neural network with a two-layer autoencoder containing 100 neurons in the first layer (240-100-10-2).

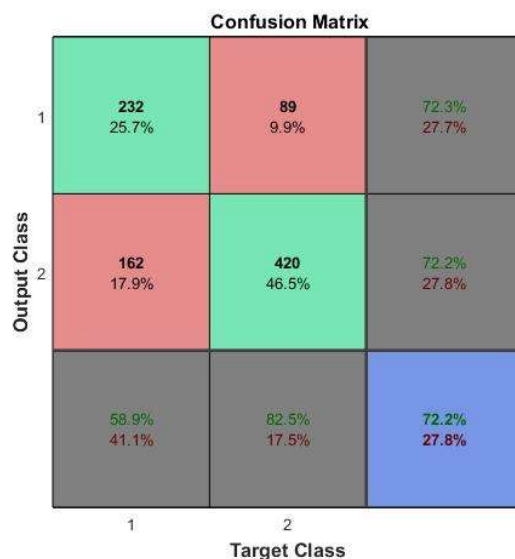


a) Calculated by 5% test samples

b) Calculated by 10% test samples

Fig. 12. Confusion matrices for the neural network with a two-layer autoencoder (240-100-10-2)

Fig. 13 shows the confusion matrix of the neural network with a two-layer autoencoder (240-100-10-2) for the Chinese language native speakers.



**Fig. 13.** Confusion matrix for the neural network with a two-layer autoencoder (240-100-10-2) for the Chinese language native speakers

A number of cepstral coefficients are calculated with account of zero coefficient removal as less informative for the analysis of speech sample intonation. There were formed two variants of data with the parameters: 1)  $nt = 5$  frames,  $nk = 50$  coefficients,  $m = 250$ ; 2)  $nt = 8$  frames,  $nk = 30$  coefficients,  $m = 240$ . The second variant with 8 frames of 30 coefficients proved to be better in accuracy of speech intonation recognition. The results of intonation recognition in the speech of the Chinese language native speakers are presented in Tab. 3.

**Table 3**

Evaluation of recognition of the chinese language native speakers' intonation

Number of speech samples and their percentage		Speech sample					
		<i>Declarative</i>		<i>Interrogative</i>		<i>Total, %</i>	
		<i>Correct</i>	<i>Incorrect</i>	<i>Correct</i>	<i>Incorrect</i>	<i>Correct</i>	<i>Incorrect</i>
Results of recognition	Declarative	232			89		
	Interrogative		162	420			
	Total, %	58.9	41.1	82.5	17.5	72.2	27.8

Comparison of NN variants used to determine the Russian language native speakers' intonation, whose properties are presented by the error matrices in Fig. 7a and Fig. 9a by Van Rijsbergen F-test: with a two-layer autoencoder,  $F = 0.959$  is slightly better compared to a single-layer autoencoder  $F = 0.926$ .

Comparison of NN variants used to determine the Chinese language native speakers' intonation, whose properties are presented by the error matrices in Fig.8 and Fig.10 by F-test: with a single-layer autoencoder,  $F = 0.73$  is slightly better compared to a two-layer autoencoder  $F = 0.712$ .

## Conclusion

The goal of the research was to measure differences between the intoning of verbal pieces by both native and international Russian language speakers. By using neural network methods, differences in recognizing intonation errors were identified for native Russian and non-native Chinese subjects learning the Russian language. Overall, the neural network method successfully measured the relative difference between the two experimental groups.

The intonation recognition among the Russian language native speakers ranged from 92 to 97% on the test array of samples which makes 10 to 5% from the general totality of intonation speech master-samples. Recognition errors are evenly distributed in declarative and interrogative intonations.

Recognition of intonation of the Chinese language native speakers is consistently recognized with probability from 70% to 73%. In the selection of speech samples of the Chinese language native speakers about 35–40% of declarative intonations are mistakenly identified as interrogative and less than 20% of interrogative intonations are identified as declarative. The possible reason can lie in the specific features of the Chinese language native speakers when the declarative intonation sounds like an interrogative one.

The research revealed that short monosyllabic words such as “кот” (cat), “дом” (house), “сок” (juice) do not contribute to the verification of data as an objective analysis of intonation can give a correct assessment in the presence of more than one syllable in a speech unit.

Due to the limitation of the teaching selection the dimension of NN was determined to be small both in the number of hidden layers and the number of neurons.

In future research work, it is planned to expand the selection of intonation samples used for teaching NN by age criterion, by number and type of samples in the form of complex words and sentences which are more true-to-fact for observing the intonation characteristics.

*The work was being completed under financial support of the Ministry of Education and Science of the Russian Federation on the project entitled «Establishment and development of a network (not less than 8) of Pushkin Institute Centers in the PRC on the basis of organizations performing education in Russian language» within the frameworks of implementation of the event entitled “Subsidy for implementation of events targeted at integral functioning and development of Russian language” of the main event entitled “Development of open education in Russian language and study of Russian language” of the field (subprogram) entitled “Development and distribution of Russian language as a foundation of civil self-identity and the language of international dialogue (“Russian language”)” of the Education Development state program of the Russian Federation, according to the Agreement between the Ministry of Education and Science of the Russian Federation and the Federal State Autonomous Educational Institution of Higher Education “South Ural State University (National Research University)”.*

*This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 3.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.*

## References

1. Drugman T., Bozkurt B., Dutoit T. Complex Cepstrum-based Decomposition of Speech for Glottal Source Estimation. INTERSPEECH, 10th Annual Conference of the International

- Speech Communication Association (Brighton, United Kingdom, September 6–10). ISCA, 2009. pp. 116–119.
2. Katsay D., Doronina E., Kazakova Y., Kharchenko E., Isupova T. Analysis of Speech Samples as a Means of Intensification of Teaching Phonetics of a Foreign Language. *ICERI2017 Proceedings: 10th International Conference of Education, Research and Innovation (Seville, Spain, November 16–18)*. IATED, 2017. pp. 2771–2778. DOI: 10.21125/ic-eri.2017.0789.
  3. Boduen de Kurtene I.A. *Izbrannye trudy po obshchemu yazykoznaniyu* [Selected Works on General Linguistics]. Moscow, Publishing of the Academy of Sciences of the USSR, 1963. vol. 1. 391 p. (in Russian)
  4. F. de Saussure. *Cours de Linguistique Generale*. Paris, Albert Sechehaye et Albert Riedlinger, Payot, 1971. 318 p.
  5. Bogorodickij V.A. *Kurs éksperimental'noj fonetiki primenitel'no k literaturnomu russkomu proiznoseniju* [A Course of Experimental Phonetics Applied to Literary Russian Pronunciation]. Kazan', Tip. Imp. un-ta, 1917. 74 p. (in Russian)
  6. Rogoznaya N.N. A Project of Automated Learning of Non-native Phonology. *Vestnik Buryatskogo gosudarstvennogo univesiteta* [Bulletin of Buryat State University]. 2017. vol. 3. pp. 87–94. (in Russian)
  7. Bondarko L.V. *Fonetika sovremennogo russkogo jazyka* [Phonetics of Russian Modern Language]. St.Petersburg, Publishing of SPbSU, 1998. 276 p. (in Russian)
  8. Bryzgunova E.A. *Prakticheskaja fonetika i intonacija russkogo jazyka* [Practical Phonetics and Intonation of the Russian Language]. Moscow, Publishing of MSU, 1963. 308 p. (in Russian)
  9. Artemov V.A. On Intoneme and Intonational version. *Intonazia i zvukovoi sostav: Sbornik nauchnyh trudov* [Intonation and Sound Composition: Colloquium Materials on Experimental Phonetics and Speech Psychology]. 1965. pp. 3–20. (in Russian)
  10. Bryzgunova E.A. Intonation. *Russkaya grammatika* [Russian Grammar]. 1980. vol. 1. pp. 96–122. (in Russian)
  11. Rogoznaya N.N. Synharmonism Influence upon the Forming of the A2. *Vestnik Buryatskogo gosudarstvennogo univesiteta* [Bulletin of Buryat State University]. 2010. vol. 10. pp. 74–79. (in Russian)
  12. Bondarko L.V., Volskaya N.B., Tananaiko S.O., Vasilieva L.A. Phonetic Properties of Russian Spontaneous Speech. *15 ICPHS Barcelona*. 2003. pp. 2973–2976. Available at: <https://pdfs.semanticscholar.org/6b1b/c7e6d5abcea5b5c7a50bf923b8d2c339bda2.pdf> (accessed: 10.08.2018).
  13. Rogoznaya N.N. *Bilingvizm. Interyazyk. Interferenciya* [Bilingualism. Interlanguage. Interference]. Irkutsk, Publishing of IrSU, 2012. 171 p. (in Russian)
  14. Shafiro V., Kharkhurin A.V. The Role of Native-Language Phonology in the Auditory Word Identification and Visual Word Recognition of Russian-English Bilinguals. *Journal of Psycholinguistic Research*. 2009. vol. 38, issue 2. pp. 93–110. DOI: 10.1007/s10936-008-9086-y.
  15. Best C.T. A Direct-Realist View of Cross-Language Speech Perception. *Speech Perception and Linguistic Experience*, W. Strange (Ed.). York Click, 1995. pp. 171–204.

16. Wade-Woolley L. First Language Influences on Second Language Word Reading: All Roads Lead to Rome. *Language Learning*, 1999, vol. 49, no.3. pp. 447–471. DOI: 10.1111/0023-8333.00096.s
17. Panova R.S. Phonetic Interference in the Russian Speech of the Chinese. *Vestnik Chelyabinskogo gosudarstvennogo universiteta. Filologiya. Iskusstvovedenie* [Bulletin of Chelyabinsk State University, Philology. History of Art]. 2009. vol. 22. pp. 231–233. (in Russian)
18. Chzao Chze. Sound Interference in the Russian Language under the Influence of the Native Language in the Context of Russian-Chinese Language Contacts. *Filologicheskie nauki. Voprosy teorii i praktiki* [Philological Sciences. Questions of Theory and Practice]. 2016. vol. 12, no. 3. pp. 179–184. (in Russian)
19. Trubchaninova I.I. *Lingvodidakticheskie osnovy obucheniya intonacii russkogo yazyka v usloviyah uchebnogo trilingvizma* [Linguodidactic Bases for Teaching the Intonation of the Russian Language in the Conditions of Educational Trilingualism]. Vladikavkaz, Publishing of DSU, 2012. 194 p. (in Russian)
20. Mazina L.Z. *Metodika obucheniya studentov-inostrancev intonacii russkogo yazyka (nachal'nyj ehtap kontakta ispanskogo yazyka s russkim)* [Methods of Teaching Foreign Students Intonation of the Russian Language (the Initial Stage of Contact of the Spanish Language with the Russian Language)]. Moscow, 1984. 257 p. (in Russian)
21. Rumyansev M. Synthesis of Chinese Tones. *Voprosy lingvistiki* [Questions of Linguistics] 1988. vol. 1. pp. 82–93. (in Russian)
22. Sorokoumova D., Korelin O., Sorokoumov A. Development and Training of the Neural Network for Voice Recognition Solution. *Transactions of NNSTU n.a. R.E. Alekseyev* 2015. no. 3 (110). pp. 77–84. (in Russian)
23. Rumyanseva N.M., Garcova D.A., Kuzhakov V.E. Electronic Means of Training as an Effective Instrument of Intensification of Studying Chinese Students of Russian Phonetics (Ethnomethic Aspect). *Scientific Research – 2016: Proceedings of Articles the International Scientific Conference (Chezh Republic, Karlovy Vary-Russia, 29–30 September 2016)*. Kirov, 2016. pp. 175–183. (in Russian)
24. Gabdrahmanova P.L., Bogatova E.N., Mustafina L.R. Development of Communicative Competence in the Online Learning Environment of Russian as a Foreign Language: Opportunities and Prospects. *Educational Technologies and Society*. 2017. vol. 2. pp. 329–345

**ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА АНАЛИЗА  
ИНТОНАЦИОННЫХ КОНСТРУКЦИЙ:  
ПРИМЕНЕНИЕ В ОБУЧЕНИИ РУССКОМУ ЯЗЫКУ  
НОСИТЕЛЕЙ КИТАЙСКОГО ЯЗЫКА**

© 2018 Я.Л. Березовская, Т.Д. Исупова, Д.А. Кацай, О.И. Шарафутдинова,  
Л.И. Шестакова, О.Б. Елагина

*Южно-Уральский государственный университет*

*(454080 Челябинск, пр. им. В.И. Ленина, д. 76)*

*E-mail: berezovskaiail@susu.ru, isupovtd@list.ru, katcaida@susu.ru,  
sharafutdinovaoi@susu.ru, shestakovali@susu.ru, elaginaob@susu.ru*

Поступила в редакцию: 24.10.2018

В статье описывается разработка программы анализа интонирования речевых отрезков в русском языке. Поставлена цель измерить различия между интонированием речевых отрезков носителями русского языка как родного и как иностранного. Методология исследования основана на применении методов нейросетевого анализа для решения задачи идентификации речевых образцов, полученных в результате записи речи инофонетиков. В эксперименте принимало участие 12 человек — носителей русского и китайского языков, мужского и женского пола, в возрасте от 20 до 35 лет. Общее количество речевых образцов составило 4800 единиц. Всего проанализировано 10 единиц в повествовательной и вопросительной интонации. Сформирована и обучена нейронная сеть, дающая оценку по соответствию речевого образца эталонному варианту интонирования. Результаты экспериментальных исследований демонстрируются в виде статистических оценок произношения речевых отрезков с различной интонацией. Рекомендуется применять эти результаты в процессе обучения русскому языку как иностранному: полученные данные рассматриваются в качестве доверительного порога распознавания интонирования, соответствующего эталону или отклоняющегося от него. Также результаты можно применять для автоматического индивидуализированного подбора рекомендаций по коррективке ошибок.

*Ключевые слова: кепстральный анализ, нейронная сеть, анализ интонации, русский язык как иностранный, интонирование.*

### **ОБРАЗЕЦ ЦИТИРОВАНИЯ**

Beresovskaya Y.L., Isupova T.D., Katsay D.A., Sharafutdinova O.I., Shestakova L.I., Elagina O.B. An Intelligent System of Analysis of Intonation Structures: Application in Teaching the Russian Language to the Chinese Language Native Speakers // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2018. Т. 7, № 4. С. 105–121. DOI: 10.14529/cmse180407.

### **Литература**

1. Drugman T., Bozkurt B., Dutoit T. Complex Cepstrum-based Decomposition of Speech for Glottal Source Estimation // INTERSPEECH, 10th Annual Conference of the International Speech Communication Association, (Brighton, United Kingdom, September 6–10). ISCA, 2009. P. 116–119.
2. Katsay D., Doronina E., Kazakova Y., Kharchenko E., Isupova T. Analysis of Speech Samples as a Means of Intensification of Teaching Phonetics of a Foreign Language // ICERI2017 Proceedings: 10th International Conference of Education, Research and Innovation (Seville, Spain, November 16–18). IATED, 2017. P. 2771–2778. DOI: 10.21125/ic-eri.2017.0789.

3. Бодуэн де Куртене И.А. Избранный труды по общему языкознанию. М.: Издательство Академии наук СССР, 1963. 391 с.
4. de Saussure F. Cours de Linguistique Generale. Paris, Albert Sechehaye et Albert Riedlinger, Payot, 1971. 318 p.
5. Богородицкий В.А. Курс экспериментальной фонетики применительно к литературному русскому произношению. Казань: Тип. Имп. ун-та, 1917. 74 с.
6. Рогозная Н.Н. Проект автоматизированного обучения фонологии неродного языка // Вестник Бурятского государственного университета. 2017. №3. С. 87–94.
7. Бондарко Л.В. Фонетика современного русского языка. СПб.: Издательство СПбГУ, 1998. 276 с.
8. Брызгунова Е.А. Практическая фонетика и интонация русского языка. М.: Издательство МГУ, 1963. 308 с.
9. Артемов В.А. Об интоне и интонационном варианте // Интонация и звуковой состав: сборник научных трудов. М.: Издательство МГУ, 1965. С. 3–20.
10. Брызгунова Е.А. Интонация // Русская грамматика. Ч. 1. М.: Наука, 1980. С. 96–122.
11. Рогозная Н.Н. Влияние сингармонизма на формирование фонологического уровня А 2 // Вестник Бурятского государственного университета. 2010. №10. С. 74–79.
12. Bondarko L.V., Volskaya N.B., Tananaiko S.O., Vasilieva L.A. Phonetic Properties of Russian Spontaneous Speech // 15 ICPHS Barselona. 2003. P. 2973–2976. URL: <https://pdfs.semanticscholar.org/6b1b/c7e6d5abcea5b5c7a50bf923b8d2c339bda2.pdf> (дата обращения: 10.08.2018).
13. Рогозная Н.Н. Билингвизм. Интерязык. Интерференция. Иркутск: Издательство ИрГУ, 2012. 171 с.
14. Shafiro V., Kharkhurin A.V. The Role of Native-Language Phonology in the Auditory Word Identification and Visual Word Recognition of Russian-English Bilinguals // Journal of Psycholinguistic Research. 2009. Vol. 38, No. 2. P. 93–110. DOI: 10.1007/s10936-008-9086-y.
15. Best C.T. A Direct-Realist View of Cross-Language Speech Perception // Speech Perception and Linguistic Experience / W. Strange (ed.). York Click, 1995. P. 171–204.
16. Wade-Woolley L. First Language Influences on Second Language Word Reading: All Roads Lead to Rome // Language Learning. 1999. Vol. 49, No.3. P. 447–471. DOI: 10.1111/0023-8333.00096.
17. Панова Р.С. Фонетическая интерференция в русской речи китайцев // Вестник Челябинского государственного университета. Филология. Искусствоведение. 2009. № 22. С. 231–233.
18. Чжэ Ч. Звуковая интерференция в русском языке под влиянием родного языка в условиях русско-китайских языковых контактов // Филологические науки. Вопросы теории и практики. 2016. № 3. С. 179–184.
19. Трубочанинова И.И. Лингводидактические основы обучения интонации русского языка в условиях учебного трилингвизма. Владикавказ, 2012. 194 с.
20. Мазина Л.З. Методика обучения студентов-иностранцев интонации русского языка (начальный этап контакта испанского языка с русским. М., 1984. 257 с.
21. Румянцев М. Синтез китайских тонов // Вопросы лингвистики. 1988. № 1. С. 82–93.



22. Сорокоумова Д., Корелин О., Сорокоумов А. Построение и обучение нейронной сети для решения задачи распознавания речи // *Transactions of NNSTU* .2015. № 3 (110). С. 77–84.
23. Румянцева Н.М., Гарцова Д.А., Кужаков В.Е. Электронные средства обучения как эффективный инструмент интенсификации процесса обучения китайских учащихся русской фонетике (этнометодический аспект) // *Scientific Research — 2016: Proceedings of Articles the International Scientific Conference (Chezh Republic, Karlovy Vary-Russia, 29–30 September 2016)*. Киров, 2016. С. 175–183.
24. Gabdrahmanova P.L., Bogatova E.N., Mustafina L.R. Development of Communicative Competence in the Online Learning Environment of Russian as a Foreign Language: Opportunities and Prospects // *Educational Technologies and Society*. 2017. № 2. С 329–345.

Березовская Ядвига Леонидовна, к.фил.н., доцент, кафедра русского языка как иностранного, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Исупова Татьяна Дмитриевна, аспирант, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Кацай Дмитрий Алексеевич, к.т.н., доцент, кафедра информационно-измерительной техники, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Шарафутдинова Олеся Ильясовна, к.фил.н., доцент, кафедра русского языка как иностранного, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Шестакова Людмила Ивановна, к.т.н., доцент, заведующая кафедрой, кафедра международных отношений и зарубежного регионоведения, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Елагина Ольга Борисовна, заместитель директора по учебно-методической работе, Институт открытого и дистанционного образования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)