Semiotic System of Musical Texts

Andrew Philippovich, Irina Golubeva, Marina Danshina

Bauman Moscow State Technical University, Moscow, Russia
{aphilippovich,igolubeva,mdanshina}@it-claim.ru

Abstract. In article authors put forward a hypothesis about existence special semiotics system in music, which is close on the structure and mechanisms to a natural language. To check the hypothesis we have selected the ancient Russian chants of XI-XVII centuries, written by Znamenny notation. Using "lingvo-musical" analogies and allocation of the corresponding semiotics designs allowed applying linguistic methods to processing and analyzing chants, identification of their musical "lexicon", syntax and semantics.

Keywords: musical semiotics, Znamenny notation, computational linguistics, thesaurus, syntactic analysis, distributed-statistical analysis.

1 Musical infocognitive technologies and Znamenny chants

The research of the mechanisms of non-verbal human consciousness is one of the promising areas of infocognitive technologies. Music and related cognitive processes which are closely connected with verbal activity hold a special place in the study of these questions.

Music as well as language are a matter of communication and do not exist outside of human communication [1]. Therefore, it is always the result of some human intermediation or performance, although various natural and technological phenomena could also be the sound sources.

The hypothesis that music and language had a common ancestor – "linguomusical system" – was offered to explain the proximity of two cognitive systems. The hypothesis determined their common features [2]. During the development the systems acquired independent and unique features, but they still interact with each other.

Russian musical compositions of XI-XVII AD were recorded using special musical system (notation) which is usually called Znamenny or semiographic. It contains hundreds of special semiographic signs ("znamyas", hooks), each of them corresponds to a certain sequence of sounds of different duration and altitude. Figure 1 presents a fragment of musical manuscript in Znamenny notation.

During the time of Peter's reforms Znamenny notation was replaced by "Italian" one which was simpler and more modern linear musical system which we still use. Unfortunately, the key to decode the melodies was lost during the transformations and this doesn't allow us to translate unambiguously many ancient chants to the contemporary presentation [4].

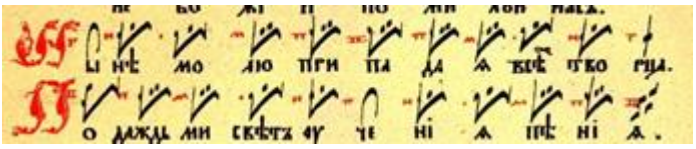


Fig. 1. A fragment of music manuscript in Znamenny notation

For the complete decryption we need to find internal laws in Znamenny notation due to which compositions contain specific signs.

While solving this task in the context of "Automated system of scientific research in the area of computer semiography" project we hypothesized that there should be some semiotic structure which is closely related by its structure and mechanisms to a natural language. This assumption allows us to apply linguistic methods to process and analyze the chants and to reveal its musical "lexicon", syntax, semantic and pragmatic.

In a case of such a hypothesis full confirmation, not merely would we possess valuable results for preserving the rich heritage of national singing culture, but also new fundamental principles of musical infocognitive technologies may be discovered.

2 Toolset development and conducting the research

To solve the problem of automated manuscript processing for selected sources, a work of several years has been carried out that included the following main stages:

- Translation of chants into a digital form;
- Carrying out basic statistical explorations;
- Informational and mathematical models development;
- Models verification, correction and application

During the first stage special computer fonts (such as "Andrew Semio") have been developed and optimized. As well, we've entered manually some semiographic chants and made necessary corrections [4].

During the second stage we've conducted statistical exploration based on an idea firstly proposed in the ancient Russian music study domain by M.V. Brazhnikov [3]. His method implies quantitative counting of semiographic signs occurrences and drawing visual graphs for subsequent analysis. We can mention a paper by B.G. Smolyakov [11] as an example of such a technology application where part one of "Dvoeznamennik 'Irmologion'"(XVII century) was analyzed with manual methods, and comparative graphs for different voices have been drawn.

First, an Andrew Tools linguistic editor was used for chants automated processing; later, a special software complex named "SemioStatistik" was developed [5-6]. It reads data in various formats (Word, Excel), parses tables and cells into constituting parts that are in their turn being written into relevant XML based data structures.

Having done some preliminary processing, SemioStatistik allows one to create various lexicographic structures such as frequency and direct concordances, vocabularies, alphabets et al. as well as to export them into different formats.

Research of dependence of frequency of a semiographic signs from its rank was conducted. It was revealed that this distribution is described by function:

$$Frec = a * e^{-b*rank},$$

a=[500;600] – depending on the manuscript
b=0,07.

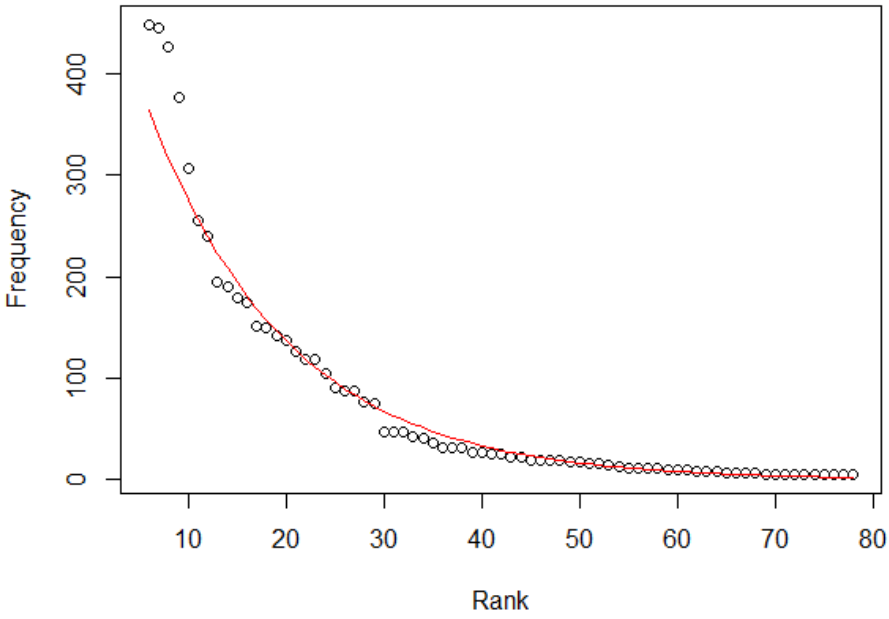


Fig. 2. Frequency dependence of the semiographic signs rank

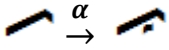
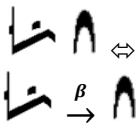
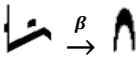
At the third stage we offered informational and mathematical models to describe components of Znamenny chants [9]. In terms of syntax we identified three types of relationships between semiographic signs (see Table 1).

The rules of semiographic signs usage could be represented in the form of Znamenny thesaurus, the structure of which includes syntax, semantic and other relationships. We present the results of building the thesaurus as dictionary entries. Every entry includes the following information: Znamya (semiographic sign), a header of dictionary entry; Basic znamya (α – relationship); Absolute frequency of znamya, numerical characteristic of znamya frequency in manuscript; Znamyas directly related to the key znamya, znamyas which go next to the key znamya (β –relationship); Znamyas context related to the key concept, znamyas which appear with the key znamya in the same context(γ – relationship).

At the fourth stage we have developed algorithms and specialized software and also have conducted a research that revealed 14 main (basic) znamyas. We could obtain other znamyas by applying the first rule (α – relationship).

To reveal syntactic relations of the second type (β-relationship) we built adjacency matrices, which contain the frequencies of znamya sequences. Further normalization concerning the overall number of znamyas allowed us to construct a stochastic table (matrix) of the transitions in Markov chain.

Table 1. Three types of relationships in the syntax of Znamenny chants

Relationship type	Description	Example
α – relationship: $Z_1 \xrightarrow{\alpha} Z_2$	Znamya Z1 is in α – relationship with znamya Z2 if Z2 is derivatived from Z1	
β –relationship: $Z_1 \xrightarrow{\beta} Z_2$	Znamya Z1 is in β – relationship with znamya Z2 if Z2 is is next to Z1	
Probabilistic β – relationship: $Z_1 \xrightarrow{\beta} Z_2 (P_i)$	If znamya Z1 could be followed by variety of znamyas then Z2 is next to Z1 the probability Pi.	 (0,56)
γ – relationship: $Z_1 \xrightarrow{\gamma} Z_2$ $Z_2 \xrightarrow{\gamma} Z_1$	Znamya Z1 is in γ – relationship with znamya Z2 if these znamyas appear in the same context (phrase, sentence, chant).	

To reveal the syntactic connections of the third type (γ-relationship) we applied a statistical distribution analysis which determined the coefficient of the «connection strength» of znamyas according to the formulas for Tanimoto metric:

$$K_{AB} = \frac{f_{AB}}{f_A + f_B - f_{AB}},$$

Figure 3 contains an example of the resulting adjacency matrix.

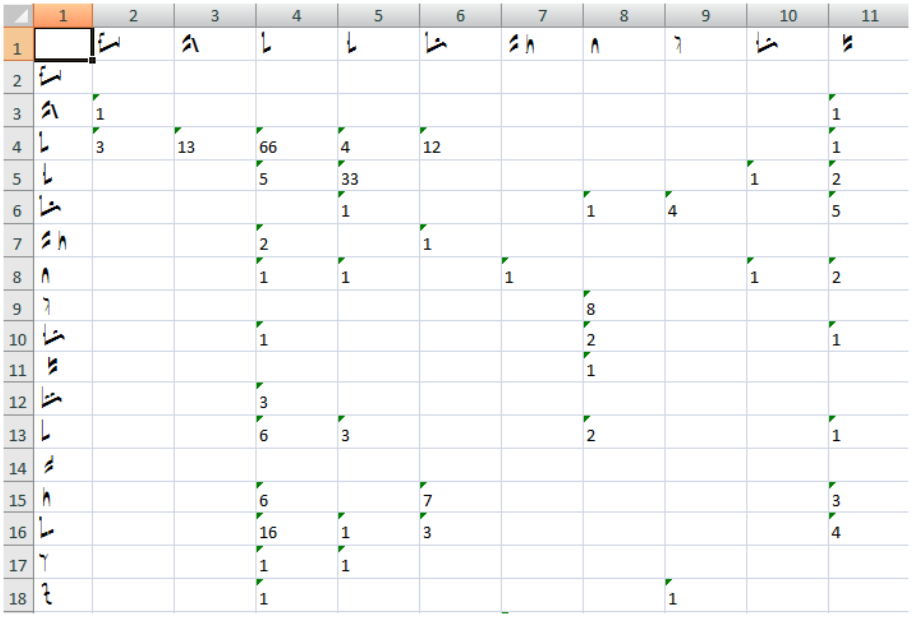


Fig. 3. Adjacency matrix (znamya sequences)

More detailed description of the model is represented in [10].

3 Conclusion

During the research we analyzed the chants from «The Ring of Ancient Znamenny Chants» book containing 24911 uses of different 722 semiographic signs (znamyas). The results of the studies support the hypothesis about the existence of complex semiotic system in Znamenny chants.

- In the general case, znamya corresponds to multiple (sequence) contemporary notes; in some cases one znamya could be replace with the group of other znamyas ("tainozamknennost"). Znamyas could be divided by typeface into main (basic) and secondary (derivative) formed by adding characteristics;
- The occurrence frequency of znamyas corresponds to the exponential law. This indicates that there is a strong spike in the probability of their usage.
- We revealed that there is a huge amount of znamya combinations that are never used; but at the same time there is a small number of combinations that are more common than others. This allows us to identify (confirm) the presence of «function» znamyas.

The application of methods of computational linguistics for the analysis of Znamenny chants, designed mathematical models and algorithms, and the results of the experiments are original and present scientific novelty in the sphere of infocognitive technologies.

The practical value of the conducted research consists of the development of software units for input, presentation and analysis of the chants, and also obtaining new statistical data about collocations of znamyas that could be used to improve data processing and to study Znamenny notation.

The obtained results provide a basis for the further studies of Znamenny chants and other musical compositions, revealing semantic and pragmatic relations, construction of new classes of personal automated systems based on infocognitive technologies.

Additional information about the project and conducted research could be found in the Internet on the website (http://it-claim.ru/semio).

References

1. Tagg P. A Short Prehistory of Western Music. Provisional course material, W310 degree course – The Institute of Popular Music, Liverpool.
2. Wallin N.L., Merker, B., Brown S. (Eds.). The Origins of Music.Cambridge, MA: MIT Press, 2000.
3. Brajnikov M.V. Ancient Russian theory of music. - Leningrad: Muzyka, 1972. (in Russian).
4. Philippovich A.Yu., Smolyakov B.G. Computational semiography // Kniga i mirovaya tsivilizatsiya: Materialy XI Mezhdunar. nauch. konf. po problemam knigovedeniya (Moskva, 20-21 apr. 2004 g.): V 4 t./ [Sost. V.I.Vasil'ev, M.A.Ermolaeva, A.Yu.Samarin; Otv. red. V.I.Vasil'ev, B.V.Lenskiy]. – M.:Nauka, 2004. — T1. – 2004. – Pp.398-401. (in Russian)
5. Philippovich A.Yu., Danshina M.V., Danshina I.V. Methods of computational semiotics on the study of ancient Russian musical writing. //XII International conference on bibliography "Biblioscience. Traditions and innovations". – Moscow: Nauka, 2009 – Pp. 359-360. (in Russian).
6. Danshina I.V. Danshina M.V. Statistical research of znamenny system based on the example of two echoes from Octoechos// Intellektual'nye tekhnologii i sistemy. Sbornik uchebno-metodicheskikh rabot i statey aspirantov i studentov. Vypusk 9. Moscow: SLC "CLAIM", 2007 – Pp. 71-80. (in Russian).
7. Smolyakov B.G. On the problem of znamenny notation deciphering. Voprosy teorii muzyki. Vypusk 3, Moscow: 1975 – Pp. 41—69.. (in Russian).
8. Danshina M.V. IPSM: A software toolset for input and processing of semiographic chants. Information technologies and written heritage: proceedings of International Scientific Conference (Ufa, October 28-31, 2010.)/ Ed. by Baranov V.A. - Ufa; Izhevsk: Vagant, 2010 – Pp. 69-74 (in Russian).
9. Philippovich A.Yu., Golubeva I.V. Syntactic research on semiographical chants // Polygraphy and publishing problems, 2012 – Pp. 147-163.. (in Russian).
10. Golubeva I.V., Philippovich A.Yu. Syntactic analysis musical texts // Analysis of Images, Social Networks, and Texts, 2013 – Pp. 196-204. (in Russian).

Семиотическая система музыкальных текстов

Андрей Филиппович, Ирина Голубева, Марина Даньшина

МГТУ им. Н. Э. Баумана, Москва, Россия
{aphilippovich,igolubeva,mdanshina}@it-claim.ru

Аннотация. Авторы статьи развивают гипотезу о существовании специальной семиотической системы в музыке, близкой по своей структуре и механизмам как естественному языку. Для проверки гипотезы взяты древнерусские песнопения XI-XVII веков, написанные в знаменной нотации. Применение «лингвомузыкальных» аналогий и размещения соответствующих семиотик позволило применить лингвистические методы для обработки и анализа песнопений, идентифицировать их музыкальный «лексикон», синтаксис и семантику.

Ключевые слова: музыкальная семиотика, знаменная нотация, компьютерная лингвистика, тезаурус, синтаксический анализ, распределённый статистический анализ.