

VILNIUS UNIVERSITY

INSTITUTE OF LITHUANIAN LITERATURE AND FOLKLORE

Justina

MANDRAVICKAITĖ

Linguistic characteristics of functional styles: a quantitative analysis

SUMMARY OF DOCTORAL DISSERTATION

Humanities,

Philology (H 004)

VILNIUS 2020

This dissertation was written between 2015 and 2019 at Vilnius University.

Academic supervisor – Assoc. Prof. Dr. Loreta Vilkienė (Vilnius University, Humanities, Philology – H 004)

Academic consultant – Prof. Dr. Tomas Krilavičius (Vytautas Magnus University, Natural Sciences, Informatics – N 009).

This doctoral dissertation will be defended in a public/closed meeting of the Dissertation Defence Panel:

Chairman – **Prof. Dr. Irena Smetonienė** (Vilnius University, Humanities, Philology – H 004).

Members:

Assoc. Prof. Dr. Jurga Cibulskienė (Vilnius University, Humanities, Philology – H 004);

Prof. Dr. Robertas Damaševičius (Kaunas University of Technology, Natural Sciences, Informatics – N 009).

Assoc. Prof. Dr. Erika Rimkutė (Vytautas Magnus University, Humanities, Philology – H 004);

Assoc. Prof. Dr. Andrius Utka (Vytautas Magnus University, Humanities, Philology – H 004).

The dissertation shall be defended at a public/closed meeting of the Dissertation Defence Panel at 15:00 on 18 September 2020 in V. Krėvė Room of the Faculty of Philology at Vilnius University.

Address: 5 Universiteto St., 01131 Vilnius, Lithuania

The text of this dissertation can be accessed at the library of Vilnius University, as well as on the website of Vilnius University: www.vu.lt/lt/naujienos/ivykiu-kalendorius

VILNIAUS UNIVERSITETAS
LIETUVIŲ LITERATŪROS IR TAUTOSAKOS INSTITUTAS

Justina
MANDRAVICKAITĖ

Kalbinė funkcinių stilių raiška: kiekybinė požymių analizė

DAKTARO DISERTACIJOS SANTRAUKA

Humanitariniai mokslai,
filologija (H 004)

VILNIUS 2020

Disertacija rengta 2015– 2019 metais Vilniaus universitete

Mokslinė vadovė – doc. dr. Loreta Vilkienė (Vilniaus universitetas, humanitariniai mokslai, filologija – H 004).

Mokslinis konsultantas – prof. dr. Tomas Krilavičius (Vytauto Didžiojo universitetas, gamtos mokslai, informatika – N 009).

Gynimo taryba:

Pirmininkė – **prof. dr. Irena Smetonienė** (Vilniaus universitetas, humanitariniai mokslai, filologija – H 004).

doc. dr. Jurga Cibulskienė (Vilniaus universitetas, humanitariniai mokslai, filologija – H 004);

prof. dr. Robertas Damaševičius (Kauno technologijos universitetas, gamtos mokslai, informatika – N 009);

doc. dr. Erika Rimkutė (Vytauto Didžiojo universitetas, humanitariniai mokslai, filologija – H 004);

doc. dr. Andrius Utka (Vytauto Didžiojo universitetas, humanitariniai mokslai, filologija – H 004).

Disertacija ginama viešame / uždaramame Gynimo tarybos posėdyje 2020 m. rugsėjo mėn. 18 d. 15 val. Vilniaus universiteto Filologijos fakulteto V. Krėvės auditorijoje. Adresas: Universiteto g. 5, Vilnius 01122, Lietuva.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU Interneto svetainėje adresu: <https://www.vu.lt/naujienos/ivykiu-kalendorius>

SUMMARY

Relevance and novelty of the research. Functional styles (FS) can be defined as language subsystems that are characterized by certain features (lexical, syntactic, others) and are related to different communicative situations in public life (Havránek, 2014 [1964]; Galperin, 1977, 2-33; Chloupek, 1993a; Canning, 2017, 63–85). Today, FS research is generally not very popular in academic world as researchers are more inclined to explore genres, e. g., studies conducted by Shlomo Argamon and colleagues (2003), Jussi Karlgren and Douglass Cutting (1994), Miroslav Kubát (2016), Peter Muntigl and Helmut Gruber (2005), Philipp Petrenz (2012), Dirk Snyman and colleagues (2012), Efsthios Stamatatos, Nikos Fakotakis and George Kokkinakis (2000), Gerard Steen (1999), Cong Zhang and Haitao Liu (2015), etc.

This has happened because FS covers a wide variety of texts in very different genres, making it quite difficult to find common denominators that distinguish as well as link them. As genres include smaller groups of texts, it is easier to find similarities or differences in this regard. Quantitative methods applied in this dissertation allow to look for similarities and differences in FS, which generalize texts of various genres by examining a large amount of textual data. Thus, the author of the dissertation continues the tradition of Lithuanian linguistics in terms of FS but assesses new possibilities with regard to quantitative linguistics. In other words, the dissertation applies new methods for the characterization of FS while definition and typology of FS is based on the works of Lithuanian school of stylistics.

Functional styles and their characteristics have often been studied from the qualitative perspective. For example, Juozas Pikčilingis distinguished 5 FS in standard Lithuanian language (Pikčilingis, 1971a), Kazimieras Župerka studied the competition of linguistic means (Župerka, 1995), the aspects of metalinguistic function in

non-linguistic texts (Župerka, 2000), the criteria for distinguishing functional styles (Župerka, 2012), the peculiarities of style norms (Župerka, 2016), the relationship between standard language and functional styles (Župerka, 2005). Although Regina Koženiauskienė did not analyze functional styles per se, but her works contributed significantly to the development of rhetoric and stylistic research in Lithuania (Koženiauskienė, 1999; 2007; 2011; 2013a; 2013b). Also, Irena Smetonienė's research on advertisement texts (Smetonienė, 2000; 2001; 2005; 2013a; 2003b; 2013c), Rūta Marcinkevičienė's academic review research (2002), Daiva Pagojienė's essay research (2006; 2009), Pranoas Kniūkšta's (2003; 2005) and Rasuolė Vladarskienė's (2005; 2006; 2007) administrative language research, among others, made important contribution to study of FS for Lithuanian language.

Audronė Bitinienė (2001; 2002; 2004; 2007a; 2008; 2009; 2010) studied FS of Lithuanian language from the quantitative perspective via parameters of sentence length and structure. Bitinienė, though studied all the FS in terms of sentence length and structure, focused her researched on publicist style. Vida Žilinskienė also applied statistical methods for analysis of the frequencies and distributions of grammatical forms in the FS (2002a; 2002b; 2002c; 2003; 2005; 2010). Furthermore, Ona Petrėnienė (2011) applied quantitative analysis for the linguistic features of the texts attributed to science popularization substyle of scientific style. One of the significant works in the application of quantitative methods in linguistics is Andrius Utkā's (2004) doctoral dissertation "Statistical Identification of Text Functions", which is related to the functional approach to language. However, the researcher did not study FS but text functions instead. Thus, it can be seen that Lithuanian researchers are gradually familiarize themselves to quantitative methods for the analysis of styles though such studies are still not common.

For other languages, quantitative studies of FS are not abundant as well. As it was already mentioned, research of genre and

individual style is more popular than of FS. However, there are several noteworthy studies, such as Hanna Gnatchuk's, which analyzed the use of English compounds in newspaper style (corresponds to the publicist style) (Gnatchuk, 2015). Also, work by Marco Guerini, Alberto Pepe, and Bruno Lepri, which investigated the influence of language use and language complexity in the abstracts of popular science articles (Guerini, Pepe, & Lepri, 2012). In addition, Solomija N. Buk and Andrij A. Rovenchak analyzed the FS (belles-lettres (fiction), colloquial, scientific, official (business), and journalistic (publicist)) in Ukrainian language via modelling word frequencies and their positions in a frequency table (Buk and Rovenchak, 2004).

Although studies of the FS are not plentiful, there are a number of linguistic studies which use methods of quantitative linguistics (QL) in their approach. QL uses quantitative data to quantify the phenomena of different language levels and to model their interrelationships (Těšitelová, 1992, 13). QL helps to better understand the interrelationships between language elements and processes and allows to study the dynamics of language change and the related problems as well. Radek Čech used quantitative approach to analyze the rhetoric of Czech presidents (Čech, 2011), while Peter Zörnig and Gabriel Altmann (2016), in a similar way, conducted quantitative analysis of the rhetoric of Italian presidents' New Year speeches. Also, Yu Fang and Haitao Liu compared the richness of vocabulary in different translations of the same texts (Fang and Liu, 2015). Ioan-Iovitz Popescu, Radek Čech, and Gabriel Altmann, used quantitative methodology to investigate linguistic properties of individual style as well as features of genres, peculiarities of children's language, and the linguistic aspects of pathological texts (Popescu, Čech, & Altmann, 2011b). Popescu (2009) also studied quantitative properties of words and together with Altmann – the differentiation of language levels (Popescu and Altmann, 2008). Altmann, together with Reinhard Köhler, analyzed different types of

repetitions and language levels in a text (Altmann and Köhler, 2015). To describe the linguistic aspects of texts and to distinguish groups of texts based on different logic, various quantitative indicators have been used by researchers. These indicators were used in a number of works, such as by Ruina Chen and Haitao Liu (2018), Chen, Liu and Altmann (2016), Köhler and Sven Naumann (2008), Miroslav Kubát (2016), Tayebah Mosavi Miangah and Mohammad Javad Rezai (2016), Arjuna Tuzzi, Popescu and Altmann (2010), Peter Zörnig, Emmerich Kelih and Ladislav Fuks (2016) and others.

Although it is clear that the study of language (and texts) via FS theoretical approach has not been popular for quite some time, however, following the tradition Lithuanian stylistic, the author of this dissertation chose the FS approach because language use is related to the situation and linguistic functions and these aspects remain relevant as society undergoes constant changes, which are also reflected in how we use our language. Lithuanian FS has been studied thoroughly enough from a qualitative perspective. However, qualitative research limits a researcher in terms of the amount of research material one can process. Thus, one of the aspects of novelty in this dissertation is a relatively large amount of research material – 3 corpora, which make up a total of ~ 36.4 million words and 19,002 texts¹. Analysis of material of this scope can help to find a more elaborate answer to the question of what the typical characteristics of administrative, scientific, and publicist FS are and to investigate the extent of those characteristics as well.

Although Lithuanian researchers have identified features of specific FS from the qualitative perspective, it is not always clear how those features are expressed. For example, as Župerka writes,

¹ For example, V. Žilinskienė relied on 300,000 words of research material (Žilinskienė, 2002b), A. Bitinienė – on 50,000 words of research material (Bitinienė, 2001), and, for comparison, Andrius Utkā relied on 10.68 million words (463 texts) of research material (Utkā, 2004).

scientific style is characterized by generalization, objectivity, clarity (Župerka, 2012, 81). However, he does not discuss either how these features are expressed or the extent of these features. Thus, the author of this dissertation has relied on the research works of other researchers, such as Vladarskienė (2000a; 2004; 2005; 2006; 2007), Smetonienė (2001; 2005; 2009; 2013a; 2013b; 2013c) and others for compiling the target FS corpora used in this study. As it is not entirely clear how FS features, distinguished in research works based on qualitative methodology, are expressed in terms of linguistic means and how to evaluate the extent of them in the texts, the author of this dissertation decided to analyze linguistic aspects of the target FS via quantitative features. Thus, another aspect of novelty in this dissertation is that although the quantitative indicators of linguistic features used in the study show only some features of linguistic expression determined by qualitative methods, they can be interpreted both mathematically (easy to calculate, estimate the extent and distribution) and linguistically (makes it possible to explain what numerical results mean in the context of language and / or text). Thus, quantitative indicators create preconditions for examining the extent, distribution, representativeness, etc. in terms of the FS properties in a large amount of research material, which would allow to make generalizations about the linguistic means of expression in the studied FS. This is a completely new step in researching the functional styles of the Lithuanian language.

In addition, it should be noted that in the statistical analysis of the FS conducted, for example, by Bitinienė (1983; 1997; 1998; 2000; 2001; 2002; 2003; 2004; 2007a; 2007b; 2007c; 2010) and Žilinskienė (2001; 2002a; 2002b; 2002c; 2003a; 2003b; 2005; 2010), the linguistic means of expression of the FS are analyzed at the sentence level rather than the full text level, while in this dissertation these means of expression are examined at the text level rather than the sentence. This principle is followed because the features of linguistic means of expression in texts may be unevenly distributed,

so it is unlikely that a random section of text of a certain length would represent, for example, full set of features of monograph or novel. Thus, the research presented in the dissertation is new in this respect as well.

Although most of the mentioned works in which the FS were researched do not belong to the field of quantitative linguistics, a considerable amount of work indicates that the topic is widely studied. On the other hand, the application of quantitative linguistic methods to the study of FS, as presented in this dissertation, is new, because so far these methods are not widely used in Lithuanian stylistics. In addition, the study of the distinctive features of the FS, in order to better understand them, can contribute to the development of automated text search and classification systems and, more generally, to new natural language processing systems and to quantitative research of textual documents as well. This is especially true in these times when the amount of electronic information is constantly growing. It should be noted that usefulness of indicators for automatic classification of texts into FS was shown in a study conducted by the author of the dissertation, which was not included in the dissertation, although published together with the co-authors².

In addition, some statistical indicators (a , TK, R_1 , RR_{mc} , VA, etc.) were applied to the Lithuanian language for the first time. It can also be said that the research of this dissertation creates preconditions for the contributions to the FS typology and typical characteristics in terms of empirical confirmation / refutation. Although, as already mentioned, the FS have been studied thoroughly enough in terms of qualitative analysis, quantitative methods have not been applied very

² Mandravickaitė, J., Krilavičius, T., & Man, K. L. (2019). Document Classification to Functional Styles (Domains of Use): Lithuanian Case. *International Journal of Design, Analysis and Tools for Integrated Circuits and Systems*, 8(1), 38-41.

actively, so this dissertation is another step in the direction of quantitative research.

The research described in this dissertation examines not only the individual features of linguistic means of expression, but also their combinations and the interrelationships. The FS are also studied not only separately, but also in comparison to each other, therefore their similarities, differences and the extent of these similarities and differences are also analyzed. Thus, new approaches are introduced for the analysis of the linguistic means of expression in texts and preconditions are created for evaluating the FS characteristics, that were extracted by qualitative methods, via quantitative ones. Finally, the research methods in this dissertation applied to the Lithuanian language and the tested methodology may allow the researchers of Lithuanian (or other language) texts to further analyze the linguistic properties of texts of different genres.

Research object, aim and objectives

The object of in this dissertation is the linguistic expression of written texts in terms of 3 functional styles – administrative, scientific and publicist. The aim is to study the linguistic expression of these 3 functional styles of the Lithuanian language via quantitative linguistic methods and to assess the similarities and differences of these functional styles according to the extent of the features of linguistic expression. This dissertation does not aim to develop new methods. In this dissertation novelty is the application and adaptation of quantitative methods in order to solve the linguistic problem – characterization of the above mentioned FS styles, i. e. in a way that has not been addressed before. The author of this dissertation relied on the research conducted by Miroslav Kubát (2016), Peter Zörnig et al. (2016), Ioan-Iovitz Popescu et al. (Popescu et al., 2010; Popescu, Čech and Altmann, 2011a; Popescu, et al., 2017), Radek Čech (2011; 2016), Czech and colleagues (Čech, Popescu and Altmann, 2014; Čech, Garabík, and Altmann, 2015),

Kubát and Jiří Milička (2013) and juxtaposed quantitative indicators with the FS features taken from qualitative analyses. This led to hypothesis that the following indicators should be appropriate for the study of the target FS linguistic expression:

- VŽFI – readability indicator or the average length of words and word forms in the text;
- a – indicator of the abundance of very frequent (usually function, but not necessarily) words and word forms in the text;
- R_1 – indicator of the diversity of words and word forms (less frequent - belonging to the main parts of speech) in the text);
- RR_{mc} – indicator of word and word form concentration – measures the quantity of words and word forms with an extremely high frequency in the text;
- MATTR – indicator of information flow or average number of different words and word forms in the text;
- TK – indicator of thematic concentration – measures the quantity of usually less frequent words and word forms which belong to the main parts of speech, with a frequency similar to function words and word forms in the text;
- A and D – indicators of text “mood” or activity and descriptivity – indicate the proportions of verbs and adjectives in the text;
- VA - indicator of the syntactic structure complexity of the of the text or the average distance, measured in words and word forms, between 2 consecutive verb forms in the text.

More about these indicators see Section 1.4.

In order to achieve the goal of this dissertation, the following objectives have been set:

1. To develop 3 target corpora. To develop target corpora of administrative and scientific style on the basis of the Corpus of the Contemporary Lithuanian Language³ and target corpus of publicist style – on the basis of the delfi.lt corpus⁴. For more information on development of target corpora, see section 2.
2. To choose a methodology for the study of linguistic expression from the quantitative perspective for large-scale corpora, evaluating the features characteristic for individual groups of texts as well as the similarities and differences between groups of texts according to the scope of quantitative features. Make recommendations for other researchers in the conclusions.
3. To analyze the linguistic expression of the target functional styles according to 9 quantitative indicators – VŽFI, *a*, *R*₁, *RR*_{mc}, *MAT*TR, *TK*, *A*, *D* and *VA* – and to assess what means of linguistic are characteristic to the target functional styles.
4. To analyze which combination of features, indicated by chosen textual indicators, is the most characteristic in terms of the linguistic expression of a particular FS.
5. To analyze the similarities and differences among the linguistic means of expression of the target functional styles (i. e. in comparison to each other) as well as the extent of them and to answer the question regarding how similar and different the target functional styles are.
6. To establish models of linguistic means of expression of target FS, i. e. systematized “portraits” of target functional styles.

³ Full text search in the Corpus of the Contemporary Lithuanian Language is available at <http://tekstynas.vdu.lt/tekstynas/>

⁴ For the needs of the PASTOVU project, the corpus of the articles of the news portal Delfi has been compiled. Project website: <http://mwe.lt/>

Research material and methods

Research material consists of 3 target corpora: corpora of administrative and scientific FS developed by the author of this dissertation on the basis of the Corpus of the Contemporary Lithuanian Language⁵ and corpus of publicist FS developed by the author of this dissertation based on delfi.lt corpus⁶, created for the needs of the project PASTOVU. The original Corpus of the Contemporary Lithuanian Language, as it was, could not be used with the selected text analysis strategies (for more details, see section 2), therefore the relevant parts (administrative and scientific texts) were adapted for research purposes, and the target publicist corpus was developed on the basis of delfi.lt corpus, as the publicist texts included in the Corpus of the Contemporary Lithuanian Language were published only up to 2009. Publicist FS records changes in society and is itself susceptible to changes (Bitinienė, 2006; Butkutė, 2008; Leonavičienė, 2010), thus delfi.lt corpus, which consists of texts, published 2014-2016, was chosen as the basis of target publicist FS corpus. Therefore, as it was already mentioned, 3 target corpora were developed for the research presented in this dissertation:

- corpus of administrative FS texts (AST), which consists of ~5.8 million words and 4,527 texts;
 - corpus of scientific FS texts (MST), which contains ~20.2 million words and 1025 texts;
-

⁵ The author of this dissertation submitted a request to the head of the Center for Computational Linguistics, Vytautas Magnus University. With the permission of Assoc. Dr. Andrius Utkas, this corpus was used for development of target corpora.

⁶ This corpus is not publicly available. The author of the dissertation worked as a junior researcher in the project PASTOVU, therefore she had access to this corpus. However, n-gram frequency lists based on delfi.lt corpus are publicly available (<http://resursai.mwe.lt/atsisiusti>).

- corpus of publicist FS texts (PST), which amounts to ~10.4 million words and 13,450 texts.

Although 5 FS (administrative, scientific, publicist, belles-lettres / fiction and colloquial) are distinguished in standard Lithuanian, in this dissertation it was decided to study linguistic expression of only 3 FS - administrative, scientific and publicist. Colloquial style is not analyzed in this dissertation, because the research is limited to a variety of written language. Also, compiling a large-scale corpus of colloquial style written texts is a complex, time-consuming work. It was decided not to study the linguistic expression of belles-lettres / fiction style due to the limitations of the dissertation in terms of length.

Administrative and scientific FS have similarities in terms of accuracy, clarity and logic (Kniūkšta, 2005; Vladarskienė, 2007), but the former style is distinctive in terms of standardization and formulaicity in the linguistic expression (Žilinskienė, 2002b; Kniūkšta, 2005), so both of them were selected for the study. Publicist style depicts events which take place in the society and influence public opinion (Bitinienė, 2007c, 12). This style has similarities to scientific style in terms of objectivity, conciseness, consistency, and logic, but information is often presented in a form similar to fiction style (Smetonienė, 2013). Therefore, publicist style was chosen to be included in the study. In addition, it should be noted that although this dissertation does not investigate the internal differentiation of FS, texts of all substyles are included in the target corpora.

Quantitative linguistics (CL) **methods** are used in this work (for more details, see Section 1.3). First of all, following the theory of functional stylistics, the author of this dissertation relied on observations by Pikčilingis (1971), Župerka (1983; 1995; 2000; 2004; 2016), Bitinienė (1983; 1997; 1998; 2000; 2001; 2002; 2003; 2004; 2005 ; 2006; 2007a; 2007b; 2007c), Kniūkšta (2003; 2005),

Vladarskienė (2000a; 2000b; 2002; 2004; 2005; 2006; 2007; 2012; 2014), Žilinskienė (2001; 2002a; 2002b; 2002c; 2003a; 2003b; 2005; 2010), Leonavičienė (2004; 2005; 2007; 2010), Petrėnienė (2003; 2011; 2013) that target FS – administrative, scientific and publicist – have characteristic, distinct features in terms linguistic expression. Second, after analysis of works by QL researchers (see Section 1.3.3), quantitative indicators, which indicated variety of features of linguistic expression, were selected. These indicators, in the opinion of the author of this dissertation, most clearly describe characteristics of target FS. A quantitative analysis of the linguistic expression of target FS was then performed. Then a mathematical assessment of the research results was performed and the numerical results were interpreted from the language perspective (in this case – in the context of FS).

Quantitative analysis of FS linguistic expression involved several steps. In the first step, the textual data under study were converted into quantitative data, i. e. for each text in each target corpus quantitative indicators, which indicate features of linguistic expression, were calculated. In the second step of the study, a descriptive analysis of the data was performed in order to evaluate the main statistical properties of the indicators in relation to the studied FS (for more details, see Sections 4.1.1, 4.2.1 and 4.3.1). The relationships between the indicator pairs for each FS were examined, i. e. correlation analysis was performed (for more details, see Sections 4.1.2, 4.2.2 and 4.3.2). A principal component analysis (PCA) was then used to determine the set of indicators that best described the linguistic expression of each target FS (for more details, see Sections 4.1.3, 4.2.3, and 4.3.3). In the last step, a multivariate analysis of variance was performed to determine the similarities and differences of target FS in terms of linguistic expression (for more details, see Section 5).

The following computer programs were used in the analysis of linguistic expression of target FS: *R*⁷, *QUITA*⁸, *MaWaTaTaRaD* and *ThematicConcentration*⁹, morphological annotator¹⁰. *R* and various libraries for this statistical package were used for text preprocessing (for more details, see Section 2.2), for descriptive analysis (for more details, see Section 3.1), for calculation of indicator VA (for more details, see Section 1.4.3.3), for correlation analysis (for more details, see Section 3.2), for principal component analysis (for more details, see Section 3.3), for multivariate analysis of variance (for more details, see Section 3.4), for visualization of results.

QUITA tool was used to calculate the values of the indicators VŽFI, *a*, *R*₁ and *RR*_{mc} (for more details, see Sections 1.4.3.1, 1.4.2.1-1.4.2.3) for each text in the target corpora. *MaWaTaTaRaD* was used to calculate the values of indicator MATTR (for more details, see Section 1.4.2.4) for each text in the target corpora. The values of the thematic concentration indicator (for more details, see Section 1.4.2.5) were calculated with the program *ThematicConcentration*. The texts were annotated morphologically (for more details, see Section 2.2), so that the values of indicators A, D and VA could be calculated (for more details, see Sections 1.4.3.2 and 1.4.3.3).

The structure of dissertation

The dissertation consists of 5 chapters, conclusions and appendix. The first chapter is of a theoretical nature. It presents theoretical

⁷ The program can be downloaded at <https://www.r-project.org/>

⁸ The programs can be downloaded at <https://miroslavkubat.webnode.cz/software/>

⁹ The programs can be downloaded at <http://milicka.cz/en/software.htm>

¹⁰ Created during the project “Automatic Identification of Lithuanian Multiword Expressions (PASTOVU)” (<http://mwe.lt/>) as a part of identification tool for multiword expressions. Morphological annotator is not publicly available. The author of the dissertation worked as a junior researcher in the project PASTOVU, therefore she had access to this tool.

approach to research, which consists of 2 parts: functional stylistics (Sections 1.1-1.2) and quantitative linguistics (Sections 1.3-1.5). The second chapter describes how target corpora were developed and preprocessed. The third chapter describes in detail the research methodology and setup. The fourth chapter, which consists of 3 subsections, describes in detail the study of linguistic expression of each of the studied FS (administrative, scientific, publicist) separately. More specifically, results of descriptive (primary) data analysis, correlation analysis and PCA are discussed. The fifth chapter discusses the similarities and differences in the linguistic expression of the studied FS in comparison to each other. These results were obtained using nonparametric methods of multivariate analysis of variance. At the end of the work, conclusions are presented, research perspectives are indicated and appendix – supplemental diagrams and tables – are attached. Other supplemental material is available at data repository¹¹. The dissertation contains 34 tables and 17 figures. At the end of the work, a reference list of 256 items is attached.

The following theses are to be defended:

1. The linguistic expression of individual functional styles, without comparing them with each other, is represented by different combinations of features.
2. Comparing the target functional styles with each other, each functional style reveals a different scope of features of linguistic expression.
3. Quantitative characteristics are interrelated, i. e. the change in one feature is influenced by the change in another feature, therefore,

¹¹ Supplemental material is available at <https://bitbucket.org/jstn00/papildomi-failai/downloads/>

when analyzing the linguistic expression from the perspective of quantitative linguistics, it is necessary to study the combinations of features, which are marked by the quantitative indicators.

4. The applied research methodology is suitable for modelling target functional styles.

Conclusions

As it was already mentioned, the aim of this work is to study the linguistic expression of 3 functional styles of the Lithuanian language – administrative, scientific and publicist – via quantitative linguistic methods and to assess the similarities and differences of these functional styles in terms of features of linguistic expression (text complexity or readability (denoted by indicator VŽFI), ratio of very frequent and less frequent words and word forms in the text (denoted by indicator a), the diversity of words and word forms (denoted by indicator R_1), the concentrations of words and word forms (denoted by indicator RR_{mc}), information flow (denoted by indicator MATTR), thematic concentration (denoted the by indicator TK), activity / dynamism (denoted by indicator A), descriptivity (denoted by indicator D), and text syntactic complexity (denoted by indicator VA), thus distinguishing generalized models of linguistic expression of target functional styles. They are revealed by total findings of the study. Thus the study of the linguistic expression of the functional styles of written texts – administrative, scientific and publicist – via application of quantitative methods allows to draw the following main conclusions:

1. The selected indicators VŽFI, a , R_1 , RR_{mc} , MATTR, TK, A, D and VA allowed to assess the linguistic expression of the administrative, scientific and publicist functional styles of Lithuanian language, therefore the assumption about their suitability for the study was confirmed.

2. Combinations of features that best represent the linguistic expression of individual functional styles were identified:

2.1. When analyzing the texts of the administrative functional style separately from the scientific and publicist functional styles, the set of features that best represent the linguistic expression of this style (the features are arranged according to the percentage of total contribution, starting with the largest) consists of:

- diversity of (less frequent, belonging to the main part of the speech) words and word forms (denoted by indicator R_1);
- activity / dynamism, or ratio of verbs and adjectives (denoted by indicator A);
- descriptivity or ratio of adjectives and verbs (denoted by indicator D);
- concentration of words and word forms, otherwise – the amount of extremely high frequency words and word forms in the text (denoted indicator RR_{mc});
- information flow, otherwise - number of different words and word forms in the text (denoted by indicator MATTR);
- thematic concentration, otherwise – the number of words and word forms that are essentially less frequent and belong to the main parts of speech, but their frequencies in the text is similar to the frequencies of very frequent (usually – function) words and word forms (denoted by indicator TK).

This combination resulted from PCA. Its summary is presented in Table 1. The higher percentage of contribution and the higher $\cos^2$ value has an indicator in particular principal component (PC1, PC2 or PC3), the larger part of expression of administrative FS it explains. For combination selected only indicators with the highest importance and contribution.

Table 1. Summary of PCA results for administrative FS: importance and contribution of indicators in PC1, PC2 and PC3

	PC1		PC2		PC3	
	Contribution (%)	Importance ($\cos^2$)	Contribution (%)	Importance ($\cos^2$)	Contribution (%)	Importance ($\cos^2$)
VŽFI	0,114	0,004	16,183	0,370	18,206	0,226
<i>a</i>	12,351	0,398	0,165	0,004	3,262	0,041
R_1	27,396	0,883	0,173	0,004	3,049	0,038
RR_{mc}	25,636	0,826	0,111	0,003	3,301	0,041
MATTR	23,760	0,766	2,261	0,052	0,000	0,000
TK	0,119	0,004	5,554	0,127	54,782	0,681
A	0,629	0,020	37,532	0,858	8,065	0,100
D	0,629	0,020	37,532	0,858	8,065	0,100
VA	9,367	0,302	0,489	0,011	1,271	0,016

2.2. Analyzing the texts of the scientific functional style separately from the administrative and publicist functional styles, the set of features that best represent the linguistic expression of this style (the features are arranged according to the percentage of total contribution, starting with the largest) consists of:

- activity / dynamism, otherwise – ratio of verbs and adjectives (denoted by indicator A);
- descriptivity, otherwise – ratio of adjectives and verbs (denoted by indicator D);
- concentration of words and word forms, otherwise – the amount of extremely high frequency words and word forms in the text (denoted by indicator RR_{mc});
- diversity of (less frequent, belonging to the main part of the speech) words and word forms (denoted by indicator R_1);
- complexity of the syntactic structure of the text, otherwise – the distance in words and word forms between two consecutive verb forms (denoted by indicator VA).

This combination resulted from PCA. Its summary is presented in Table 2. The higher percentage of contribution and the higher $\cos^2$ value has an indicator in particular principal component (PC1, PC2 or PC3), the larger part of expression of scientific FS it explains.

Table 2. Summary of PCA results for scientific FS: importance and contribution of indicators in PC1, PC2 and PC3

	PC1		PC2		PC3	
	Contribution (%)	Importance ($\cos^2$)	Contribution (%)	Importance ($\cos^2$)	Contribution (%)	Importance ($\cos^2$)
VŽFI	18,049	0,592	1,207	0,027	9,512	0,103
<i>a</i>	0,678	0,022	19,511	0,434	1,000	0,011
R_1	6,098	0,200	27,682	0,616	4,871	0,053
RR_{mc}	0,007	0,000	31,519	0,701	1,706	0,019
MATTR	5,947	0,195	16,792	0,374	15,094	0,164
TK	13,521	0,443	2,622	0,058	21,511	0,234
A	24,137	0,791	0,092	0,002	9,154	0,100
D	24,137	0,791	0,092	0,002	9,154	0,100
VA	7,426	0,243	0,483	0,011	27,997	0,304

2.3. Analyzing the texts of the publicist functional style separately from the administrative and scientific functional styles, the set of features that best represent the linguistic expression of this style (the features are arranged according to the percentage of total contribution, starting with the largest) consists of:

- diversity of (less frequent, belonging to the main part of the speech) words and word forms (denoted by indicator R_1);
- complexity of the syntactic structure of the text, otherwise – the distance in words and word forms between two consecutive verb forms (denoted by indicator VA);
- activity / dynamism, otherwise - ratio of verbs and adjectives (denoted by indicator A);

- descriptivity, otherwise - ratio of adjectives and verbs (denoted by indicator D);
- thematic concentration, otherwise – the number of words and word forms that are essentially less frequent and belong to the main parts of speech, but their frequencies in the text is similar to the frequencies of very frequent (usually – function) words and word forms (denoted by indicator TK);
- concentration of words and word forms, otherwise – the amount of extremely high frequency words and word forms in the text (denoted by indicator RR_{mc}).

This combination resulted from PCA. Its summary is presented in Table 3. The higher percentage of contribution and the higher $\cos^2$ value has an indicator in particular principal component (PC1, PC2 or PC3), the larger part of expression of publicist FS it explains.

Table 3. Summary of PCA results for publicist FS: importance and contribution of indicators in PC1, PC2 and PC3

	PC1		PC2		PC3	
	Contribution (%)	Importance ($\cos^2$)	Contribution (%)	Importance ($\cos^2$)	Contribution (%)	Importance ($\cos^2$)
VŽFI	8,278	0,265	10,777	0,307	8,173	0,091
<i>a</i>	1,474	0,047	15,736	0,448	8,296	0,093
R_1	7,313	0,234	22,612	0,643	2,328	0,026
RR_{mc}	8,110	0,259	18,920	0,538	7,913	0,088
MATTR	10,628	0,340	13,080	0,372	3,596	0,040
TK	8,114	0,259	0,951	0,027	51,281	0,573
A	18,574	0,593	8,386	0,239	9,206	0,103
D	18,574	0,593	8,386	0,239	9,206	0,103
VA	18,934	0,605	1,151	0,033	0,000	0,000

3. Similarities and differences among target functional styles were identified (see Fig. 1, where the higher value of relative

treatments effects, the higher probability for higher values of particular indicator):

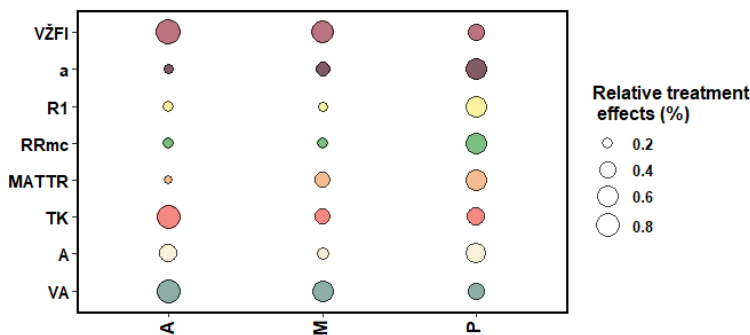


Fig. 1 Summary of quantitative characteristics of the target FS (A – administrative FS, M – scientific FS, P – publicist FS)

3.1. The administrative style, in comparison to other target functional styles, is characterized by:

- longer average length of words and word forms (otherwise – more complex texts in terms of readability; denotes indicator VŽFI);
- higher quantity of very frequent (usually function, but not necessarily) words and word forms (denotes indicator *a*);
- lower concentration of words and word forms – according to this feature administrative and scientific texts are similar (denotes indicator RR_{mc});
- lesser information flow (or on average lesser number of different words and word forms in the text; denotes indicator MATTR);
- higher thematic concentration (higher number of words and word forms, which are essentially less frequent and belong to the main parts of speech, but their frequencies in the case of particular text are similar to the frequencies of high frequency (usual – function) words and word forms;

- more complex syntactic structure (otherwise – a greater distance in words and word forms between 2 consecutive verb forms; denotes indicator VA);
 - according to diversity of words and word forms (less frequent ones, belonging to the main parts of speech; denotes indicator R₁)), administrative texts are similar to scientific ones and lesser diversity of rarer words and word forms is more characteristic in them;
 - according to activity / dynamism (or ratio of verbs and adjectives in the text; denotes by indicator A), administrative texts are more similar to publicist texts (higher extent of activity is more characteristic) than to scientific ones.
- 3.2. Scientific style is between administrative and publicist in terms of some of the features of linguistic expression. Thus, in the following cases, the linguistic expression of scientific texts is assessed according to its closeness to texts of opposite poles – administrative and publicist styles (see Fig. 1):
- lower diversity of words and word forms (less frequent ones, belonging to the main parts of speech) is more typical for scientific texts (denotes by indicator R₁);
 - lower thematic concentration is more typical for scientific texts, i. e. a smaller number of words and word forms, which are essentially less frequent and belong to the main parts of speech, but their frequencies in the case of particular text are similar to the frequencies of high frequency (usual – function) words and word forms;
 - scientific texts are characterized by a lower activity / dynamism (denotes indicator A), otherwise – a smaller number of verbs in comparison to adjectives;
 - in terms of text complexity (readability), otherwise – average length of words and word forms; denotes indicator VŽFI), scientific texts are closer to administrative texts (more complex

texts or on average longer words and forms of words are more typical) than to publicist texts (simpler texts in terms of readability or on average shorter words and forms of words are more characteristic for these texts);

- lower concentration of words and word forms (denotes indicator RR_{mc}) is more typical to scientific texts and in this regard scientific texts are similar to administrative texts;
- in terms of ratio of high frequency and less frequent words and word forms (denotes indicator a), scientific texts are closer to administrative ones (higher quantity of high frequency words and word forms is more characteristic);
- according to information flow (or average number of different words and word forms in the text; denotes indicator MATTR) scientific texts are closer to administrative ones (more typical higher flow of information or higher number of different words and word forms on average);
- according to syntactic structure complexity (denotes indicator VA), scientific texts are closer to administrative ones as well (more complex syntax is more characteristic).

3.3. Publicist style, in comparison to other target functional styles, is characterized by (see Fig. 1):

- on average shorter words and word forms or, alternatively, simpler texts in terms of readability;
- lower quantity of high frequency words and word forms (denotes indicator a);
- higher diversity of words and word forms (less frequent ones, belonging to the main parts of speech);
- higher concentration of words and word forms, i. e. few words and word forms with extremely high frequencies;
- higher information flow (or on average higher number of different words and word forms);

- simpler syntactic structure, otherwise – lesser distance in words and word forms between 2 consecutive verb forms;
 - in terms of thematic concentration, publicist texts are closer to scientific ones (lower thematic concentration, otherwise - a smaller number of words and word forms, which are essentially less frequent and belong to the main parts of speech, but their frequencies in the case of particular text are similar to the frequencies of high frequency (usual – function) words and word forms, is more typical);
 - in terms of activity / dynamism (ratio of verbs to adjectives in the text; denotes indicator A), more dynamic texts are more characteristic of publicist style, although this tendency is not very strong as higher activity of administrative texts has only slightly lesser probability than publicist ones.
4. Relationships between the studied features of linguistic expression were established (see Fig. 2-4):

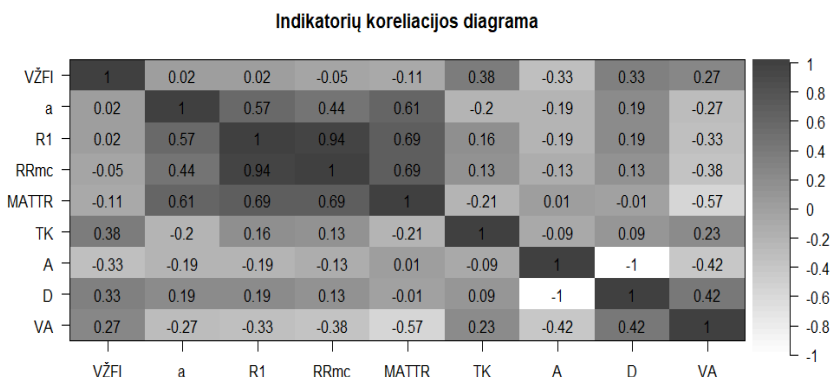


Fig. 2 Correlation of indicators for administrative FS

- 4.1. A very strong relationship was found for all 3 target functional styles between:

- diversity of words and word forms (less frequent ones, belonging to the main parts of speech) and concentration of words and word forms (indicators R_1 and RR_{mc});
- activity / dynamism and descriptivity (indicators A and D).

4.2. Medium strength relationships were identified:

4.2.1. For administrative style – 8 feature pairs (see Fig. 2):

- the diversity of words and word forms (less frequent ones) and ratio of high frequency and less frequent words and word forms in the text (indicators R_1 and a); concentration of words and word forms and ratio of high frequency and less frequent words and word forms in the text (indicators RR_{mc} and a); information flow and ratio of high frequency and less frequent words and word forms in the text (indicators MATTR and a); information flow and diversity of words and word forms (less frequent ones) (indicators MATTR and R_1); information flow and concentration of words and word forms (indicators MATTR and RR_{mc}); information flow and syntactic complexity of the text (indicators MATTR and VA); activity / dynamism and syntactic complexity (indicators A and VA); descriptivity and syntactic complexity (indicators D and VA).

4.2.2. For scientific style - 10 feature pairs (see Fig. 3):

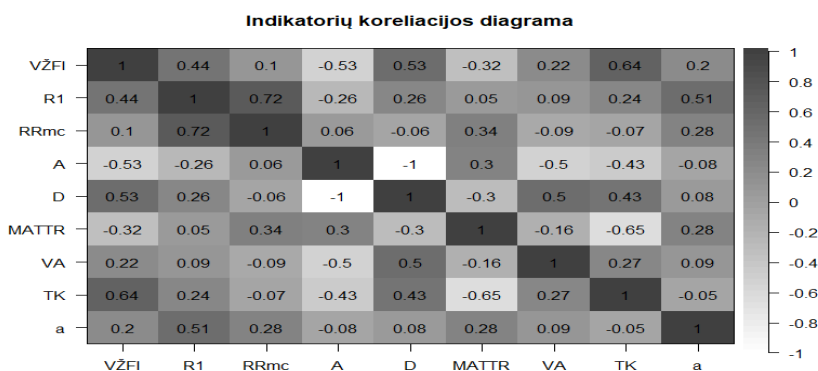


Fig. 3 Correlation of indicators for scientific FS

- text complexity (readability) and diversity of words and word forms (less frequent ones) (indicators VŽFI and R₁); text complexity (readability) and activity / dynamism (indicators VŽFI and A); text complexity (readability) and descriptivity (indicators VŽFI and D); the diversity of word forms (less frequent ones) and ratio of high frequency and less frequent words and word forms in the text (indicators R₁ and a); activity / dynamism and syntactic complexity (indicators A and VA); activity / dynamism and thematic concentration (indicators A and TK); descriptivity and syntactic complexity (indicators D and VA); descriptivity and thematic concentration (indicators D and TK); information flow and thematic concentration (indicators MATTR and TK); text complexity (readability) and thematic concentration (indicators VŽFI and TK).

4.2.3. For publicist style - 11 feature pairs (see Fig. 4):

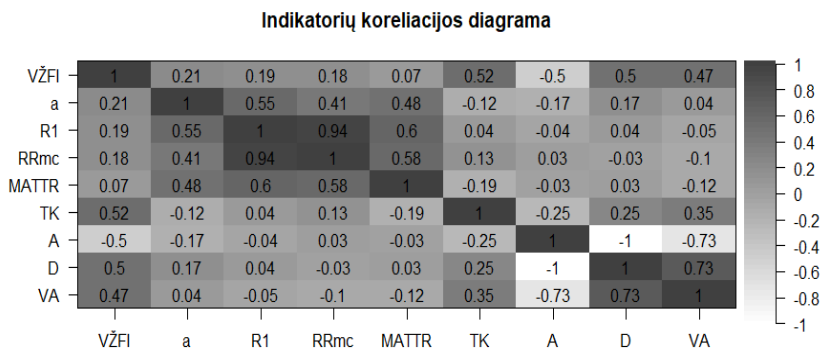


Fig. 4 Correlation of indicators for publicist FS

- text complexity (readability) and thematic concentration (indicators VŽFI and TK); text complexity (readability) and activity / dynamism (indicators VŽFI and A); text complexity (readability) and descriptivity (indicators VŽFI and D); syntactic complexity and text complexity (readability)

(indicators VA and VŽFI); the diversity of word forms (less frequent ones) and ratio of high frequency and less frequent words and word forms in the text (indicators R_1 and a); concentration of words and word forms and ratio of high frequency and less frequent words and word forms in the text (indicators RR_{mc} and a); information flow and ratio of high frequency and less frequent words and word forms in the text (indicators MATTR and a); information flow and the diversity of word forms (less frequent ones) (indicators MATTR and R_1); information flow and concentration of words and word forms (indicators MATTR and RR_{mc}); activity / dynamism and syntactic complexity (indicators A and VA); descriptivity and syntactic complexity (indicators D and VA).

The study of the linguistic expression of administrative, scientific and publicist functional styles carried out in this work allowed to examine linguistic expression of these styles both individually and in comparison to each other as well, thus distinguishing the generalized linguistic expression models of the targeted functional styles. The proposed methodology for studying linguistic expression in large-scale corpora from quantitative perspective was appropriate for assessing the similarities and differences in individual functional styles and in comparison to each other according to the extent of selected quantitative features.

On the other hand, there are **limitations** to the study as well. Traditionally 5 functional styles are distinguished for standard Lithuanian language, and only 3 have been studied in this dissertation. It should also be noted that in this study publicist style was represented by the texts of news portal Delfi, which, although diverse, do not necessarily cover specialized, newspaper or regional varieties of publicist style. In addition, linguistic expression consists of a variety of features at different language levels, and only 9 were used in this study, which, although sufficient to distinguish one

functional style from another, did not cover the full extent of linguistic expression.

The study described in this thesis suggests that **researchers may benefit from the following observations:**

- Functional styles are part of a general functional style system, so it would be more useful to study them not individually, but by comparing them with each other. This would allow to study the extent and dynamics of features, to record the change of functional styles, relations among features and functional styles themselves as well as internal differentiation of functional styles.
- It is useful to refer to a combination rather than an individual feature when analyzing the linguistic expression of functional styles (or texts grouped by other criteria). This would allow for a more accurate assessment of the extent of linguistic expression, as the attributes are often interrelated.

The results of this research could be useful for automatic classification and categorization of texts as automatic grouping of texts according to functional styles may be relevant for libraries, bookstores, planning and development of language resources (texts, databases, etc.), assessing the abundance of texts available online, searching for information by user criteria, document management systems (document / text recommendations, document compliance with security requirements, etc.) and similar purposes. In addition, the results of the analysis of linguistic expression of functional styles can be used in the field of automatic text generation, such as the automatic generation of product/product descriptions from a database entry or image. The results of this study would also make it possible to automatically change the style of texts from one to another, which could be useful for businesses when one style of text, such as administrative text, can be automatically converted to another style, such as publicist text, thus saving time, financial and human resources.

SANTRAUKA

Temos aktualumas ir naujumas

FS galima apibrėžti kaip kalbos posistemes, kurioms būdingi tam tikri požymiai (leksiniai, sintaksiniai, kiti) ir kurios yra susijusios su skirtingomis visuomenės gyvenimo komunikacinėmis situacijomis (Havránek, 2014 [1964]; Galperin, 1977, 2—33; Chloupek, 1993a; Canning, 2017, 63—85). Šiandien FS tyrimai pasaulyje apskritai nėra itin populiarūs, mokslininkai labiau linkę tirti žanrus, pavyzdžiui, Shlomo Argamon ir kolegų (2003), Jussi Karlgrén ir Douglass Cutting (1994), Miroslav Kubát (2016), Peter Muntigl ir Helmut Gruber (2005), Philipp Petrenz (2012), Dirk Snyman ir kolegų (2012), Efsthios Stamatatos, Nikos Fakotakis ir George Kokkinakis (2000), Gerard Steen (1999), Cong Zhang ir Haitao Liu (2015) ir panašūs. Taip yra todėl, kad FS apima didelę labai skirtingų žanrų tekstų įvairovę, tad rasti bendrus juos skiriančius ir siejančius vardiklius yra gana sudėtinga. O žanrai apima mažesnes tekstų grupes, tad šiuo požiūriu rasti panašumus ar skirtumus yra paprasčiau. Kiekybiniai metodai, kurie taikomi šioje disertacijoje, leidžia ieškoti panašumų ir skirtumų FS, kurie subendrina įvairių žanrų tekstus, požiūriu tiriant didelį kiekį tekstinių duomenų, tad disertacijos autorė tęsia lietuvių kalbotyros FS tradiciją, įvertindama naujas kiekybinės lingvistikos galimybes. Kitaip tariant, disertacijoje FS charakterizavimui pritaikomi nauji metodai, o FS apibrėžimui ir išskyrimui remiamasi Lietuvos stilistikos mokyklos mokslininkų darbais.

FS ir jų savybės neretai tirtos ir tiriamos taikant kokybinius tyrimo metodus. Pavyzdžiui, Juozas Pikčilingis išskyrė 5 lietuvių kalbos FS (Pikčilingis, 1971a), Župerka tyrė kalbos priemonių konkurenciją (Župerka, 1995), metalingvistinės funkcijos raišką nelingvistiniuose tekstuose (Župerka, 2000), taip pat šis mokslininkas analizavo funkcinių stilių skyrimo kriterijus (Župerka,

2012), stiliaus normų savitumą (Župerka, 2016), bendrinės kalbos ir funkcinių stilių santykį (Župerka, 2005). Regina Koženiauskiene atskirai funkcinių stilių neanalizavo, tačiau jos darbai daug prisidėjo prie retorikos ir stilistikos mokslo tyrimų plėtotės Lietuvoje (Koženiauskiene, 1999; 2007; 2011; 2013a; 2013b). Taip pat paminėtini Irenos Smetonienės reklamos tekstų tyrimai (Smetonienė, 2000; 2001; 2005; 2013a; 2003b; 2013c), Rūtos Marcinkevičienės akademinių recenzijų tyrimai (2002), Daivos Pagojienės esė tyrimai (2006; 2009), Prano Kniūkštos (2003; 2005) bei Rasuolės Vladarskienės (2005; 2006; 2007) administracinės kalbos tyrimai ir pan.

Lietuvoje kiekybiniais metodais FS per sakinio ilgio ir struktūros parametrus tyrė Audronė Bitinienė (2001; 2002; 2004; 2007a; 2008; 2009; 2010). Ši mokslininkė, nors tyrė ir kitų FS sakinio ilgio ir struktūros parametrus, koncentravosi į publicistinių stilių. Statistinius metodus taip pat taiko Vida Žilinskienė, analizavusi gramatinių formų dažnumus ir pasiskirstymus FS (2002a; 2002b; 2002c; 2003; 2005; 2010). Paminėtina ir tai, kad kiekybiniai metodais mokslinio stiliaus mokslo populiarinamojo postilio tekstų kalbos ypatybes analizavo Ona Petrėnienė (2011). Vienas reikšmingų šių metodų taikymo darbų – Andriaus Utkos (2004) daktaro disertacija „Statistinis tekstų funkcijų nustatymas“, kuri yra susijusi su funkcinio požiūriu į kalbą, tačiau joje minėtas mokslininkas tiria ne FS, o tekstų funkcijas. Taigi matyti, kad Lietuvos mokslininkai pamažu imasi ir analizuoti stilius ir kiekybiniais metodais, tačiau reikia pasakyti, kad tokių tyrimų nėra daug.

Užsienyje FS statistinių tyrimų atlikta taip pat ne tiek ir daug. Kaip jau minėta, žanrų ir individualaus stiliaus tyrimai populiariesni nei FS analizė. Kita verta, galima paminėti Hannos Gnatchuk tyrimą, kuriame analizuotas samplaikų vartojimas anglų kalbos laikraštiniame (angl. *newspaper style*) (vadinasi, publicistiniame) stiliuje (Gnatchuk, 2015). Taip pat minėtini Marco Guerini, Alberto

Pepe ir Bruno Lepri darbai, kuriuose buvo tiriama kalbos priemonių vartojimo ir kalbos sudėtingumo įtaka mokslinių straipsnių (santraukų) populiarumui (Guerini, Pepe ir Lepri, 2012). Be to, Solomija N. Buk ir Andrij A. Rovenchak analizavo ukrainiečių kalbos FS (beletristinį (meninį), šnekamąjį, mokslinį, oficialųjį (verslo) ir publicistinį) per žodžių dažnių ir jų pozicijų dažninėje lentelėje modelį (Buk ir Rovenchak, 2004). Tad, kaip matyti, FS tyrimų nėra gausu.

Kita vertus, kalbos (ir tekstų) tyrimų, kuriose taikomi kiekybinės lingvistikos (toliau – KL) metodai, esama nemažai. KL tyrimuose naudojami kiekybiniai duomenys, kiekybiškai įvertinami skirtingų kalbos lygmenų fenomenai ir modeliuojami jų tarpusavio ryšiai (Těšitelová, 1992, 13). Tai padeda geriau suprasti kalbos elementų ir procesų tarpusavio ryšius, ištirti kalbos kitimo dinamiką ir pan. Taikydami kiekybinius metodus Radek Čech analizavo Čekijos prezidentų retoriką (Čech, 2011), Peter Zörnig ir Gabriel Altmann (2016) – Italijos prezidentų Naujųjų metų kalbų retoriką, Yu Fang ir Haitao Liu lygino teksto vertimų žodyno turtingumą (Fang ir Liu, 2015). Ioan-Iovitz Popescu, Radek Čech ir Gabriel Altmann, naudodami kiekybinę metodologiją, tyrė individualaus stiliaus raišką, žanrų požymius, vaikų kalbos ypatumus, patologinių tekstų kalbinę raišką (Popescu, Čech ir Altmann, 2011b). Popescu (2009) taip pat tyrė kiekybines žodžių savybes, kartu su Altmann – kalbos lygmenų diferenciaciją (Popescu ir Altmann, 2008). Altmann kartu su Reinhard Köhler analizavo įvairių tipų ir kalbos lygmenų pasikartojimus tekste (Altmann ir Köhler, 2015). Įvairūs kiekybiniai indikatoriai tyrėjų buvo naudojami tekstų kalbinės raiškos apibūdinimui ir įvairia logika parentų tekstų grupių atskyrimui, pavyzdžiui, tokius tyrimus atliko Ruina Chen ir Haitao Liu (2018), Chen, Liu ir Altmann (2016), Köhler ir Sven Naumann (2008), Miroslav Kubát (2016), Tayebbeh Mosavi Miangah ir Mohammad Javad Rezai (2016), Arjuna Tuzzi, Popescu ir Altmann (2010), Peter Zörnig, Emmerich Kelih ir Ladislav Fuks (2016) ir kiti.

Nors iš to, kas buvo rašyta, aiškėja, kad kalbos (ir tekstų) tyrimas taikant FS teorines prieigas pasaulyje nėra itin populiarius, tačiau, sekant Lietuvos stilistikos tradicija, šios disertacijos autorės buvo pasirinkta FS prieiga, nes kalbos vartojimas, siejamas su situacija ir kalbos atliekamomis funkcijomis, vis dėlto išlieka aktualus, kadangi visuomenėje vyksta nuolatiniai pokyčiai, kurie atsispindi ir kalbos vartojime. Lietuvių kalbos FS yra neblogai ištirti kokybiniu aspektu, tačiau atliekant kokybinį tyrimą mokslininkas gali apdoroti tik ribotą tiriamosios medžiagos kiekį. Tad vienas iš naujų šioje disertacijoje dalykų – palyginti didelės apimties tiriamoji medžiaga – 3 tekstynai, kuriuos iš viso sudaro ~36,4 mln. žodžių ir 19 002 tekstai. Tokios apimties medžiagos analizė gali padėti geriau atsakyti į klausimą, kokie požymiai būdingi administraciniam, moksliniam ir publicistiniam FS, bei ištirti tų požymių mastą.

Nors taikydami kokybinius metodus lietuvių mokslininkai identifikavo konkrečių FS požymius, ne visada yra iki galo aišku, kaip tie požymiai pasireiškia. Pavyzdžiui, kaip rašo Župerka, moksliniam stiliui būdingas apibendrinimas, objektyvumas, aiškumas (Župerka, 2012, 81), tačiau minėtas mokslininkas neaptaria, kokia yra šių požymių raiška ir koks šių požymių mastas. Taigi sudarant šiame tyrime naudotus tikslinius FS tekstynus disertacijos autorės buvo remiamasi ir kitų tyrėjų – pvz., Vladarskienės (2000a; 2004; 2005; 2006; 2007), Smetonienės (2001; 2005; 2009; 2013a; 2013b; 2013c) ir kitų – išskirtais FS požymiais, tačiau, kadangi neviseškai aišku, kaip atrodo šių požymių raiška tekste, kaip ją įvertinti, šios disertacijos autorės buvo nuspręsta, kad disertacijoje analizuojamų FS kalbinė raiška bus tirama per kiekybinius požymius. Tad kitas naujas šioje disertacijoje aprašomo tyrimo bruožas yra toks: nors tyrime naudojami kiekybiniai kalbinę raišką žymintys indikatoriai rodo tik kai kuriuos kokybiniais metodais nustatytus kalbinės raiškos požymius, tačiau jie gali būti interpretuojami tiek matematiškai (paprasta suskaičiuoti, įvertinti

mastą ir pasiskirstymą), tiek lingvistiškai (sudaro galimybes paaiškinti, ką skaitiniai rezultatai reiškia kalbos ar/arba teksto kontekste). Tad kiekybiniai indikatoriai sudaro prielaidas patikrinti FS kalbinės raidos požymių mastą, pasiskirstymą, tipiskumą ir pan. didelės apimties tiriamojame medžiagoje, kas leistų daryti apibendrinimus apie tiriamųjų FS kalbinę raidą ir jos ypatumus. Tai visiškai naujas žingsnis tiriant funkcinius lietuvių kalbos stilius.

Be to, atkreiptinas dėmesys į tai, kad statistiniuose FS tyrimuose, atliktuose, pavyzdžiui, tiek Bitinienės (1983; 1997; 1998; 2000; 2001; 2002; 2003; 2004; 2007a; 2007b; 2007c; 2010), tiek Žilinskienės (2001; 2002a; 2002b; 2002c; 2003a; 2003b; 2005; 2010), FS kalbinė raida analizuota sakinio, o ne viso teksto lygmeniu¹². Šioje disertacijoje FS kalbinė raida tiriama teksto, o ne sakinio lygmeniu. Šio principo laikomasi, nes kalbinės raidos požymiai tekstuose gali būti pasiskirstę netolygiai, tad mažai tikėtina, kad atsitiktinė tam tikro dydžio teksto atkarpa reprezentuotų, pavyzdžiui, monografijos ar romano raidos ypatybes. Taigi ir šiuo požiūriu disertacijos tyrimas yra naujas.

Nors dauguma paminėtų darbų, kuriuose tirti FS, nepriklauso kiekybinės lingvistikos sričiai, tačiau nemažas darbų kiekis rodo, kad tema plačiai tyrinėjama. Kita vertus, kiekybinės lingvistikos metodų taikymas FS tirti yra nauja, nes iki šiol stilistikoje šie metodai nėra plačiai taikomi. Be to, skiriamųjų FS ypatybių tyrimas, siekiant jas geriau suprasti, gali prisidėti kuriant automatines tekstų paieškos ir klasifikavimo bei apskritai naujas natūralios kalbos apdorojimo sistemas ir atliekant tekstinių dokumentų kiekybinius tyrimus. Tai ypač aktualu šiais laikais, kai elektroninės informacijos kiekis pastoviai auga. Paminėtina, jog tai, kad indikatoriai gali būti naudojami automatiniam tekstų klasifikavimui į FS, parodė

¹² Pavyzdžiui, A. Bitinienė mokomųjų tekstų sakinio lygmens požymių tyrime rėmėsi 3100 sakinių stratifikuota imtimi (Bitinienė, 2007a).

disertacijos autorės atliktas, bet į disertaciją netilpęs tyrimas, publikuotas kartu su bendraautorais¹³. Be to, kai kurie statistiniai indikatoriai (*a*, TK, R₁, RR_{mc}, VA, etc.) lietuvių kalbai buvo pritaikyti pirmą kartą.

Taip pat galima sakyti, kad šios disertacijos tyrimas sudaro prielaidas FS tipologijos ir požymių papildymui bei empiriniam patvirtinimui/paneigimui. Nors, kaip jau buvo minėta, FS kokybiniu požyriu yra neblogai ištirti, kiekybiniai metodai nebuvo labai aktyviai taikyti, tad ši disertacija – dar vienas žingsnis kiekybinių tyrimų linkme. Atliekant šioje disertacijoje aprašomą tyrimą, analizuojami ne tik atskiri kalbinės raiškos požymiai, bet ir jų kombinacijos, požymių tarpusavio ryšiai. Taip pat FS tiriami ne tik atskirai, bet analizuojami ir jų tarpusavio panašumai, skirtumai bei panašumų ir skirtumų mastas. Tad atveriamos naujos prieigos tekstų raiškos priemonių analizei bei sudaromos prielaidos patikrinti kokybiniais metodais išskirtus FS charakteristikas. Be to, disertacijoje lietuvių kalbai pritaikyti tyrimo būdai, patikrinta metodika gali leisti lietuvių kalbos tekstų tyrėjams toliau analizuoti ir skirtingų žanrų tekstų raiškos požymius.

Darbo objektas, tikslas ir uždaviniai

Šioje disertacijoje aprašomo tyrimo objektas yra 3 funkcinių stilių – publicistinio, mokslinio ir administracinio – rašytinių tekstų kalbinė raiška. Tikslas – kiekybinės lingvistikos metodais ištirti 3 lietuvių kalbos funkcinių stilių – administracinio, mokslinio ir publicistinio – kalbinę raišką ir įvertinti šių funkcinių stilių panašumus ir skirtumus pagal kalbinės raiškos požymių mastą. Šioje

¹³ Mandravickaitė, J., Krilavičius, T., & Man, K. L. (2019). Document Classification to Functional Styles (Domains of Use): Lithuanian Case. *International Journal of Design, Analysis and Tools for Integrated Circuits and Systems*, 8(1), 38-41.

disertacijoje nėra siekiama kurti naujus metodus. Joje pritaikomi metodai lingvistinei problemai – minėtų FS stilių charakterizavimui – spręsti taikant kiekybinės lingvistikos metodus, t. y. taip, kaip ji anksčiau nebuvo sprendžiama. Šio darbo autorė, pasirėmusi Miroslav Kubát (2016), Peter Zörnig ir kolegų (2016), Ioan-Iovitz Popescu ir kolegų (Popescu et al., 2010; Popescu, Čech ir Altmann, 2011a; Popescu, et al., 2017), Radek Čech (2011; 2016), Čech ir kolegų (Čech, Popescu ir Altmann, 2014; Čech, Garabík, ir Altmann, 2015), Kubát ir Jiří Milička (2013) mokslininkų darbais ir sugretinusi kiekybinius požymius (indikatorius) su kokybinei prieigai skirtais požymiais, iškėlė prielaidą, kad FS kalbinei raiškai tirti, turėtų tikti šie indikatoriai:

- VŽFI – teksto sudėtingumo įvertis arba vidutinis žodžių ir žodžių formų ilgis tekste;
- a – labai dažnų (paprastai – funkcinių, bet nebūtinai) žodžių ir žodžių formų gausos tekste įvertis;
- R_1 – žodžių ir jų formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovės tekste įvertis);
- RR_{mc} – žodžių ir žodžių formų koncentruotumo įvertis – vertina žodžių ir žodžių formų, turinčių itin didelį dažnumą, kiekį tekste;
- MATTR – informacijos srauto įvertis arba vidutinis skirtingų žodžių ir žodžių formų kiekis tekste);
- TK – teminės koncentracijos įvertis – vertina iš esmės retesnių, pagrindinėms kalbos dalims priklausančių žodžių ir žodžių formų, dažnumu bemaž prilygstančių funkciniais žodžiams ir žodžių formoms, kiekį tekste);
- A ir D – teksto „tono“ arba aktyvumo ir deskriptyvumo įverčiai – žymi veiksmažodžių ir būdvardžių santykį tekste;
- VA – teksto sintaksinės struktūros kompleksiskumo įvertis arba vidutinis atstumas tekste tarp 2 iš eilės einančių veiksmažodžių.

Plačiau apie šiuos indikatorius rašoma 1.4 sk.

Kad tikslas būtų pasiektas, kelti tokie uždaviniai:

1. Parengti 3 tikslinius tekstynus. Dabartinės lietuvių kalbos tekstyno pagrindu parengti administracinio ir mokslinio stiliaus tekstynus, delfi.lt tekstyno¹⁴ pagrindu – publicistinio stiliaus tekstyną. Apie tekstynų parengimą plačiau žr. 2 sk.
2. Parinkti metodiką kalbinei raiškai didelės apimties tekstuose tirti kiekybiniu aspektu, įvertinant tiek atskiroms tekstų grupėms būdingus požymius, tiek tekstų grupių tarpusavio panašumus ir skirtumus pagal kiekybinių požymių raiškos mastą. Formuluoiant išvadas pateikti rekomendacijų, į ką reiktų atkreipti dėmesį.
3. Išanalizuoti tiriamiems funkciniais stiliams būdingą kalbinę raišką pagal 9 kiekybinius indikatorius – VŽFI, a , R_1 , RR_{mc} , MATTR, TK, A, D ir VA – ir nustatyti, kas būdinga tiriamiems funkciniais stiliams.
4. Ištirti, kokia tiriamųjų požymių kombinacija būdingiausia konkrečiam FS kalbinei raiškai.
5. Ištirti tikslinių funkcinių stilių kalbinės raiškos tarpusavio panašumus ir skirtumus požymių raiškos mastą ir atsakyti į klausimą, kuo tarpusavyje panašūs ir kuo – skiriasi tiriamieji funkciniai stiliai.
6. Pateikti tekstų tyrėjams FS kalbinės raiškos modelius, t. y. susistemintus tikslinių funkcinių stilių „portretus“.

Tiriamoji medžiaga ir metodai

Tiriamoji medžiaga – Dabartinės lietuvių kalbos tekstyno pagrindu disertacijos autorės parengti administracinio ir mokslinio FS

¹⁴ Projekto PASTOVU reikmėms sudarytas naujienų portalo Delfi straipsnių tekstynas, plačiau žr. <http://mwe.lt/>

tekstynai bei projekto PASTOVU reikmėms sukurto delfi.lt tekstyno pagrindu disertacijos autorės parengtas publicistinio FS tekstynas. Originalaus Dabartinės lietuvių kalbos tekstyno (DLKT) tekstams nebuvo galima taikyti disertacijoje numatytų tekstų analizės strategijų (plačiau žr. 2 sk.), todėl atitinkamos jo dalys (administraciniai ir moksliniai tekstai) buvo pritaikytos tyrimo reikmėms, o publicistinis tekstynas parengtas delfi.lt tekstyno pagrindu, nes publicistiniai tekstai, įtraukti į DLKT publikuoti tik iki 2008 m. imtinai. Publicistinis FS, fiksuodamas pokyčius visuomenėje, pats yra imlus pokyčiams (Bitinienė, 2006; Butkutė, 2008; Leonavičienė, 2010), tad tikslinio publicistinio FS tekstyno pagrindu pasirinktas delfi.lt tekstynas, kurį sudaro 2014-2016 m. publikuoti tekstai. Tad, kaip jau minėta, šios disertacijos autorės tyrimui buvo parengti 3 tiksliniai tekstynai:

- administracinio FS tekstynas (toliau – AST), kurį sudarė apie 5,8 mln. žodžių, 4 527 tekstai;
- mokslinio FS tekstynas (toliau – MST), jame buvo apie 20,2 mln. žodžių, 1025 tekstai;
- publicistinio FS tekstynas (toliau – PST), kurį sudarė apie 10,4 mln. žodžių, 13 450 tekstų).

Nors lietuvių kalboje skiriami 5 FS (administracinis/dalykinis, mokslinis, publicistinis, meninis/beletristinis ir šnekamasis/buitinis), šioje disertacijoje apsispręsta tirti tik 3 FS – administracinio, mokslinio ir publicistinio – kalbinę raišką. Šnekamojo/buitinio stiliaus kalbinė raiška šioje disertacijoje neanalizuojama, nes tyrime apsiribota rašytinės kalbos atmaina, o surinkti didelės apimties buitinio stiliaus rašytinių tekstų tekstyną yra sudėtingas, daug laiko reikalaujantis darbas. Meninio/beletristinio stiliaus kalbinės raiškos apsispręsta netirti dėl apribojimų disertacijos apimčiai.

Administracinis ir mokslinis FS turi tarpusavio panašumų tikslumo, aiškumo ir logiškumo požiūriu (Kniūkšta, 2005;

Vladarskienė, 2007), tačiau juos vieną nuo kito skiria administracinio stiliaus standartiškumas/sablioniškumas (Žilinskienė, 2002b; Kniūkšta, 2005), tad tyrimui pasilikti abu. Publicistinis stilius atvaizduoja visuomenėje vykstančius įvykius bei daro poveikį visuomenės nuomonei (Bitinienė, 2007c, 12). Jį su moksliniu stiliumi sieja dalykiškumas, glaustumas, nuoseklumas, logiškumas, tačiau informacija dažnai pateikiama panašia į meninio stiliaus forma (Smetonienė, 2013). Tad į tyrimą pasirinkta įtraukti ir publicistinį stilių. Be to, pasakytina, kad nors šioje disertacijoje netiriama FS vidinė diferenciacija, tačiau į tikslinius tekstynus įtraukti visų postilių tekstų.

Šiame darbe taikomi kiekybinės lingvistikos (KL) metodai (plačiau žr. 1.3 skyrių). Visų pirma, sekant funkcinės stilistikos teorija, šios disertacijos autorės buvo pasiremta Pikčilingio (1971), Župerkos (1983; 1995; 2000; 2004; 2016), Bitinienės (1983; 1997; 1998; 2000; 2001; 2002; 2003; 2004; 2005; 2006; 2007a; 2007b; 2007c), Kniūkštos (2003; 2005), Vladarskienės (2000a; 2000b; 2002; 2004; 2005; 2006; 2007; 2012; 2014), Žilinskienės (2001; 2002a; 2002b; 2002c; 2003a; 2003b; 2005; 2010), Leonavičienės (2004; 2005; 2007; 2010), Petrėnienės (2003; 2011; 2013) įžvalgomis, kad FS stiliai – administracinis, mokslinis ir publicistinis – turi jiems tipiskų, skirtingų kalbinės raiškos požymių. Antra, išanalizavus KL mokslininkų darbus (žr. 1.3.3 sk.), buvo pasirinkti kiekybiniai indikatoriai, kurie, šios disertacijos autorės nuomone, akivaizdžiausiai apibūdina tiriamuosius FS. Tada buvo atlikta tiriamųjų FS kalbinės raiškos kiekybinė analizė. Tuomet buvo atliktas matematinis tyrimo rezultatų įvertinimas, skaitiniai rezultatai interpretuoti kalbos požiūriu (šiuo atveju – FS kontekste).

FS kalbinės raiškos kiekybinė analizė apėmė keletą etapų. Pirmame etape tiriamieji tekstiniai duomenys buvo paversti kiekybiniais duomenimis, t. y. kiekvienam kiekvieno tikslinio tekstyno tekstui buvo suskaičiuojami kalbinės raiškos požymius žymintys kiekybiniai indikatoriai. Antrame tyrimo etape atlikta

aprašomoji duomenų analizė, kad būtų įvertintos pagrindinės statistinės indikatorių savybės tiriamųjų FS atžvilgiu (plačiau žr. 4.1.1, 4.2.1 ir 4.3.1 sk.). Tada buvo tiriami ryšiai tarp indikatorių kiekviename FS, t. y. atliekama koreliacinė analizė (plačiau žr. 4.1.2, 4.2.2 ir 4.3.2 sk.). Paskui taikyta pagrindinių komponentių analizė (PKA), kad būtų nustatytas kiekvieno tiriamojo FS kalbinę raišką geriausiai apibūdinančių indikatorių rinkinys (plačiau žr. 4.1.3, 4.2.3 ir 4.3.3 sk.). Paskutiniu etapu buvo atlikta dispersinė analizė tiriamųjų FS kalbinės raiškos tarpusavio panašumų ir skirtumų nustatymui (plačiau žr. 5 sk.).

Darbe atliekant FS kalbinės raiškos analizę naudojamos šios kompiuterinės programos: R^{15} , *QUITA*¹⁶, *MaWaTaTaRaD* ir *ThematicConcentration*¹⁷, morfologinis anotatorius¹⁸. Programa R ir įvairios šiam statistiniam paketui skirtos bibliotekos naudotos tekstų sutvarkymui (plačiau žr. 2.2 sk.), pirminei analizei (plačiau žr. 3.1 sk.), indikatoriaus VA (plačiau žr. 1.4.3.3 sk.) verčių skaičiavimui, koreliacinei analizei (plačiau žr. 3.2 sk.), pagrindinių komponentių analizei (PKA) (plačiau žr. 3.3 sk.), dispersinei analizei (plačiau žr. 3.4 sk.), rezultatų vizualizavimui. Įrankis *QUITA* pasitelktas indikatorių VŽFI, a , R_1 , RR_{mc} (plačiau žr. 1.4.3.1, 1.4.2.1-1.4.2.3 sk.) verčių skaičiavimui kiekvienam tiriamosios medžiagos tekstui. *MaWaTaTaRaD* naudotas indikatoriaus MATTR (plačiau žr. 1.4.2.4 sk.) vertėms kiekvienam tiriamų tekstynų tekstui suskaičiuoti. Teminės koncentracijos indikatoriaus (plačiau žr. 1.4.2.5 sk.) vertės apskaičiuotas programa *ThematicConcentration*. Tekstai buvo anotuoti morfologiškai (plačiau žr. 2.2 sk.), kad būtų galima vėliau rasti indikatorių A, D ir VA vertes (plačiau žr. 1.4.3.2-1.4.3.3 sk.).

¹⁵ Pasiikiama adresu <https://www.r-project.org/>

¹⁶ Pasiikiama adresu <https://miroslavkubat.webnode.cz/software/>

¹⁷ Pasiikiama adresu <http://milicka.cz/en/software.htm>

¹⁸ Morfologinis anotatorius buvo sukurtas projekto PASTOVU metu kaip pastoviųjų žodžių junginių atpažinimo įrankio dalis. Nėra viešai prieinamas.

Darbo struktūra

Disertaciją sudaro 5 skyriai, išvados ir priedai. Pirmas skyrius yra teorinio pobūdžio, jame pristatomos teorinės tyrimo prieigos, kurias sudaro 2 dalys: funkcinė stilistika (1.1-1.2 sk.) ir kiekybinė lingvistika (1.3-1.5 sk.). Antrame skyriuje apžvelgiama tyrimo medžiaga ir jos parengimas, trečiajame nuodugnai aprašoma tyrimo metodika ir eiga. Ketvirtajame skyriuje, kurį sudaro 3 poskyriai, išsamiai aprašomas kiekvieno iš tiriamųjų FS (administracinio, mokslinio, publicistinio) atskirai kalbinės raiškos tyrimas, o konkrečiau: aptariama aprašomoji (pirminė) duomenų analizė ir jos rezultatai, koreliacinė analizė ir PKA. Penktame skyriuje aptariami tiriamųjų FS kalbinės raiškos tarpusavio panašumai ir skirtumai, gauti taikant nparametrinius dispersinės analizės metodus. Darbo pabaigoje teikiamos išvados, nurodomos tyrimo perspektyvos ir pridedami priedai – diagramos ir lentelės. Kiti priedai dėl didelės duomenų apimties pateikiami duomenų talpykloje¹⁹. Disertacijoje yra 34 lentelės ir 17 paveikslų. Darbo pabaigoje pridedamas 256 pozicijų naudotos literatūros sąrašas.

Ginamieji teiginiai

Gynimui teikiami šie teiginiai:

1. Atskirų funkcinų stilių kalbinę raišką, nelyginant jų tarpusavyje, reprezentuoja skirtingos požymių kombinacijos.
2. Lyginant tikslinius funkcinus stilius tarpusavyje, kiekviename funkciname stiliuje išryškėja skirtingas požymių raiškos mastas.
3. Kiekybiniai požymiai yra tarpusavyje susiję, t. y. vieno požymio raiškos kitimui turi įtakos kito požymio raiškos kitimas, todėl,

¹⁹ Pasiekama <https://bitbucket.org/jstn00/papildomi-failai/downloads/>

analizuojant kalbinę raišką kiekybinės lingvistikos požiūriu, privalu tirti požymių, kuriuos žymi tyrime naudojami kiekybiniai indikatoriai, kombinacijas.

4. Taikyta analizės metodika yra tinkama tikslinių funkcinių stilių modelių išskyrimui.

Išvados

Kaip jau buvo minėta, šio darbo tikslas – kiekybinės lingvistikos metodais ištirti 3 lietuvių kalbos funkcinių stilių – administracinio, mokslinio ir publicistinio – kalbinę raišką ir įvertinti šių funkcinių stilių panašumus ir skirtumus pagal kalbinės raiškos požymių (teksto sudėtingumo (žymi indikatorius VŽFI), labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (žymi indikatorius a), žodžių ir žodžių formų (retesnių) įvairovės (žymi indikatorius R_1), žodžių ir žodžių formų koncentracijos (žymi indikatorius RR_{mc}), informacijos srauto (žymi indikatorius MATTR), teminės koncentracijos (žymi indikatorius TK), dinamiškumo (žymi indikatorius A), deskriptyvumo (žymi indikatorius D) ir teksto sintaksinio kompleksiško (žymi indikatorius VA)) mastą, taip išskiriant apibendrintus tikslinių funkcinių stilių kalbinės raiškos modelius. Juos atskleidžia tyrimo išvadų visuma. Atliktas rašytinių tekstų funkcinių stilių – administracinio, mokslinio ir publicistinio – kalbinės raiškos tyrimas taikant kiekybinės lingvistikos metodus leidžia daryti tokias pagrindines išvadas:

1. Pasirinkti indikatoriai VŽFI, a , R_1 , RR_{mc} , MATTR, TK, A, D ir VA leido įvertinti lietuvių kalbos administracinio, mokslinio ir publicistinio funkcinių stilių kalbinę raišką, tad prielaida apie jų tinkamumą kalbinės raiškos tyrimui iš kiekybinės perspektyvos pasitvirtino.
2. Nustatyti atskirų tiriamųjų funkcinių stilių kalbinę raišką geriausiai reprezentuojančių požymių kombinacijos:

2.1 administracinio funkcinio stiliaus tekstus analizuojant atskirai nuo mokslinio ir publicistinio funkcinų stilių, šio stiliaus kalbinę raišką geriausiai reprezentuojančių požymių rinkinį (požymiai išdėstyti pagal bendrą procentinę indėlį, pradedant nuo didžiausio) sudaro:

- žodžių ir žodžių formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovė (žymi indikatorius R_1);
- aktyvumas / dinamiškumas, kitaip – veiksmažodžių ir būdvardžių santykis (žymi indikatorius A);
- deskriptyvumas / aprašomumas, kitaip – būdvardžių ir veiksmažodžių santykis (žymi indikatorius D);
- žodžių ir žodžių formų koncentruotumas, kitaip – itin didelio dažnio žodžių ir žodžių formų kiekis tekste (žymi indikatorius RR_{mc});
- informacijos srautas, kitaip – skirtingų žodžių ir žodžių formų kiekis tekste (žymi indikatorius MATTR);
- teminė koncentracija, kitaip – iš esmės retesnių, pagrindinėms kalbos dalims priklausančių, tačiau tekste labai dažniems (funkciniams) dažnumu bemaž prilygstančių žodžių ir žodžių formų kiekis (žymi indikatorius TK).

2.2 Mokslinio funkcinio stiliaus tekstus analizuojant atskirai nuo administracinio ir publicistinio funkcinų stilių, šio stiliaus kalbinę raišką geriausiai reprezentuojančių požymių rinkinį (požymiai išdėstyti pagal bendrą procentinę indėlį, pradedant nuo didžiausio) sudaro:

- aktyvumas / dinamiškumas, kitaip – veiksmažodžių ir būdvardžių santykis (žymi indikatorius A);
- deskriptyvumas / aprašomumas, kitaip – būdvardžių ir veiksmažodžių santykis (žymi indikatorius D);
- žodžių ir žodžių formų koncentruotumas, kitaip – itin didelio dažnio žodžių ir žodžių formų kiekis tekste (žymi indikatorius RR_{mc});

- žodžių ir žodžių formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovė (žymi indikatorius R_1);
- teksto sintaksinės struktūros kompleksiskumas – atstumas žodžiais tarp dviejų iš eilės einančių veiksmažodžio formų (žymi indikatorius VA).

2.3 Publicistinio funkcinio stiliaus tekstus analizuojant atskirai nuo administracinio ir mokslinio funkcinį stilių, šio stiliaus kalbinę raišką geriausiai reprezentuojančių požymių rinkinį (požymiai išdėstyti pagal bendrą procentinį indėlį, pradedant nuo didžiausio) sudaro:

- žodžių ir žodžių formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovė (žymi indikatorius R_1);
- teksto sintaksinės struktūros kompleksiskumas – atstumas žodžiais tarp dviejų iš eilės einančių veiksmažodžio formų (žymi indikatorius VA);
- aktyvumas / dinamiškumas, kitaip – veiksmažodžių ir būdvardžių santykis (žymi indikatorius A);
- deskriptyvumas / aprašomumas, kitaip – būdvardžių ir veiksmažodžių santykis (žymi indikatorius D);
- teminė koncentracija, kitaip – iš esmės retesnių, pagrindinėms kalbos dalims priklausančių, tačiau tekste labai dažniems (funkciniams) dažnumu bemaž prilygstančių žodžių ir žodžių formų kiekis (žymi indikatorius TK);
- žodžių ir žodžių formų koncentruotumas, kitaip – itin didelio dažnio žodžių ir žodžių formų kiekis tekste (žymi indikatorius RR_{mc}).

3. Nustatyti panašumai ir skirtumai tarp tiriamųjų funkcinį stilių:

3.1. Administraciniam stiliui, lyginant visus tiriamuosius funkcinį stilius, būdingiau:

- ilgesnis vidutinis žodžių ir žodžių formų ilgis (kitaip – suprantamumo prasme sudėtingesni tekstai; žymi indikatorius VŽFI);

- didesnė gausa labai dažnų (paprastai – funkcinių, bet nebūtinai) žodžių ir žodžių formų (žymi indikatorius a);
- mažesnė žodžių ir žodžių formų koncentracija – pagal šį rodiklį administraciniai ir moksliniai tekstai yra panašūs (indikatorius RR_{mc});
- mažesnis informacijos srautas (arba vidutiniškai mažesnis kiekis skirtingų žodžių ir žodžių formų; žymi indikatorius $MATTR$);
- didesnė teminė koncentracija (didesnis kiekis iš esmės retesnių, pagrindinėms kalbos dalims priklausančių, tačiau administraciniuose tekstuose labai dažniems (funkciniams) dažnumu bemaž prilygstančių žodžių ir žodžių formų);
- kompleksiškesnė sintaksinė struktūra (kitaip – didesnis atstumas žodžiais tarp 2 iš eilės einančių veiksmažodžių; žymi indikatorius VA);
- pagal žodžių ir žodžių formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovės raišką administraciniai tekstai yra panašesni į mokslinius – būdingesnė mažesnė retesnių žodžių ir žodžių formų įvairovė (žymi indikatorius R_1);
- pagal aktyvumo (veiksmažodžių ir būdvardžių santykį tekste; žymi indikatorius A) raišką administraciniai tekstai yra panašesni į publicistinius (būdingesnė didesnė aktyvumo raiška) tekstus nei į mokslinius.

3.2 Mokslinis stilius pagal dalį kalbinės raiškos požymių yra tarp administracinio ir publicistinio. Tad šiais atvejais mokslinių tekstų kalbinė raiška įvertinama pagal jos artimumą priešingų polių – administracinio ir publicistinio – stilių tekstams:

- moksliniams tekstams tipiškesnė mažesnė žodžių ir žodžių formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovė (žymi indikatorius R_1);
- moksliniams tekstams tipiškesnė mažesnė teminė koncentracija, t. y. mažesnis kiekis iš esmės retesnių, pagrindinėms kalbos

dalims priklausančių, tačiau moksliniuose tekstuose labai dažniems (funkciniams) dažnumu bemaž prilygstančių žodžių ir žodžių formų;

- moksliniams tekstams būdingesnė mažesnė aktyvumo raiška (žymi indikatorius A), kitaip – lyginant su būdvardžiais, mažesnis kiekis veiksmazodžių;
- suprantamumo sudėtingumo atžvilgiu (kitaip – pagal vidutinį žodžių ir žodžių formų ilgį; žymi indikatorius VŽFI) moksliniai tekstai yra artimesni administraciniams (būdingesni suprantamumo atžvilgiu sudėtingesni tekstai – vidutiniškai ilgesni žodžiai ir žodžių formos) negu publicistiniams (būdingesni suprantamumo atžvilgiu paprastesni tekstai – vidutiniškai trumpesni žodžiai ir žodžių formos);
- mažesnė žodžių ir žodžių formų koncentracija (žymi indikatorius RR_{mc}) – pagal šį rodiklį administraciniai ir moksliniai tekstai yra panašūs;
- pagal labai dažnų žodžių ir žodžių formų kiekį (indikatorius a) moksliniai tekstai taip pat yra artimesni administraciniams (būdingesnė didesnė labai dažnų žodžių ir žodžių formų gausa);
- pagal informacijos srautą (vidutinį skirtingų žodžių ir žodžių formų kiekį tekste; žymi indikatorius MATTR) moksliniai tekstai artimesni administraciniams (būdingesnis didesnis informacijos srautas – didesnis vidutinis skirtingų žodžių ir žodžių formų kiekis);
- pagal sintaksinės struktūros kompleksiskumo rodiklį (indikatorius VA) moksliniai tekstai artimesni administraciniams (tekstams būdingesnė kompleksiškesnė sintaksė).

3.3. Publicistiniam stiliui, lyginant visus tiriamuosius funkcinis stilius, būdingiau:

- vidutiniškai trumpesni žodžiai ir žodžių formos arba, kitaip, suprantamumo prasme paprastesni tekstai;

- mažesnė labai dažnų (paprastai – funkcinių) žodžių ir žodžių formų gausa (žymi indikatorius a);
- didesnė žodžių ir žodžių formų (retesnių – priklausančių pagrindinėms kalbos dalims) įvairovė;
- didesnis žodžių ir žodžių formų koncentruotumas, t. y. mažesnis kiekis žodžių ir žodžių formų, turinčių itin didelius dažnius;
- didesnis informacijos srautas (arba vidutiniškai didesnis kiekis skirtingų žodžių ir žodžių formų);
- paprastesnė sintaksinė struktūra – mažesnis atstumas tarp 2 iš eilės einančių veiksmažodžių;
- teminės koncentracijos požiūriu publicistiniai tekstai artimesni moksliniams (būdingesnė mažesnė teminė koncentracija, kitaip – mažesnis kiekis iš esmės retesnių, pagrindinėms kalbos dalims priklausančių, tačiau moksliniuose tekstuose labai dažniems (funkciniams) dažnumu bemaž prilygstančių žodžių ir žodžių formų);
- aktyvumo (veiksmažodžių ir būdvardžių santykis tekste; žymi indikatorius A) atžvilgiu publicistiniams tekstams būdingesni dinamiškesni tekstai, nors ši tendencija nėra labai stipri, nes administracinių tekstų aktyvumo raiška mastu nedaug nusileidžia publicistiniams.

4. Nustatyti ryšiai tarp tiriamųjų požymių:

4.2 Labai stiprus ryšys visiems trimis tiriamiesiems funkciniais stiliams nustatytas tarp:

- žodžių ir žodžių formų (retesnių, priklausančių pagrindinėms kalbos dalims) įvairovės ir žodžių ir žodžių formų koncentruotumo (indikatoriai R_1 ir RR_{mc});
- dinamiškumo ir deskriptyvumo (indikatoriai A ir D).

4.3 Vidutinio stiprumo ryšiai nustatyti:

4.3.1 Administraciniam stiliui – 8 požymių poroms:

- žodžių ir žodžių formų (retesnių) įvairovės ir labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai R_1

ir a); žodžių ir žodžių formų koncentracijos ir labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai RR_{mc} ir a); informacijos srauto ir labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai MATTR ir a); informacijos srauto ir žodžių ir žodžių formų (retesnių) įvairovės (indikatoriai MATTR ir R_1); informacijos srauto ir žodžių ir žodžių formų koncentracijos (indikatoriai MATTR ir RR_{mc}); informacijos srauto ir sintaksinio kompleksiskumo (indikatoriai MATTR ir VA); dinamiškumo ir sintaksinio kompleksiskumo (indikatoriai A ir VA); deskriptyvumo ir sintaksinio kompleksiskumo (indikatoriai D ir VA).

4.3.2 Moksliniam stiliui – 10 požymių porų:

- teksto sudėtingumo ir žodžių ir žodžių formų (retesnių) įvairovės (indikatoriai VŽFI ir R_1); teksto sudėtingumo ir dinamiškumo (indikatoriai VŽFI ir A); teksto sudėtingumo ir deskriptyvumo (indikatoriai VŽFI ir D); žodžių formų (retesnių) įvairovės ir labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai R_1 ir a); dinamiškumo ir sintaksinio kompleksiskumo (indikatoriai A ir VA); dinamiškumo ir teminės koncentracijos (indikatoriai A ir TK); deskriptyvumo ir sintaksinio kompleksiskumo (indikatoriai D ir VA); deskriptyvumo ir teminės koncentracijos (indikatoriai D ir TK); informacijos srauto ir teminės koncentracijos (indikatoriai MATTR ir TK); teksto sudėtingumo ir teminės koncentracijos (indikatoriai VŽFI ir TK).

4.3.3 Publicistiniam stiliui – 11 požymių porų:

- teksto sudėtingumo ir teminės koncentracijos (indikatoriai VŽFI ir TK); teksto sudėtingumo ir dinamiškumo (indikatoriai VŽFI ir A); teksto sudėtingumo ir deskriptyvumo (indikatoriai VŽFI ir D); teksto sintaksinio kompleksiskumo ir teksto sudėtingumo (indikatoriai VA ir VŽFI); žodžių formų (retesnių) įvairovės ir labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai R_1 ir a); žodžių ir žodžių formų koncentracijos ir

labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai RR_{mc} ir a); informacijos srauto ir labai dažnų ir retesnių žodžių ir žodžių formų santykio tekste (indikatoriai $MATTR$ ir a); informacijos srauto ir žodžių ir žodžių formų (retesnių) įvairovės (indikatoriai $MATTR$ ir R_1); informacijos srauto ir žodžių ir žodžių formų koncentracijos (indikatoriai $MATTR$ ir RR_{mc}); dinamiškumo ir sintaksinio kompleksiskumo (indikatoriai A ir VA); deskriptyvumo ir sintaksinio kompleksiskumo (indikatoriai D ir VA).

Šiame darbe atliktas administracinio, mokslinio ir publicistinio funkcinų stilių kalbinės raiškos tyrimas leido ištirti šių funkcinų stilių raišką tiek kiekvieno atskirai, tiek lyginant ir priešpriešinant vieną su kitu, taip išskiriant apibendrintus tikslinių funkcinų stilių kalbinės raiškos modelius. Pasiūlyta metodika kalbinei raiškai tirti didelės apimties tekstynuose kiekybiniu aspektu tiko tiek atskirų tekstų grupių (atskirų funkcinų stilių) būdingų požymių, tiek tekstų grupių (funkcinų stilių) tarpusavio panašumų ir skirtumų pagal kiekybinių požymių raiškos mastą įvertinimui.

Kita vertus, tyrime esama ir **ribotumų**. Lietuvių kalboje tradiciškai išskiriami 5 funkciniai stiliai, o šiame tirti tik 3. Analizuojant visą funkcinų stilių sistemą sudarančių stilių kalbinės raiškos požymių dinamiką (stilių tarpusavio panašumai ir skirtumai, jų mastas) rezultatai gali šiek tiek pasikeisti. Taip pat paminėtina, kad tyrime publicistikos stilių reprezentavo naujienų portalo Delfi tekstai, kurie, nors ir įvairialypiai, nebūtinai aprėpia specializuotą, laikraščių ar regioninę publicistiką. Be to, kalbinę raišką sudaro įvairių kalbos lygmenų požymių įvairovė, o šiame tyrime naudoti tik 9, kurių nors ir pakako vieno funkcinio stiliaus atskyrimui nuo kito, tačiau neapima visos kalbinės raiškos.

Atlikus šioje disertacijoje aprašomą tyrimą manoma, kad **tekstų tyrėjams gali būti naudingi šie pastebėjimai:**

- Funkciniai stiliai yra bendros funkcinų stilių sistemos dalis, tad juos ir tirti būtų naudingiau ne po vieną, o lyginant juos tarpusavyje. Tai leistų tirti požymių raiškos mastą ir dinamiką, užfiksuoti funkcinų stilių kitimą, ryšius, jų vidinę diferenciaciją.
- Tiriant funkcinų stilių (ar kita logika grupuojamų tekstų) kalbinę raišką naudinga remtis kombinacija, o ne pavieniu požymiu. Tai leistų tiksliau įvertinti kalbinės raiškos mastą, kadangi požymiai dažnai yra susiję tarpusavyje.

Šio tyrimo rezultatai galėtų būti panaudojami automatiniam tekstų klasifikavimui ir kategorizavimui – tekstų automatinis suskirstymas pagal funkcinus stilius gali būti aktualus bibliotekoms, knygynams, kalbinių resursų (tekstynų, duomenų bazių ir pan.) planavimui ir kūrimui, internete egzistuojančių tekstų gausos įvertinimui, informacijos paieškai pagal vartotojo pateiktus kriterijus, dokumentų valdymo sistemoms (dokumentų/tekstų rekomendacijoms, dokumentų atitikimas saugumo reikalavimams ir pan.) bei panašioms tikslams. Be to, funkcinų stilių kalbinės raiškos tyrimo rezultatai gali būti panaudojami automatinio tekstų generavimo srityje, pavyzdžiui, automatiniam prekių/produktų aprašymų generavimui iš duomenų bazės įrašo ar paveikslėlio. Šio tyrimo rezultatai taip pat sudarytų prielaidas tekstų stiliaus automatiniam keitimui iš vieno į kitą – šis taikymas galėtų būtų naudingas verslui, kai vieno stiliaus tekstas, pavyzdžiui, administracinio, gali būti automatiškai konvertuotas į kito stiliaus, pavyzdžiui, publicistinio, taip taupant laiką, finansinius ir žmogiškuosius resursus.

ARTICLES ON THE TOPIC OF DISSERTATION
PUBLISHED IN PEER-REVIEWED PUBLICATIONS

DISERTACIJOS TEMA PARENGTI STRAIPSNIAI
RECENZUOJAMUOSE LEIDINIUOSE

1. Mandravickaitė Justina, Krilavičius Tomas and Man Ka Lok (2019). Document Classification to Functional Styles (Domains of Use): Lithuanian Case. *International journal of design, analysis and tools for integrated circuits and systems (IJDATICS)*. Hong Kong: Solari Co, 2019, vol. 8, no. 1 (*Proceedings of the International Conference on Recent Advancements in Computing in AI, IoT and Computer Engineering Technology (CICET 2019)*).
2. Mandravickaitė Justina and Krilavičius Tomas (2018). Quantitative Analysis of Textual Genres: Comparison of English and Lithuanian. *Proceedings of Int. Master and PhD students conference „Information Society and University Studies“ (IVUS 2018)*, p. 61—67.
3. Mandravickaitė Justina, Krilavičius Tomas and Man Ka Lok (2018). Initial Analysis of Characteristics of Textual Genres: Comparison of Lithuanian and English. *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Vol. 2.
4. Stanikūnas Daumantas, Mandravickaitė Justina and Krilavičius Tomas (2017). Comparison of Distance and Similarity Measures for Stylometric Analysis of Lithuanian Texts. *Proceedings of Int. Master and PhD students conference „Information Society and University Studies“ (IVUS 2017)*, p. 1—7.
5. Mandravickaitė Justina and Krilavičius Tomas (2017). Stylometric Analysis of Parliamentary Speeches: Gender Dimension. *Proceedings of the 6th workshop on BSNLP*, p. 102—107.

6. Mandravickaitė Justina and Oakes Michael (2016). Multiword Expressions for Capturing Stylistic Variation Between Genders in the Lithuanian Parliament. *Proceedings of the LREC 2016 workshop „CCURL 2016 – Towards an Alliance for Digital Language Diversity“*, p. 80—85.
7. Mandravickaitė Justina and Krilavičius Tomas (2015). Lietuvos Respublikos Seimo narių kalbinė raida atsižvelgiant į jų politinę orientaciją. *Darbai and dienos* 64, p. 133—151.

PAPERS ON THE TOPIC OF DISSERTATION
PRESENTED AT INTERNATIONAL ACADEMIC
CONFERENCES

TARPTAUTINĖSE MOKSLINĖSE KONFERENCIJOSE
SKAITYTI PRANEŠIMAI DISERTACIJOS TEMA

1. Mandravickaitė Justina and Krilavičius Tomas. Similarities and Differences of Lithuanian Functional Styles: a Quantitative Perspective. *The 9th International Conference „Human Language Technologies – the Baltic Perspective“ (BALTIC-HLT 2020)*, 2020 09 22-23, Kaunas.
2. Mandravickaitė Justina and Krilavičius Tomas. Document Classification to Domains of Use: Lithuanian Case. *Tarpt. konferencija „Duomenų analizės metodai programų sistemoms“ (DAMSS)*, 2019 11 28-30, Druskininkai.
3. Mandravickaitė Justina, Krilavičius Tomas and Man Ka Lok. Document Classification to Functional Styles (Domains of Use): Lithuanian Case. *The International Conference on Recent Advancements in Computing in AI, IoT and Computer Engineering Technology (CICET 2019)*, 2019 10 21-23, Sinbėjus, Taivanas.
4. Mandravickaitė Justina. Kiekybiniai indikatoriai mokslinio ir publicistinio funkcinų stilių analizėje. *5-oji taikomosios*

kalbotyros konferencija „Kalbos ir žmonės: komunikacija daugiakalbiame pasaulyje“, 2019 09 26-28, Vilnius.

5. Mandravickaitė Justina and Krilavičius Tomas. Quantitative Indicators in the Analysis of Functional Styles. *Tarpt. konferencija „Duomenų analizės metodai programų sistemoms“ (DAMSS)*, 2018 11 29-12 01, Druskininkai.
6. Mandravickaitė Justina. Kiekybiniai indikatoriai rašytinių tekstų funkcinį stilių analizėje. *25-oji tarptautinė mokslinė Jono Jablonskio konferencija „Variantiškumas kalbose ir jų atmainose“*, 2018 09 27-28, Vilnius.
7. Mandravickaitė Justina and Krilavičius Tomas. Quantitative Analysis of Textual Genres: Comparison of English and Lithuanian. *Int. Master and PhD students conference „Information Society and University Studies“ (IVUS 2018)*, 2018 04 27, Kaunas, Lithuania.
8. Mandravickaitė Justina, Krilavičius Tomas and Man Ka Lok. Initial Analysis of Characteristics of Textual Genres: Comparison of Lithuanian and English. *International MultiConference of Engineers and Computer Scientists (IMECS)*, 2018 03 14-16, Honkongas.
9. Stanikūnas Daumantas, Mandravickaitė Justina and Krilavičius Tomas. Comparison of Distance and Similarity Measures for Stylometric Analysis of Lithuanian Texts. *Int. Master and PhD students conference „Information Society and University Studies“ (IVUS 2017)*, 2017 04 28, Kaunas.
10. Mandravickaitė Justina and Krilavičius Tomas. Statistical Analysis of Word Frequency Distribution in Texts of Different Genres: Comparison of Lithuanian and English. *Tarpt. konferencija „Duomenų analizės metodai programų sistemoms“ (DAMSS)*, 2017 11 30-12 02, Druskininkai.
11. Mandravickaitė Justina and Krilavičius Tomas. Stylometric Analysis of Parliamentary Speeches: Gender Dimension. *The 6th*

Workshop on Balto-Slavic Natural Language Processing (BSNLP), 2017 04 04, Valencija, Ispanija.

12. Mandravickaitė Justina and Oakes Michael. Multiword Expressions for Capturing Stylistic Variation Between Genders in the Lithuanian Parliament. *LREC 2016 Workshop „CCURL 2016 – Towards an Alliance for Digital Language Diversity“*, 2016 05 23, Portorož, Slovėnia.
13. Mandravickaitė Justina and Krilavičius Tomas. Visualization of Differences in Language Usage of Parliamentary Position and Opposition in the Lithuanian Parliament 2008-2012. *Tarpt. Konferencija „Duomenų analizės metodai programų sistemoms“ (DAMSS)*, 2015 12 03-05, Druskininkai.

THE CANDIDATE'S BIOGRAPHICAL NOTE

TRUMPOS ŽINIOS APIE DISERTANTĄ

EDUCATION / IŠSILAVINIMAS

2015–2019 Doctoral studies in Philology; Vilnius University.

2008–2010 Master of Communication and Information; Vytautas Magnus University.

2004–2008 Bachelor of Political Sciences; Vytautas Magnus University.

PROFESSIONAL ACTIVITY/ PROFESINĖ VEIKLA

2011–present Junior Researcher; Baltic Institute of Advanced Technology.

2020–2020 Lecturer; Faculty of Philology, Vilnius University.

2015–2017 Assistant/Lecturer; Faculty of Informatics, Vytautas Magnus University.

2013–2014 Junior Researcher; Faculty of Informatics, Vytautas Magnus University.

Other activities / Kita veikla

Teaching / Dėstymas

- Introduction to Digital Humanities / Skaitmeninės humanitarikos įvadas, Vilnius University, 2020.
- Technologies and Innovations in Knowledge Professions / Viešosios komunikacijos technologijos ir inovacijos, Vytautas Magnus University, 2016.
- Internet Real-Time Process Engineering / Realaus laiko sistemų internete inžinerija, Vytautas Magnus University, 2015; 2016.

NOTES

NOTES

Vilniaus universiteto leidykla
Saulėtekio al. 9, III rūmai, LT-10222 Vilnius
El. p. info@leidykla.vu.lt,
www.leidykla.vu.lt
Tiražas 35 egz.