

Multidimensional Analysis of Linguistic Variation in the Russian Speeches Before War in Ukraine

Bahaa-eddin Abulhassan Hassan 

bahaaeldin_ali@art.sohag.edu.eg

Sohag University, Egypt

Hassan, Bahaa-eddin Abulhassan (2024). Multidimensional Analysis of Linguistic Variation in the Russian Speeches Before War in Ukraine. *Language Value*, 17(2), 100-126. Universitat Jaume I ePress: Castelló, Spain. <http://www.languagevalue.uji.es>.

December 2024

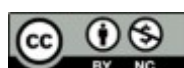
DOI: <https://doi.org/10.6035/languagev.7999>

ISSN 1989-7103

ABSTRACT

Linguistic variation is a trending topic to study genre analysis on the grounds of theory, methodology, and practicality. This study sets out to explore the genre analysis and text type of Russian speeches at the Munich Security Conferences MSC from 2007 to 2018 using multidimensional analysis. Through comparative Multi-dimensional (MD) analysis homogeneous and heterogeneous linguistic features are found. Data for the current research is taken from MSC websites by the selection of 13 Russian speeches allocated into the sub-categories of expository texts. For comparative MD analysis of sub-genre of the English version of the Russian speeches at MSC MAT tagger 1.3.3 findings were taken. The outcome indicates that there are high scores on D3 and D5 and low scores on D1 which suggests that the expository genre is utilized in the Russian speeches in MSC. The possible reasons for this linguistic variation are also discussed in the conclusion.

Keywords: Multidimensional analysis; Linguistic variation; Genre; Register; Text-type; MAT



I. INTRODUCTION

The tradition of systemic functional grammar contains some of the most sophisticated concepts on genre and register. According to this tradition, a language variety is defined by its linguistic characteristics and functionally associated with specific contextual or situational parameters of variation. Alternatively, it can be thought of as a particular configuration of field, tenor, and mode choices (in terms of Hallidayan grammar). Genre is more abstractly defined:

A genre is known by the meanings associated with it. In fact the term “genre” is a short form for the more elaborate phrase “genre-specific semantic potential” ... Genres can vary in delicacy in the same way as contexts can. But for some given texts to belong to one specific genre, their structure should be some possible realisation of a given GSP Generic Structure Potential ... It follows that texts belonging to the same genre can vary in their structure; the one respect in which they cannot vary without consequence to their genre-allocation is the obligatory elements and dispositions of the GSP. (Halliday & Hasan, 1985, p. 108)

This study retains emphasis on the practically reliable model of multidimensional analysis. The study of linguistic variations in language or texts is common in the field of Linguistics. The multidimensional analysis approach for investigating linguistic variation is a reliable method in corpus-based research. It is becoming increasingly popular for its authentic results which save time when evaluating various forms of linguistic variation (Huang, 2013). The method of MD analysis can be applied to investigate linguistic differences among language varieties, genres, registers, etc. In the interim, MD analysis has become a reliable method to study linguistic variations.

Most corpus-based studies depend directly or indirectly on the notion of genre or the related concept register and text type. Many studies mainly use the term genre (e.g. Halliday, 1978; Swales 1990, 2004; Bhatia 1993; Devitt et al., 2004; Halliday & Matthiessen, 2004; Nesi & Gardner, 2012), others use both the terms text type and genre (e.g. Biber, 1988, 1989; Stubbs, 1996). In general, studies on linguistic variation focus on a specific *genre* or a manageable set of genres, in order to limit

generalizations and linguistic features. Genre is based on external, non-linguistic, “traditional” criteria and text type is based on the interior, linguistic qualities of texts (Biber, 1988, p. 70, 170; EAGLES, 1996). According to this perspective, a genre is a conventionally recognized grouping of texts based on properties other than lexical or grammatical (co-)occurrence features. Instead, a genre is defined as a category assigned on the basis of external criteria such as intended audience, purpose, and activity type. According to Biber (1988, p. 170), “genre categories are determined on the basis of external criteria relating to the speaker’s purpose and topic; they are assigned on the basis of use rather than on the basis of form”. Paltridge (1996) considers “text types” as “discourse/rhetorical structure types,”

Table 1. *Paltridge’s Examples of Genres and “Text Types” (based on Hammond, Burns, Joyce, Brosnan & Gerot, 1992)*

Genre	Text type
recipe	procedure
Personal letter	anecdote
advertisement	description
Police report	description
Student essay	exposition
Formal letter	exposition
Formal letter	Problem-solving
News item	recount
Health brochure	procedure
Student assignment	recount
Biology textbook	report
Film review	review

As cultural structures that shift with the times, with fashion, and with social movements of ideology, genres can appear and disappear or evolve. As a result, it has been noted that several English subgenres of “official documents” have evolved recently, becoming more conversational, intimate, and familiar—sometimes purposefully, with the intention of manipulating others (Fairclough, 1992). Since genre labels are descriptions of socially constructed, functional text categories, they remain unchanged despite the fact that the genres have altered in terms of the registers invoked (a feature of intertextuality), among other modifications.

II. MULTIDIMENSIONAL ANALYSIS

A number of researches have dealt with MD analyses. Biber (1988) first introduced the MD method which incorporates quantitative and qualitative comparative analysis of various dimensions of registers. The most influential work on text typology is Biber’s (1989) work employing his factor-analysis-based multi-dimensional (MD) method; nonetheless, his categories do not appear to have been adopted by other linguists. It is asserted that the linguistic features of his eight text types—such as “informational interaction,” “learned exposition,” and “involved persuasion”—are as distinctive as possible. Biber (1993) points out that since situational characteristics of linguistic variation may be ascertained before texts are collected, it is more crucial to concentrate on covering them all as a first step in building a corpus, whereas

there is no a priori way to identify linguistically defined types ... [however,] the results of previous research studies, as well as on-going research during the construction of a corpus, can be used to assure that the selection of texts is *linguistically* as well as *situationally representative* [italics added]. (1993: 245)

The most perplexing terms are genre and register, which are frequently used synonymously due to their some degree of overlap. One distinction between the two is that, while register is associated with the organization of situation or immediate context, genre tends to be more closely tied to considerations of ideology and power

and is more closely associated with the organization of culture and social purposes around language (Bhatia, 1993; Swales, 1990). According to Crystal (1991, p. 295) *register* is as “a variety of language defined according to its use in social situations, e.g. a register of scientific, religious, formal English.”

Couture (1986) provides a new perspective on the differentiation between genre and register.

While registers impose explicitness constraints at the level of vocabulary and syntax, genres impose additional explicitness constraints at the discourse level ... Both literary critics and rhetoricians traditionally associate genre with a complete, unified textual structure. Unlike register, genre can only be realized in completed texts or texts that can be projected as complete, for a genre does more than specify kinds of codes extant in a group of related texts; it specifies conditions for beginning, continuing, and ending a text (p.82).

This article utilizes a collection of texts that are assembled as corpora to test the results of MD analysis based on Biber (1988). These speeches can be analyzed in terms of text types and are linked to power, ideological, and discourse—all of which are dynamic, negotiated elements of language use in specific contexts. Steen (1999) conceptualizes genre using the notion of basic-level categories:

It is presumably the level of *genre* that embodies the basic level concepts, whereas *subgenres* are the conceptual *subordinates*, and more abstract *classes of discourse* are the *superordinates*. Thus the *genre* of an advertisement is to be contrasted with that of a sermon, a recipe, a poem, and so on. These genres differ from each other on a whole range of attributes ... The *subordinates* of the genre of the advertisement are less distinct from each other. The press advertisement, the radio commercial, the television commercial, the Internet advertisement, and so on, are mainly distinguished by *one* feature: their medium. The *superordinate* of the genre of the ad, advertising, is also systematically distinct from the other superordinates by means of only *one* principal attribute, the one of domain: It is “business” for advertising, but it exhibits the respective values of “religious”, “domestic” and “artistic” for the other examples. [all italics added] (p. 112)

Basically, Steen suggests that we can identify genres by their cognitive basic-level status: genres are most different from one another in terms of specific attributes, whereas sub-genres, which operate on a prototype basis, have fewer distinctions among themselves. According to Steen, a genre is characterized by a set of seven attributes:

- *domain (e.g., art, science, religion, government),*
- *medium (e.g., spoken, written, electronic),*
- *content (topics, themes),*
- *form (e.g., generic superstructures, à la van Dijk (1985), or other text-structural patterns),*
- *function (e.g., informative, persuasive, instructive),*
- *type (the rhetorical categories of “narrative,” “argumentation,” “description,” and “exposition”) and*
- *language (linguistic characteristics: register/style[?]). Steen introduces a preliminary sketch of this approach to genre (and hence to a taxonomy of discourse),*

MD was also applied to different corpora. For example, Cao and Xiao (2013) study the differences between abstracts written by native British and Chinese authors in academic research articles. Gray (2013) uncovers multidimensional patterns of variation in academic research articles. Friginal and Mustafa (2017) also examine linguistic differences between abstracts published in the United States and Iraq. Gardner et al. (2019) use MD to analyze university student writing.

III. RESEARCH QUESTIONS

The study is a content analysis of the Russian speeches at MSC from 2007 till 2018. The key theme of these speeches focused on declaring the end of the unipolar world. Russian speakers advocated new multipolar world to end US hegemony. The study

argues that linguistic variation analysis could reveal the intentions of Russian policy makers to enter war in Ukraine. The present study is an attempt to find out answers to the following questions:

- 1) How far is the method of MD analysis reliable in describing linguistic variation?
- 2) What are the linguistic variations found in the Russian speeches in MSC?
- 3) What are the text types of the Russian speeches in MSC as provided by MAT analyses?

This paper suggests that it is fruitful to start by looking at the genre (categories of texts) of the Russian speeches in MSC, or even “text type” in Biber’s sense (categories of texts empirically based on linguistic characteristics) and relate it (through induction) to the existence of register (situation and discourse).

IV. RESEARCH FRAMEWORK AND METHODOLOGY

This research is a mixed method research as the quantitative and qualitative approaches were adopted to find out the Russian speeches in MSC. I chose these texts because they summarize the Russian attitude before its war on Ukraine. Biber’s (1988) theoretical model for multidimensional MD analysis was applied in the current study. MD analysis is a method that looks into linguistic variation on six dimensions by calculating co-occurring linguistic features using statistical techniques.

Table 2. *Biber’s Dimensions and text types*

Dimension	Description
1	Dimension 1 is the opposition between Involved and Informational discourse. Low scores on this variable indicate that the text is informationally dense, as for example academic prose, whereas high scores indicate that the text is affective and interactional, as for example a casual conversation. A high score on this Dimension means that the text presents many verbs and pronouns (among other features) whereas a low score on this Dimension means that the text presents many nouns, long words and adjectives (among other features).

Dimension	Description
2	Dimension 2 is the opposition between Narrative and Non-Narrative Concerns. Low scores on this variable indicate that the text is non-narrative whereas high scores indicate that the text is narrative, as for example a novel. A high score on this Dimension means that the text presents many past tenses and third person pronouns (among other features).
3	Dimension 3 is the opposition between Context-Independent Discourse and Context Dependent Discourse. Low scores on this variable indicate that the text is dependent on the context, as in the case of a sport broadcast, whereas a high score indicate that the text is not dependent on the context, as for example academic prose. A high score on this Dimension means that the text presents many nominalizations (among other features) whereas a low score on this Dimension means that the text presents many adverbs (among other features).
4	Dimension 4 measures Overt Expression of Persuasion. High scores on this variable indicate that the text explicitly marks the author's point of view as well as their assessment of likelihood and/or certainty, as for example in professional letters. A high score on this Dimension means that the text presents many modal verbs (among other features).
5	Dimension 5 is the opposition between Abstract and Non-Abstract Information. High scores on this variable indicate that the text provides information in a technical, abstract and formal way, as for example in scientific discourse. A high score on this Dimension means that the text presents many passive clauses and conjuncts (among other features).
6	Dimension 6 measures On-line Informational Elaboration. High scores on this variable indicate that the text is informational in nature but produced under certain time constraints, as for example in speeches. A high score on this Dimension means that the text presents many post modifications of noun phrases (among other features).

According to Biber (1989), each dimension is scored with positive and negative factor. In each factor there is a set of co-occurring features. The present study was done on the following six dimensions:

Table 3. *Biber's (1989) text types*

Text Type	Characterizing Genres	Characterizing Dimensions	Description
Intimate interpersonal interaction	telephone conversations between personal friends	high score on D1, low score on D3, low score on D5, unmarked scores for the other Dimensions	Texts belonging to this text type are typically interactions that have an interpersonal concern and that happen between close acquaintances

Text Type	Characterizing Genres	Characterizing Dimensions	Description
Informational Interaction	face-to-face interactions, telephone conversations, spontaneous speeches, personal letters	high score on D1, low score on D3, low score on D5, unmarked scores for the other Dimensions	Texts belonging to this text type are typically personal spoken interactions that are focused on informational concerns
Scientific Exposition	academic prose, official documents	low score on D1, high score on D3, high score on D5, unmarked scores for the other Dimensions	Texts belonging to this text type are typically informational expositions that are formal and focused on conveying information and very technical
Learned Exposition	official documents, press reviews, academic prose	low score on D1, high score on D3, high score on D5, unmarked scores for the other Dimensions	Texts belonging to this text type are typically informational expositions that are formal and focused on conveying information
Imaginative Narrative	romance fiction, general fiction, prepared speeches	high score on D2, low score on D3, unmarked scores for the other Dimensions	Texts belonging to this text type are typically texts that present an extreme narrative concern
General Narrative Exposition	Press reportage, press editorials, biographies, non-sports broadcasts, science fiction	Low score on D1, high score on D2, unmarked scores for the other Dimensions	Texts belonging to this text type are typically texts that use narration to convey information
Situated Reportage	Sports broadcasts	low score on D3, low score on D4, unmarked scores for the other Dimensions	Texts belonging to this text type are typically on-line commentaries of events that are in progress
Involved Persuasion	spontaneous speeches, professional letters, interviews	high score on D4, unmarked scores for the other Dimensions	Texts belonging to this text type are typically persuasive and/or argumentative

Data for the present research were taken from the Russian speeches at MSC. In the present research linguistic variation only in the Russian speeches in MSC was investigated. The current study is limited to the findings of MAT tagger 1.3.3.

V. MULTIDIMENSIONAL ANALYSIS TAGGER

The Multidimensional Analysis Tagger (MAT) is a computer program created by Nini (2019) to replicate Biber's (1988) tagger for the MD analysis of English texts. It is generally utilized in studies on text type or genre variation. It generates a grammatically annotated version of the selected text or the corpus as well as the statistics needed to perform a text-type or genre analysis. It reproduces the input text or corpus according to the linguistic features used in Biber's (1988) Dimensions and it displays its closest text type, as proposed by the patterns described in Biber (1989). It offers a tool for visualising the Dimensions features of an input text.

For example, the program adds tags to linguistic items to facilitate the identification of Biber's (1988) linguistic features, e.g. the word 'to' used as an infinitive marker is differentiated from the preposition *to*. Examples of tags are: (1) indefinite pronouns (INPR): *anybody, anyone, anything, everybody, everyone, everything, nobody, none, nothing, nowhere, somebody, someone, something*; (2) quantifiers (QUAN): *each, all, every, many, much, few, several, some, any*; (3) quantifier pronouns (QUPR): *everybody, somebody, anybody, everyone, someone, anyone, everything, something, anything*. A full list of variables and tags is given below.

Next to the name of the variable is the tag used by the present tagger to identify it. An asterisk appears next to the name of the variables for which Biber (1988) manually checked the results. The present version of the tagger does not allow any manual intervention in the tagging process. However, the texts can be manually checked before the analysis takes place.

Past tense (VBD)

Perfect aspect (PEAS)

Present tense (VPRT)

Place adverbials (PLACE)

Time adverbials (TIME)

First person pronouns (FPP1)

First person pronouns (FPP1)

Third person pronouns (TPP3)

Pronoun it (PIT)

Demonstrative pronouns (DEMP)

Indefinite pronouns (INPR)

Pro-verb do (PROD)

Direct WH-questions (WHQU)

Nominalizations (NOMZ)

Gerunds (GER)

Total other nouns (NN)

Agentless passives (PASS)

By-passives (BYPA)

Be as main verb (BEMA)

Existential there (EX)

That verb complements (THVC)

That adjective complements (THAC)

WH-clauses (WHCL)

Infinitives (TO)

Present participial clauses (PRESP)

Past participial clauses (PASTP)

Past participial WHIZ deletion relatives (WZPAST)

Present participial WHIZ deletion relatives (WZPRES)

That relative clauses on subject position (TSUB)

That relative clauses on object position (TOBJ)
WH relative clauses on subject position (WHSUB)
WH relative clauses on object position (WHOBJ)
Pied-piping relative clauses (PIRE)
Sentence relatives (SERE)
Causative adverbial subordinators (CAUS)
Concessive adverbial subordinators (CONC)
Conditional adverbial subordinators (COND)
Other adverbial subordinators (OSUB)
Total prepositional phrases (PIN)
Attributive adjectives (JJ)
Predicative adjectives (PRED)
Total adverbs (RB)
Type-token ratio (TTR)
Word length (AWL)
Conjuncts (CONJ)
Downtoners (DWNT)
Hedges (HDG)
Amplifiers (AMP)
Emphatics (EMPH)
Discourse particles (DPAR)
Demonstratives (DEMO)
Possibility modals (POMD)
Necessity modals (NEMD)

Predictive modals (PRMD)

Public verbs (PUBV)

Private verbs (PRIV)

Suasive verbs (SUAV)

Seem/appear (SMP)

Contractions (CONT)

Subordinator that deletion (THATD)

Stranded preposition (STPR)

Split infinitives (SPIN)

Split auxiliaries (SPAU)

Phrasal coordination (PHC)

Independent clause coordination (ANDC)

Synthetic negation (SYNE)

Analytic negation (XX0)

The reliability of the program was tested on the Lancaster-Oslo/Bergen (LOB) Corpus, a one-million-word collection of British English texts compiled in the 1970s. Results indicate that MAT is mainly successful in replicating Biber's (1988) analysis. When MAT on the Brown corpus, a one million words (500 samples of 2000+ words each) of running text of edited English prose, significant differences can be observed between MAT scores and Biber's (1988) scores. However, MAT results on the Brown corpus indicate that the Dimensions found by Biber (1988) are still valid for those genres even when taking into account different texts. Results suggest that MAT can be used to analyze texts according to scores. Furthermore, it can also categorise a text for its text type, as provided by Biber (1989). AntConc is a useful tool in MAT for finding clusters (frequency patterns of word sequences of words within corpus or selected texts and calculating type token ratio (TTR). TTR is the ratio obtained by dividing the types (the

total number of different words, occurring in a text by its token (the total number of words). A high TTR indicates a high degree of linguistic variation while a low TTR indicates the opposite.

VI. RESULTS AND DISCUSSION

In this section MD analysis of the Russian speeches at MSC is made on the basis of MAT results. The following table illustrates the ratios of the six dimensions.

Table 4. *Dimension ratios*

Dimension	Ratio
1	-14.03
2	-1.9
3	5.82
4	2.51
5	2.06
6	0.78
Closest Text Type	Learned exposition

Table (5) shows the ratios of the variables in the Multidimensional Analysis tagger.

Table 5. *Variables ratios*

Variable	Score
Tokens	28556
AWL	5.07
TTR	234

Variable	Score
AMP	0.16
ANDC	0.57
BEMA	1.48
BYPA	0.15
CAUS	0.07
CONC	0.03
COND	0.18
CONJ	0.40
CONT	0.05
DEMO	1.20
DEMP	0.57
DPAR	0.02
DWNT	0.22
EMPH	0.47
EX	0.23
FPP1	2.85
GER	0.32
HDG	0
INPR	0.04
JJ	9.09
NEMD	0.31

Variable	Score
NN	22.00
NOMZ	4.86
OSUB	0.13
PASS	0.92
PASTP	0.07
PEAS	0.74
PHC	1.17
PIN	12.44
PIRE	0.09
PIT	1.00
PLACE	0.25
POMD	0.38
PRED	0.76
PRESP	0.21
PRIV	1.18
PRMD	0.80
PROD	0.08
PUBV	0.57
RB	3.19
SERE	0.30
SMP	0.05
SPAU	0.56

Variable	Score
SPIN	0.07
SPP2	0.18
STPR	0.04
SUAV	0.49
SYNE	0.17
THAC	0.11
THATD	0.09
THVC	0.41
TIME	0.51
TO	2.10
TOBJ	0.13
TPP3	0.63
TSUB	0.20
VBD	1.27
VPRT	5.22
WHCL	0.07
WHOBJ	0.01
WHQU	0.06
WHSUB	0.09
WZPAST	0.28
WZPRES	0.18
XX0	0.82

A measure of how many standard deviations below or above the mean or raw score is called z-score. It will be positive if the value lies above the mean and negative if it lies below the mean.

Table 6. *Z-score*

Variable	Score
AMP	-0.42
ANDC	0.25
AWL	1.43
CAUS	-0.24
CONC	-0.25
COND	-0.32
CONJ	1.75
DEMO	0.50
DEMP	0.23
DPAR	-0.43
DWNT	0.12
EMPH	-0.38
EX	0.06
FPP1	0.05
GER	-1.00
HDG	-0.46
INPR	-0.50
JJ	1.61

Variable	Score
NEMD	0.48
NN	1.11
NOMZ	1.99
OSUB	0.27
PHC	3.07
PIN	0.55
PIT	-0.04
PLACE	-0.18
POMD	-0.57
PRED	1.12
PRMD	0.57
RB	-1.91
SPP2	-0.59
SYNE	0.00
THAC	1.33
THVC	0.28
TIME	-0.03
TO	1.09
TOBJ	0.45
TPP3	-1.05
TSUB	2.00
TTR	1.42

Variable	Score
VBD	-0.90
VPRT	-0.74
XX0	-0.05
BEMA	-1.42
BYPA	0.54
CONT	-0.70
PASS	-0.06
PASTP	1.50
PEAS	-0.23
PIRE	0.18
PRESP	0.65
PRIV	-0.60
PROD	-0.63
PUBV	-0.37
SERE	7.25
SMP	-0.30
SPAU	0.04
SPIN	7000.00
STPR	-0.59
SUAV	0.65
THATD	-0.54
WHCL	0.10

Variable	Score
WHOBJ	-0.76
WHQU	0.67
WHSUB	-0.60
WZPAST	0.10
WZPRES	0.11
Underused variables	none
Overused variable	PHC [SERE] [SPIN]

The following figure indicates the low score (-14.3) of Dimension 1. It means the speeches are informationally dense.

Figure 1. Dimension 1

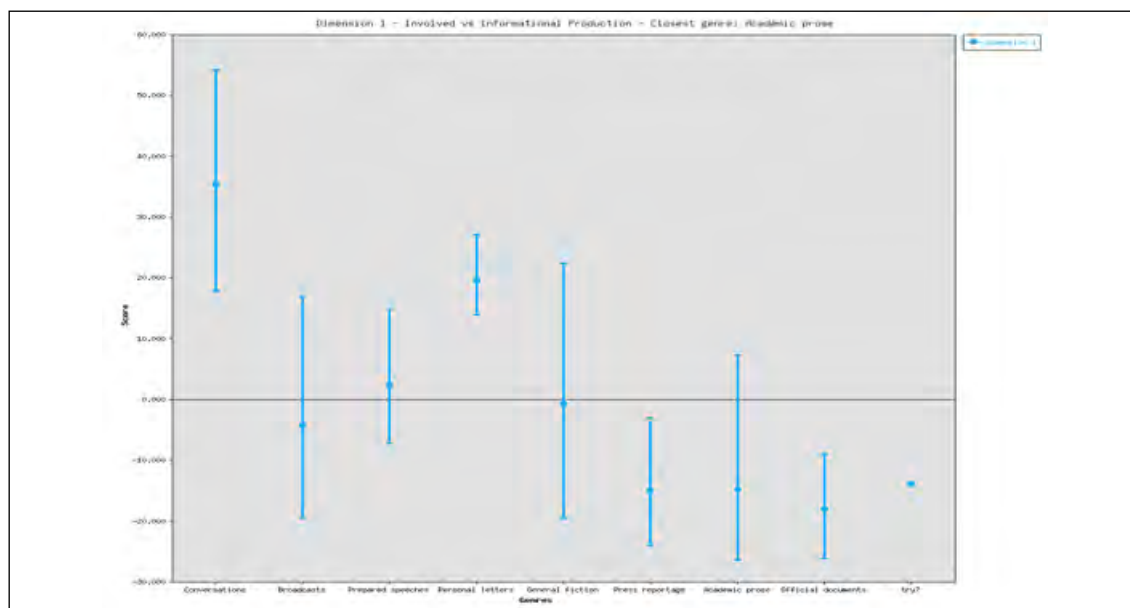
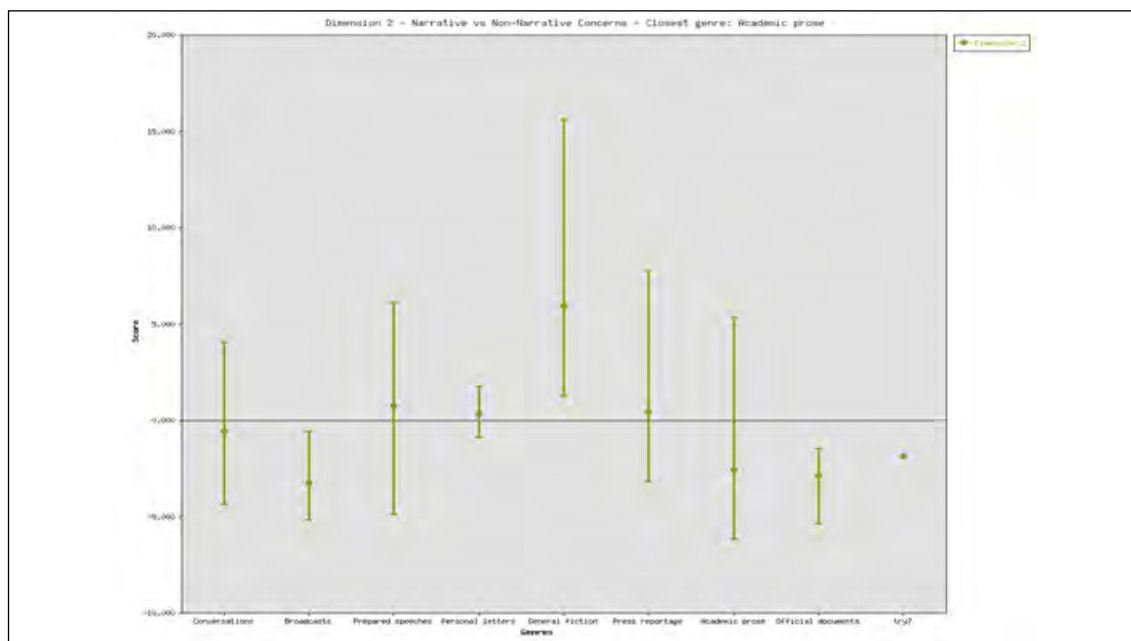


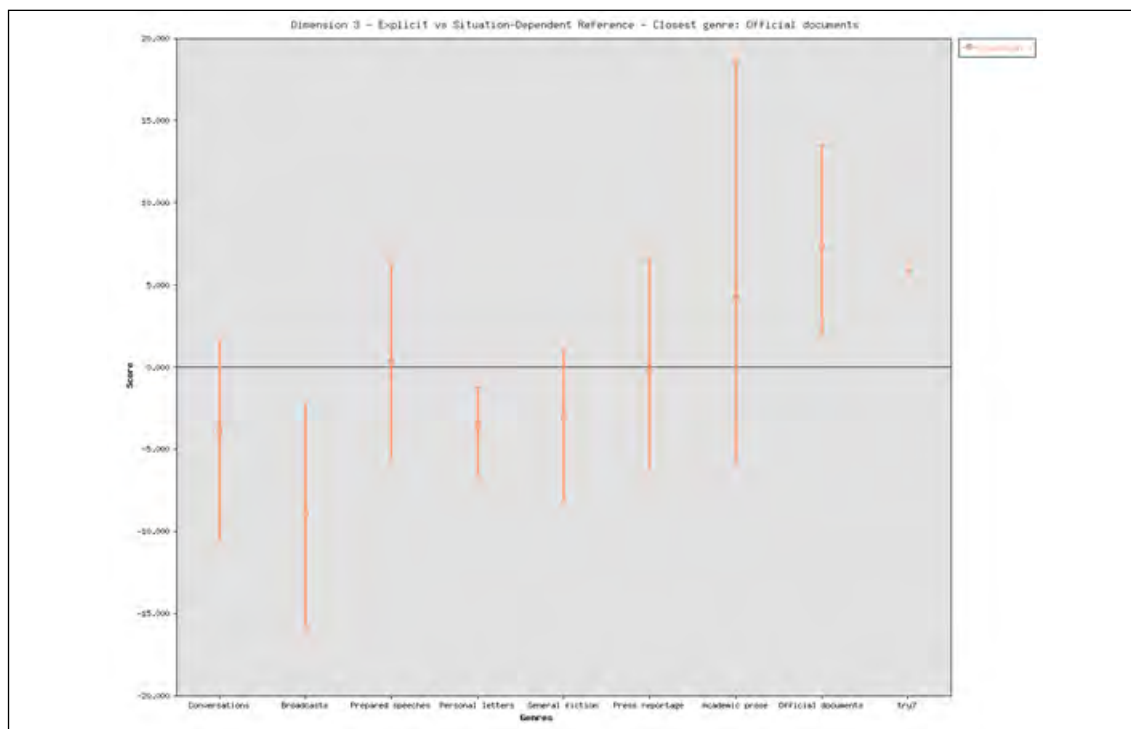
Figure (2) indicates a low score (-1.9) of Dimension 2. This low score is a feature of non-narrative speeches.

Figure 2. Dimension 2



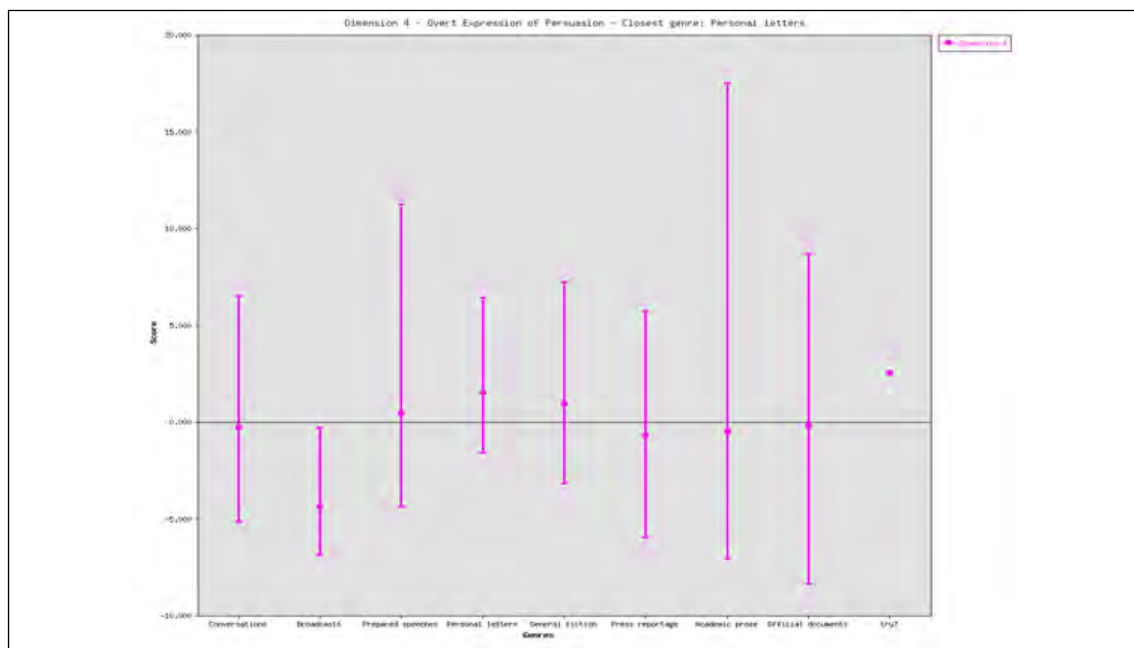
Dimension 3 in Figure (3) indicates a high score (5.82). This means that the speeches are not dependent on context, as if it is close to academic prose.

Figure 3. Dimension 3



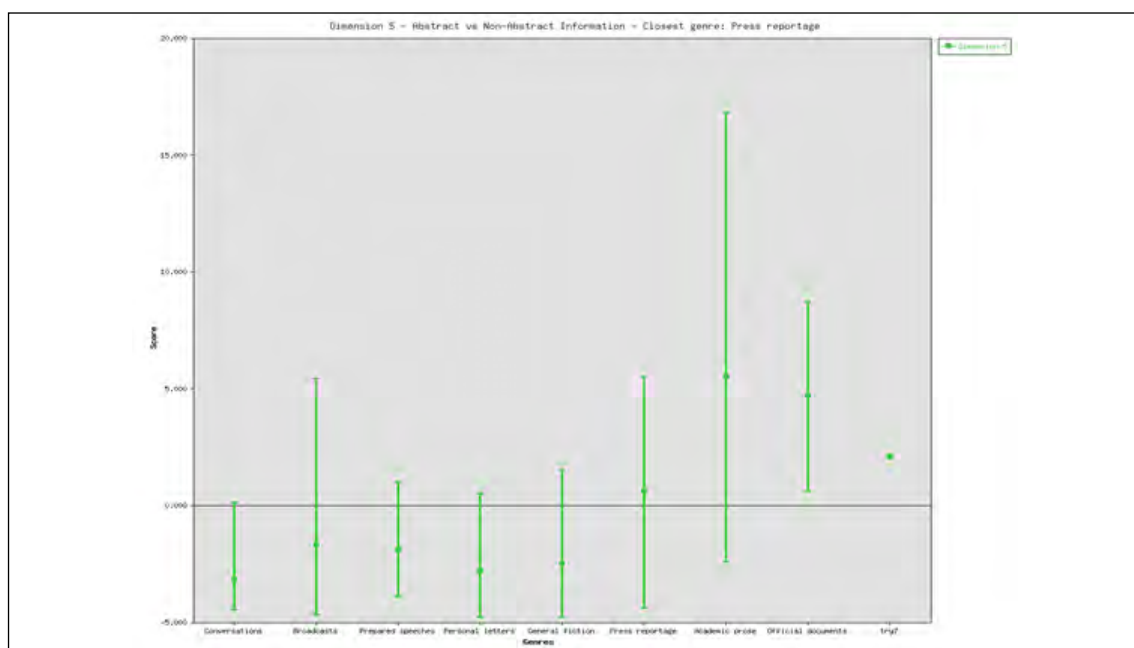
In the following figure Dimension 4 indicates a high score on this variable. It shows that the speeches explicitly mark the speakers' points of view. Texts belonging to this text type are typically persuasive in the Russian speeches.

Figure 4. Dimension 4



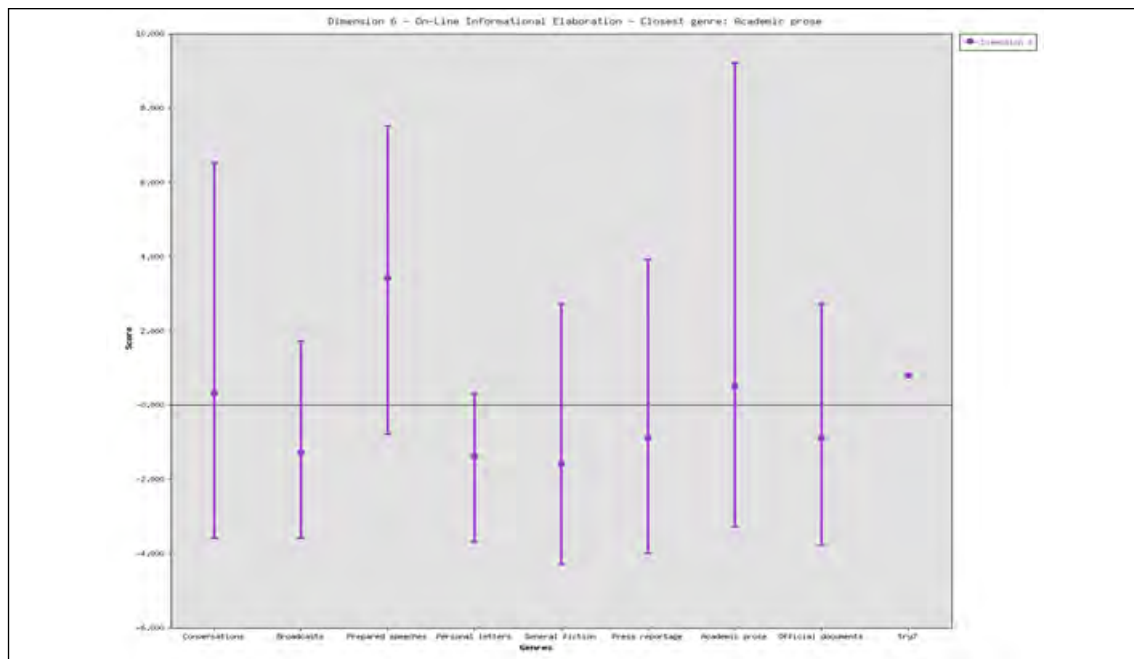
Dimension 5 in Figure (5) is measured (2.06). This indicates that the speeches provide information in a technical, abstract and formal way.

Figure 5. Dimension 5



In the following figure Dimension 6 is (0.78). This indicates that the speeches are informational.

Figure 6. Dimension 6



The following figure indicates that the text type of the Russian speeches is learned exposition. The whole features of the dimensions indicate that the Russian leaders chose to use a text type close to academic prose.

Figure 7. Text Type



The study argues that the linguistic variations found in the Russian speeches in MSC are numerous and include six dimensions according to MD analysis. Thus, the model of multidimensional analysis has proven reliable in determining the text type used in the Russian speeches. As results indicate, learned exposition is the text type of the Russian speeches in MSC as provided by MAT analysis.

VII. CONCLUSION

Ten years before the launch of the 2022 Russian invasion of Ukraine, the Russian speeches at the Munich Security Conferences were revealing moments of Putin's later intentions. Expository genre has been used to inform Western leaders about intentions to wage war on Ukraine and legitimize it. Putin's speech in 2007 was an implicit declaration of war and Western leaders fail to recognize the situational and discourse analyses of the speeches.

Russian speeches are informational as there is Low scores on D1. This variable indicates that the text is informationally dense. **D3** a high score indicate that the text is not dependent on the context, as for example academic prose. A high score on this Dimension means that the text presents many nominalizations. This type of discourse creates conciseness, objectivity, formality and thematic connections which help speakers achieve their goals. **D5** High scores on this variable indicate that the text provides information in a technical, abstract and formal way, as for example in scientific discourse. A high score on this Dimension means that the text presents many passive clauses and conjuncts (among other features).

REFERENCES

- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Longman.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

- Biber, D. (1989). A typology of English texts. *Linguistics*, 27(1), 3–43.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243–257
- Cao Y., Xiao R. (2013). A multidimensional contrastive study of English abstracts by native and non-native writers. *Corpora*, 8(2), 209–234
- Couture, B. (1986). Effective ideation in written text: A functional approach to clarity and exigence. In B. Couture (ed.), *Functional approaches to writing: Research perspectives* (pp. 69–91). Ablex.
- Crystal, D. (1991). *A Dictionary of Linguistics and Phonetics*. Oxford, UK: Basil Blackwell.
- Devitt, A. J., Reiff, M. J., & Bawarshi, A. (2004). *Scenes of writing: Strategies for composing with genres*. Longman.
- EAGLES (Expert Advisory Group on Language Engineering Standards). (1996, June). *Preliminary recommendations on text typology*. EAGLES Document EAG-TCWG-TTYP/P. [Available at <http://www.ilc.pi.cnr.it/EAGLES96/texttyp/texttyp.html>]
- Fairclough, N. (1992). *Discourse and social change*. Cambridge, UK: Polity Press.
- Friginal, E. & Mustafa, S. S. (2017). A comparison of U.S.-based and Iraqi English research article abstracts using corpora. *Journal of English for Academic Purposes*, 25, 45–57
- Garner, S., Nesi, H. & Biber, D. (2019). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics*, 40(4), 646–674.
- Gray, B. (2013). More than discipline: Uncovering multi-dimensional patterns of variation in academic research articles. *Corpora*, 8 (2), 153–181.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.

- Halliday, M. A. K., & Hasan, R. (1985). *Language context and text: Aspects of language in a social semiotic perspective*. Oxford University Press.
- Halliday, M. A. K., & Matthiessen, C. (2004). *An introduction to functional grammar* (3rd ed.). Edward Arnold.
- Hammond, J., Burns, A., Joyce, H., Brosnan, D., & Gerot, L. (1992). *English for social purposes: A handbook for teachers of adult literacy*. National Centre for English Language Teaching and Research, Macquarie University.
- Huang, C.-T. James. (2013). Variations in non-canonical passives. In A. Alexiadou & F. Schaefer (eds.), *Non-canonical passives*, 95-114. John Benjamins.
- Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.
- Nini, A. (2019). The Multi-Dimensional Analysis Tagger. In T. Berber Sardinha & M. Veirano Pinto (eds), *Multi-Dimensional Analysis: Research Methods and Current Issues* (pp. 67-94). Bloomsbury Academic.
- Paltridge, B. (1996). Genre, text type, and, and the language classroom. *ELT Journal*, 50 (3), 237–243.
- Steen, G. (1999). Genres of Discourse and the Definition of Literature. *Discourse Processes*, 28 (2), 109–120.
- Stubbs, M. (1996). *Text and corpus analysis*. Oxford: Blackwell.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Swales, J. M. (2004). *Research genres*. Cambridge University Press.

Received: 30 March 2024

Accepted: 23 December 2024