Markers of Ethnocultural Identity in Medical Research Papers

Irina Torubarova[1*], *Anna* Stebletsova[1], *Tanya* Linaker[2]

[1]Foreign Languages Department, Voronezh N.N. Burdenko State Medical University, 10 Studencheskaya Street, Voronezh 394036, Russian Federation
[2] Modern Language Centre, King's College London, UK

Abstract. Cross-cultural differences in academic discourse as a field of applied linguistic research became mainstream at the end of the 20th century. Two fundamental concepts emerged at that period: one highlighted universal features of academic discourse, the other focused on ethnocultural peculiarities of cognitive and textual structures. Research in the field of academic discourse becomes particularly significant when applied to training and education of academic staff. Accepting the idea of universal features of academic discourse, we, however, tend to support the concept claiming that academic writing style may contain markers of ethnocultural identity of their authors. So, the aim of our research was to identify markers of ethnocultural identity in medical research texts. The study included comparative discourse analysis of research focused medical texts created by English-speaking (ES) and Russian-speaking (RS) authors. The results obtained have drawn us to the conclusion that despite universal features that scientific medical texts of the same genres share, they nevertheless manifest specific properties depending on the linguistic and cultural affiliation of their authors, i.e. markers of ethnocultural identity. This knowledge can contribute to international academic communication.

1 Introduction

Cross-cultural differences in academic discourse as a research topic became relevant at the end of the 20th century. Two fundamental concepts emerged in that period: one highlighted universal features of academic discourse, the other focused on cultural identity of cognitive and textual structures [1].

There have been a number of studies that support the following assumption: the differences in academic style may be associated with the author's national and cultural identity. Kaplan [2] compared translations from different languages into English, as well as English-language texts created by authors of different nationalities, emphasising the fact that, despite the expected "linear pattern" of constructing an English paragraph, there were distinctive features in the structure of the paragraph arrangement and development in the texts of native and non-native English-speaking authors. Kaplan has categorized the

---

* Corresponding author: torubarova69@mail.ru

identified features as five types of text construction, namely, English, Russian, Semitic, Oriental and Roman [2].

In another study Clyne examined the problem of the degree with which national identity manifested itself in academic texts. The author addressed the question whether this manifestation could affect the writer's performance in international academic communication [3].

Academic discourse is a multidimensional interdisciplinary phenomenon that has been studied by linguists both in Russia and abroad [3-7]. Findings from this continuing research can be a significant contribution to teaching English for Specific Purposes (ESP) and English for Academic Purposes (EAP) to Russian speakers including the academic community who participate in international exchange.

Medical research discourse as part of a global academic domain has been broadly investigated recently. The genres of medical discourse are analysed from various aspects: structure - with the emphasis on language description and use in a professional context; genre variation - in terms of genre evolution and a particular genre use in cross-cultural and cross-linguistic contexts; communication - with an emphasis on the doctor-patient relationship etc. [4; 5; 8; 9; 6]

A substantial bulk of academic research is closely connected with higher medical education, which has to satisfy increasing requirements for the qualification and competence of healthcare professionals and researchers. There is an obvious demand for academic writing skills for medical researchers who regularly contribute to international publications. The practice of international academic exchange demonstrates that non-native speakers of English can encounter certain challenges which may partly be attributed to national differences in their academic styles. To address these challenges, Russian academic discourse scholars focus on various aspects of medical research papers created by native and non-native English speakers [10-12]. The present paper is in the framework of this research. Sharing the idea of universal features of academic discourse, we also support the concept of national identity in academic discourse and argue that cultural identity markers may be manifested in a research text.

The aim of this study is to identify markers of ethnocultural identity in medical research texts. This will be done by comparative analysis of the two corpora of medical research texts created in English by the Russian-speaking (RS) and English-speaking (ES) authors.

2   Materials and methods

The corpus of study consisted of full-text articles published in journals included in the International databases. The journals are published in English. They cover a wide spectrum of topics and subjects, e.g. cardiology, pulmonology, surgery, etc. All articles were original research, had an "Introduction, Method, Results and Discussion" (IMRaD) structure and are open access papers. These criteria correspond to the universal genre format and stylistic norms established and applied to writing research articles in recent decades.

Article titles, figures, tables, diagrams and their legends were excluded from the research analysis.

The authors' attribution to English-speaking (ES) or Russian-speaking (RS) groups was applied according to their personal information and affiliation with medical universities or research institutions at the moment of publication of their articles.

We used descriptive analysis and statistical research methods for text examination and interpretation. For statistical processing, the Compleat Lexical Tutor v.8.3 program was employed.

3 Results

The accurate selection and examination of national identity markers in both corpora can only be done according to relevant criteria accounting for academic discourse features. Such criteria for quantitative analysis were type-token ratio, lexical density, and lexical coverage parameters, which were calculated for English Native Speaking (NS) and Russian NS corpora. Type-token ratio is defined as the proportion of different words or words from different word families used in a text to all words in a text. [13] «Lexical diversity is a measure of how many different words are used in a text, while lexical density provides a measure of the proportion of lexical items (i.e. nouns, verbs, adjectives and some adverbs) in the text» [13, p. 61]. "Lexical coverage" is defined as the percentage of words that characterized the "apprehensibility" of the text for the reader [14; 15]. This criterion is determined by comparing the words used in the texts of articles with a General Word List (GWL) and an Academic Word List (AWL).

Qualitative analysis criteria were selected through a set of parameters which reflect formality as a basic stylistic characteristic of academic writing [7]. These parameters covered terminological vocabulary, including abbreviations and acronyms, abstract nouns and noun groups, linguistic and stylistic repetitions, 1st person plural, complex sentences, passive constructions and cohesion phrases. Terms, abstract nouns, noun groups, sentences containing passive voice and cohesion phrases were selected and statistically processed. There was a statistical analysis of 1st person plural usage which was also calculated for each corpus.

3.1 Quantitative analysis findings

The results of the quantitative analysis demonstrate that the type-token ratio was higher in the texts created by Russian authors: the lowest parameter for RS texts was 0.25, the highest was 0.31; for ES texts these parameters were 0.17 and 0.28, respectively (see Table 1). This data demonstrates that Russian authors' texts were lexically more diverse, contained a wide range of various vocabulary, including synonymic lexical items, which are commonly used to avoid repetition (Table 1).

Lexical density was similar in texts of both corpora (Table 1). This meant that the texts used lexical items belonging to particular group of words or word families, and the texts mainly consisted of content words.

Lexical coverage is measured in frequency level. In ES corpus this indicator ranged from K-4 to K-12. This data means that the reader should know between 4,000 to 12,000 words to understand 60 per cent of the text. The same parameter in RS corpus is significantly higher and ranges from K-12 to K-19 frequency level. This statistical difference demonstrates that RS texts are more lexically complicated than ES texts. The table below presents these findings (Table 1).

Table 1. Quantitative analysis of research texts of the medical corpus.

	ES texts	RS texts
Type-token ratio	0.17 min 0.28 max	0.25 min 0.31 max
Lexical density	0.64 min 0.70 max	0.63 min 0.71 max
Lexical "coverage"	K-4 frequency level - min K-12 frequency level - max	K-12 frequency level – min K-19 frequency level – max
1st person pronouns	0.01% min 2.2% max	0.05% min 0.6% max
Passive sentences	21% min 53% max	26.7% min 69.4% max

3.2 Qualitative analysis findings

Formality of the academic style is manifested through a number of text markers at all levels of the language system. The vocabulary markers, which generally cover terms, including abbreviations and acronyms, noun groups and verbal nouns were represented in both ES and RS corpora. Here are some examples below:

(*1*)   *AGEs - Advanced glycation end products, SAF - Skin autofluorescence, 25(OH)D 3 - 25-hydroxyvitamin D 3, eGFR - Estimated glomerular filtration rate, dementia risk prediction models (ES).*
(*2*)   *Chronic obstructive pulmonary disease (COPD), final TBA-active LPO products, the development of OS in patients with AH, in the regulation of BP, the presence of such complications as typhlitis (RS);*

It should be emphasized that RS corpus demonstrates a regular use of multi-word noun groups with an *of-nouns* chain to modify the head noun, which is a characteristic feature of academic texts created in Russian. The regular use of *of-noun* groups for modifying the head noun in RS corpus of the present study can be interpreted as a marker of ethnocultural identity of Russian authors.

The comparison of 1st person pronouns in ES and RS corpora (2.2% and 0.6% respectively) shows that RS authors were less likely to use personal pronouns in the subject position to express the active role of the researcher. On the contrary, the proportion of passive voice used by RS authors was higher compared to the ES corpus (69.4 % and 53.0 % respectively). The data is given in Table 1. Here are some examples below:

*(3) **We** investigated the effects on PPV and the numbers of cases ascertained by implementing additional criteria. (ES).*
*(4) The asthma diagnosis **was based** on the integral assessment of symptoms. (ES)*
*(5) **We** analyzed complaints, anamnesis data, objective status data, and laboratory and instrumental data (spirometry with a 400-mg salbutamol test), and **we** measured levels of NPY, adiponectin, leptin. (RS)*
*(6) The age and sex distribution of patients **is shown** in Table 1. (RS)*

Another significant marker of cultural differences in compared corpora was lexical and syntactic repetitions. Lexical repetition is defined as the repetition of a word, phrase or sentence within one statement (sentence, paragraph) and in larger units of communication.

[16] Syntactic repetition is based on repetition of the syntactic structure - a phrase or a sentence. From a pragmatic point of view syntactic repetition or parallelism in the academic discourse functions as an effort saving device: the author can be clear and concise using similar syntax, which is beneficial to the reader, making the text reader-friendly.

The results of our analysis showed that lexical and syntactic repetitions, which are a typical property of the academic style, were less commonly used by the RS authors compared to ES authors. This can be explained by the RS authors' attempt to avoid tautology and diversify the text lexically and syntactically which is consistent with the norms of the Russian discourse. The RS corpus demonstrated the use of synonymous verbs when denoting the same action within the same sentence and paragraph. This can be considered a clear marker of Russian identity in the academic discourse. The following examples illustrate the described phenomenon:

(7) *According to anthopometrical reference data, 30/16.8% patients* **had normal weight**; *90/50.6% patients* **were overweight***; 49/27.5% patients* **had first-degree obesity***. (RS)*

(8)   **For all-cause dementia***, Alzheimer's disease and vascular dementia, we calculated the PPV for each dataset separately, and for all three combined.*

For all-cause dementia*, true positive cases were those where the adjudicator recorded dementia as being present, with or without meeting diagnostic criteria.*

For dementia subtypes*, cases were true positives if the adjudicator indicated that a particular subtype diagnosis could be made, with or without meeting the particular diagnostic criteria. (ES)*

The next comparison criterion concerned complex sentences, relative clauses, participial/gerund phrases, homogeneous parts of the sentence, which are typical markers of academic style. We have observed a large number of complex sentences in both RS and ES corpora, e.g.:

(9) *It has been confirmed that the complete CC of the BB from RA to the ejection from the LV has a length equal to two completed heart contraction cycles. Interaction of venous and arterial boluses, with differentiated external myocardial exposure, along the path "venous block of the heart–lung–arterial block of the heart," forms averaged (compensated by the flexible septum) variable pressure values between the stages of intracardiac routes of BBs (unidirectional, synchronous, but spread in the space. (RS, 79 words)*

(10) *A key unanswered question is whether the association between higher BW and increased breast cancer risk seen in observational studies is largely due to the BW-increasing effect of nongenetic factors such as maternal hormones and nutrition and we are unable to rule out an association between the nongenetically-determined component of BW and breast cancer. (ES, 55 words)*

It should be noted, though, that the longest sentences in the texts created by RS contained more words than the longest sentences in the texts created by ES (79 words and 55 words respectively). This may be due to the traditional Russian perception of an academic style which should normally be complex and contain deliberately complicated sentences.

Linking words and phrases are essential components of an academic text and most crucial means of discourse cohesion and coherence. The comparative analysis shows that linking words and phrases were broadly used by both ES and RS authors, e.g.:

(11) *In conclusion, we have estimated the PPV of using UK routinely-collected healthcare datasets to identify cases of all-cause dementia, Alzheimer's disease and vascular dementia during follow-up in large, prospective studies in the UK. (ES)*

(12) *Briefly, between 2006 and 2010, UKB recruited 500,000 participants aged 40–69 years. (ES)*

(13) *Thus, in ISIAH rats there was an increase in toxic TBA-active products and a decrease in TTA, α-tocopherol and retinol in comparison with normotensive animals. (RS)*

(14) On the other hand, adiponectin plays an opposite role, inhibiting inflammation. (RS)

The results of linking words' usage in the RS corpus show little diversity: the most frequently used word was "thus". For example, in the article about peroxidation it was used 5 times within one paragraph in the section "Results". We have recorded one example of a subjective linking phrase "In our opinion", which is not typical for academic writing and was not recorded in the ES corpus, as seen in the illustration below.

(15) In our opinion, CMIP is a nonlinear, multidimensional regulatory structure of the heart as a whole. (RS)

4   Conclusion

A comparative analysis of English texts created by English and Russian speaking authors is a relevant way to identify markers of ethnocultural identity of academic style, since it is based on universal quantitative and qualitative criteria of academic discourse. All analysed texts of both corpora were comparable in genre, stylistic properties and formal parameters.

The analysis demonstrated that although there is similar usage of terminology, passive voice and complex sentences, there are several markers in RS in academic style features, which are different to ES, namely:
• RS texts were lexically more diverse and significantly more complicated for the reader in terms of perception;
• RS texts demonstrated a dominant use of 'of-noun' phrases as a means of modifying the head noun in a noun phrase;
• RS authors tended to overuse passive voice as a means of making their language objective;
• RS authors tended to avoid lexical and syntactic repetition and used synonyms to identify actions, events and phenomena;
• RS authors used longer and more sophisticated sentences which might be a transfer from the Russian academic style.

These ethnocultural markers have clearly shown that alongside the universal features that scientific medical texts of the same genres share, these texts nevertheless manifest specific or ethnocultural properties depending on the linguistic and cultural affiliation of their authors. The awareness of ethnocultural identity and its manifestation in academic discourse can contribute to international academic communication.

Acknowledgments

References

1.   D. Siepmann, Academic Writing and Culture: An Overview of Differences between English, French and German. *Meta*, 51 (1), pp. 131–150 (2006)

2.   R. B. Kaplan, "Cultural Thought Patterns in Intercultural Education," Croft, Kenneth (ed.) Readings on English as a Second Language Cambridge, Winthrop, pp. 399-418 (1980)

3.   M. Clyne, Cultural differences in the organization of the academic texts. Journal of pragmatics, 11, pp. 211-247 (1987)

4. A.O. Stebletsova, Meta-professional business discourse: typology and national specifics (2017) (In Russ.)

5. A.O. Stebletsova, Pragmalinguistic approach to discourse description: a model of descriptive analysis and application practice. *Bulletin of Voronezh State University. Series: Philology*. Journalism, №1. pp. 90-93 (2014) (In Russ.)

6. M. Gotti, Insights into medical discourse: diachronic and synchronic perspectives. *ESP today*, Vol. 3(1), pp. 5-24 (2015):

7. A.O. Stebletsova, I.I. Torubarova, Formality as a Style-Forming Property of Academic English: Synchronic Analysis of Medical Research Texts. *Science Journal of Volgograd State University. Linguistics*, vol. 18, no. 4, pp. 157-173 (2019) (in Russ.).

8. E.V. Vinogradova, I.Y. Kolesnikova, Reflection of the peculiarities of national mentality in medical discourse. *Medical discourse: questions of theory and practice*: Proceedings of the 7th International Scientific, Practical and Educational Conference. Edited by E.V. Vinogradova (2019) (In Russ.)

9. E.A. Vishnyakova, O.D. Vishnyakova, A.V. Kiseleva, Medical professional discourse in the context of cognitive-linguistic analysis. *Discourse of professional communication*. №1-1, pp. 55-69 (2019) (In Russ.)

10. L.N. Ponomarenko, E.A. Mishutinskaya, I.S. Zlobina, Linguo-stylistic features of medical texts in the translation aspect. *Humanitarian paradigm*, №1 (4), (2018) (In Russ.)

11. A.O. Stebletsova, Y.N. Naumenko, Titles of English and Russian scientific articles: discursive-cognitive features. *Bulletin of the Moscow State Linguistic University. Humanitarian sciences*. № 4 (820). P. 161-179 (2019) (In Russ.)

12. A.O. Stebletsova, I.I. Torubarova, Some features of translation of medical texts: post-translation analysis of a scientific article. *Mir nauki, kultury, obrazovaniya* [Journal of science, culture, education]. № 1 (74). pp. 407-409 (2019) (In Russ.)

13. V. Johansson, Lexical diversity and lexical density in speech and writing: a developmental perspective // Lund University, Dept. of Linguistics and Phonetics. Working Papers. Vol. 53, pp. 61–79 (2008)

14. L. Batia, What percentage of text lexis is necessary for comprehension? // Special Language: from humans thinking to thinking machines / ed. by Ch. Lauren, M. Nordman. L. : Multilingual Matters. pp. 316–323 (1989)

15. I. S. P. Nation, How large a vocabulary is needed for reading and listening? in *Canadian Modern Language Review*, Vol. 63, iss. 1. pp. 59–82 (2006)

16. I.R. Galperin, Essays on the style of the English language (1958) (In Russ.)