Yaskevich Daniil Mikhailovich
student
Gladyshev Aleksandr Olegovich
student
Sokolov Daniil Antonovich
student
Misolin Timofey Iurevich
student
Scientific adviser
Ioffe Nina Evgenyevna
senior lecturer

MIREA – Russian Technological University
Moscow

FUNCTIONAL STYLES OF MODERN ENGLISH: A CORPUS-BASED COMPARISON OF THEIR LINGUISTIC FEATURES

Abstract: *functional styles of English are usually described with selected examples rather than measurements. In the study, scientific, official, newspaper, publicistic, belles-lettres and colloquial texts from the Brown Corpus and the NPS Chat Corpus (about 450 thousand words) were compared on eight quantitative features. Scientific and official texts had the longest sentences and the highest frequencies of nominalizations and passive constructions, while colloquial texts contained 13 times more first- and second-person pronouns than scientific prose.*

Keywords: *functional styles, stylistics, register, scientific style, colloquial style, corpus linguistics, nominalization, passive voice.*

Яскевич Даниил Михайлович
студент
Гладышев Александр Олегович
студент

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

Соколов Даниил Антонович
студент
Мисолин Тимофей Юрьевич
студент
Научный руководитель
Иоффе Нина Евгеньевна
старший преподаватель

ФГБОУ ВО «МИРЭА – Российский технологический университет»
г. Москва

ФУНКЦИОНАЛЬНЫЕ СТИЛИ СОВРЕМЕННОГО АНГЛИЙСКОГО ЯЗЫКА: КОРПУСНОЕ СРАВНЕНИЕ ИХ ЯЗЫКОВЫХ ПРИЗНАКОВ

Аннотация: *функциональные стили английского языка обычно описываются на отдельных примерах, а не с помощью измерений. В исследовании научные, официально-деловые, газетные, публицистические, художественные и разговорные тексты из корпусов Brown и NPS Chat (около 450 тыс. слов) сравнивались по восьми количественным признакам. Научные и официально-деловые тексты отличались наибольшей длиной предложений и частотой номинализаций и пассивных конструкций, а в разговорных текстах местоимения первого и второго лица встречались в 13 раз чаще, чем в научной прозе.*

Ключевые слова: *функциональные стили, стилистика, регистр, научный стиль, разговорный стиль, корпусная лингвистика, номинализация, пассивный залог.*

Introduction.

The theory of functional styles describes how the choice of language means depends on the sphere of communication. I. R. Galperin distinguished the belles-lettres, publicistic, newspaper, scientific prose and official document styles of English, while colloquial speech is usually considered separately [7]; I. V. Arnold described a style as a system of interrelated language means that serves a definite aim

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

in communication [1]. These descriptions are mostly qualitative and are based on selected examples. Corpus linguistics allows the same features to be measured in large collections of texts: D. Biber showed that the registers of English differ systematically in the frequencies of nouns, passives, pronouns and many other features [2; 3].

The aim of this study is to compare the main functional styles of English on a set of measurable linguistic features using publicly available corpora.

Materials and methods.

Five styles were represented by genres of the Brown Corpus, a corpus of written American English with part-of-speech tags [6]: scientific prose (category «learned», 80 texts), official documents («government», 30 texts), newspaper style («news», 44 texts), publicistic style («editorial», 27 texts) and belles-lettres («fiction», 29 texts). Colloquial style was represented by the NPS Chat Corpus, which contains posts from online chat rooms [5]; system messages and user names were excluded. Both corpora were used in the versions distributed with the NLTK library [4] and were processed with Python scripts. For every style, eight features were calculated: mean sentence length (for the chat, mean post length), mean word length, lexical density (the share of nouns, full verbs, adjectives and adverbs), and the frequencies of nominalizations (nouns in -tion, -sion, -ment, -ness, -ity, -ance, -ence), passive constructions (a form of be followed by a past participle), first- and second-person pronouns and contractions per 1,000 words.

Results.

The values of the features are given in Table 1. The features arrange the styles along a gradient from the bookish to the colloquial pole. Scientific prose and official documents had the longest sentences (about 21 words), the longest words and the highest frequencies of nominalizations and passives: official texts contained 4.7 times more nominalizations than fiction, and scientific texts 16 times more passives than chat. Newspaper and publicistic texts occupied the middle position, but the newspaper style had the highest lexical density (54.4%), which reflects its compressed, information-rich manner. The publicistic style differed from the newspaper style mainly

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)

by more personal pronouns and questions (5.0% of sentences against 1.0%), i.e. by its direct appeal to the reader. Fiction and especially colloquial texts showed the opposite profile: short sentences and words, few nominalizations and many personal pronouns and contractions. Long words of seven letters or more made up 30.9% of official and 28.7% of scientific texts but only 8.5% of chat posts. The standardized type-token ratio, calculated on segments of 1,000 words, was highest in newspaper (48.0%) and publicistic texts (47.8%) and lowest in official (39.3%) and scientific texts (40.9%). This shows that bookish styles repeat their terms for precision, while newspapers prefer a varied vocabulary.

Table 1

Quantitative features of functional styles of English (‰ – per 1,000 words)

Feature	Scientific	Official	Newspaper	Publicistic	Fiction	Colloquial
Words, thousand	159.9	61.1	87.0	54.1	58.3	29.7
Sentence length, words	21.0	21.1	18.9	18.3	13.7	3.9
Word length, letters	5.02	5.13	4.82	4.79	4.34	3.69
Lexical density, %	51.7	52.7	54.4	50.8	47.3	42.1
Nominalizations, ‰	49.7	61.7	30.8	33.6	13.1	3.6
Passives, ‰	19.3	18.2	12.2	11.5	6.5	1.2
Pronouns I/we/you, ‰	6.9	10.2	6.0	15.8	22.7	92.4
Contractions, ‰	0.4	0.1	2.9	2.9	8.9	31.7

Discussion.

The results confirm the qualitative descriptions of English styles [1; 7]: the impersonal, nominal and passive character of scientific and official texts and the personal, dialogic character of colloquial speech are clearly visible in the figures. They also agree with the multidimensional analysis of D. Biber, in which informational registers rich in nouns and long words are opposed to involved registers rich in pronouns and contractions [3]. At the same time, the boundaries between the styles are not sharp: scientific and official texts have almost identical profiles, and the newspaper style is closer to them than to fiction. For learners of English, this means that the style of a text depends not on single markers but on the frequencies of many features at once. For example, a scientific article written with contractions and personal pronouns at the level of fiction would be perceived as informal, while an excessive num-

Содержимое доступно по лицензии Creative Commons Attribution 4.0 license (CC-BY 4.0)

ber of nominalizations and passives makes any text heavy and difficult to read. The study is limited by the material: the Brown Corpus represents American English of the 1960s, and the chat corpus represents written online conversation rather than spoken colloquial speech.

Conclusion.

Simple quantitative features clearly distinguish the functional styles of English. Nominalizations, passive constructions and long sentences mark scientific and official texts, whereas personal pronouns and contractions mark colloquial speech. Such measurements can complement traditional stylistic analysis and help learners to control the style of their own texts.

Список литературы

1. Арнольд, И. В. Стилистика. Современный английский язык : учебник для вузов / И. В. Арнольд. – 13-е изд., стер. – Москва : Флинта, 2016. – 384 с.

2. Biber, D. Register, genre, and style / D. Biber, S. Conrad. – Cambridge : Cambridge University Press, 2009.

3. Biber, D. Variation across speech and writing / D. Biber. – Cambridge : Cambridge University Press, 1988.

4. Bird, S. Natural language processing with Python / S. Bird, E. Klein, E. Loper. – Sebastopol : O'Reilly Media, 2009.

5. Forsyth, E. N. Lexical and discourse analysis of online chat dialog / E. N. Forsyth, C. H. Martell // Proceedings of the First IEEE International Conference on Semantic Computing (ICSC 2007). – Irvine : IEEE, 2007. – P. 19–26.

6. Francis, W. N. Brown corpus manual / W. N. Francis, H. Kučera. – Revised ed. – Providence : Brown University, 1979.

7. Galperin, I. R. Stylistics / I. R. Galperin. – Moscow : Higher School, 1981. – 334 p.

Content is licensed under the Creative Commons Attribution 4.0 license (CC-BY 4.0)