

A Corpus-Based Investigation of Language and Writing Framework Style in the Works of Russian Researchers of the Late XVIII-Early XX Centuries in the Bronze Age¹

Elena Vorobeva², Svetlana Baksheeva³, Albina Imamutdinova⁴, & Ekaterina Plakhova⁴

Abstract

Corpus linguistics is one of the fastest-growing methodologies in contemporary linguistics. In a conversational format, this article answers a few questions that corpus linguists regularly face from linguists who have not used corpus-based methods so far. It discusses some of the central assumptions ('formal distributional differences reflect functional differences'), notions (corpora, representativity and balancedness, markup and annotation), and methods of corpus linguistics (frequency lists, concordances, collocations), and discusses a few ways in which the discipline still needs to mature. When studying the language and writing framework style in the works of Russian researchers of the late XVIII- early XX centuries are of considerable interest. The materials of the article can be useful to journalists, local lore specialists, those who deal with problems of house building and the oldest history of the Forest Volga region population.

Keywords: Corpus-Based Investigation; Writing Framework Style; Bronze Age.

1. Introduction

Corpus linguistics is the study of language as expressed in corpora (samples) of "real world" text. Corpus linguistics proposes that reliable language analysis is more feasible with corpora collected in the field in its natural context ("realia"), and with minimal experimental-interference. The field of corpus linguistics features divergent views about the value of corpus annotation. These views range from John McHardy Sinclair, who advocates minimal annotation so texts speak for themselves

¹ Please cite this paper as follows:

Vorobeva, E., Baksheeva, S., Imamutdinova, A., & Plakhova, E. (2019). A corpus-based investigation of language and writing framework style in the works of Russian researchers of the late XVIII-early XX centuries in the Bronze Age. *Journal of Research in Applied Linguistics*, 10 (SP), 1237-1243.

² Institute of international relations, Kazan (Volga Region) Federal University, Kazan, Russia; veer47@mail.ru

³ Russian Academy of Sciences, Vladivostok, Russia.

⁴ Institute of international relations, Kazan (Volga Region) Federal University, Kazan, Russia.

(Sinclair, 1992) to the Survey of English Usage team (University College, London), who advocate annotation as allowing greater linguistic understanding through rigorous recording (Wallis, 2007; Abedi, Keshmirshakan, & Namaziandost, 2019). The text-corpus method is a digestive approach that derives a set of abstract rules that govern a natural language from texts in that language, and explores how that language relates to other languages. Originally derived manually, corpora now are automatically derived from source texts. In addition to linguistics research, assembled corpora have been used to compile dictionaries (starting with The American Heritage Dictionary of the English Language in 1969) and grammar guides, such as A Comprehensive Grammar of the English Language, published in 1985.

Corpus research is an emerging and ever-growing field in second language (L2) pedagogy. The conspicuous development of computer technology has popularized corpus linguistics in language research. Corpus linguistic research enables extensive text analysis and provides empirical language description based on large quantities of authentic texts. An increasing number of studies have recently applied the insights and methodologies of corpus linguistics to L2 pedagogy and research (e.g., Biber & Conrad, 2001; Conrad, 2000; McCarthy & Carter, 2001; Abedi, Namaziandost, & Akbari, 2019).

Put simply, corpus research analyzes a “corpus” by using a “concordance program” in order to identify patterns of language use (Conrad, 2000; Nasri, & Namaziandost, 2019). A corpus is a large collection of natural texts, chosen to represent a language or a variety, which, nowadays, is almost always in a computer-readable form (Sinclair, 1991; Nasri, Namaziandost, & Akbari, 2019). Through concordance programs, corpora provide multiple examples of lexical or grammatical features in a variety of contexts and offer such textual information as word frequency counts and collocation patterns. The compilation of corpora has immensely expanded in number, size, and types (see Flowerdew, 2002, for an extensive overview of corpus development).

Corpus-based research shares several underlying principles. One fundamental assumption is that empirical data is needed for more accurate descriptions of language use. Language description, which is the basis of language teaching, has been grounded on human intuitions or “armchair reflections” (Johns, 1991) rather than on empirical observation. While such intuitions are often observed to be unreliable because of their subjective nature, the corpus-based approach has been regarded as a viable source to represent actual language use since it produces objective data about language use (Biber & Conrad, 2001; McCarthy, 2001; Shakibaei, Shahamat, & Namaziandost, 2019).

The style in writing can be defined as the way a writer writes. It is the technique that an individual author uses in his writing. It varies from author to author, and depends upon one’s syntax, word choice, and tone. It can also be described as a

“voice” that readers listen to when they read the work of a writer. There are four basic literary styles used in writing. These styles distinguish the works of different authors, one from another. Here are four styles of writing: Expository writing style is a subject-oriented style. The focus of the writer in this type of writing style is to tell the readers about a specific subject or topic, and in the end the author leaves out his own opinion about that topic (Namaziandost, Neisi, Mahdavi-rad, & Nasri, 2019).

Style in literature is the literary element that describes the ways that the author uses words — the author's word choice, sentence structure, figurative language, and sentence arrangement all work together to establish mood, images, and meaning in the text. Style describes how the author describes events, objects, and ideas. One easy way to understand literary style is to think about fashion styles. Clothes can be formal and dressy, informal and casual, preppy, athletic, and so forth. Literary style is like the clothes that a text puts on. By analogy, the information underneath is like the person's body, and the specific words, structures, and arrangements that are used are like the clothes. Just as we can dress one person in several different fashions, we can dress a single message in several different literary styles.

In general, it should be noted that the researchers of the late XVIII - early XX centuries repeatedly noted in their studies the preservation of a number of traditions of the Forest Volga region population. Those concern the basic principles of residential construction, mainly the coastal river arrangement of settlements, a certain arrangement of manor houses in the villages (“nesting” or in two parallel rows, saving multiple entrance constructions).

2. Methods

The leading approaches to the study of this problem are historical-comparative and structural-system methods.

3. Results and Discussion

Of special note is that corpus-based studies have concentrated on written production, although spoken corpora are increasing in number (Flowerdew, 2002). This is partly due to the ease of access to written texts as opposed to the difficulty in collecting speech samples, but also because text analysis is an indispensable methodology of corpus research. Accordingly, corpus research has revealed useful information about reading and writing, and, not surprisingly, corpus studies have often been associated with the teaching of L2 literacy skills. In this vein, Jabbour (2001) claims that “a corpus approach befits teaching second language reading and writing, since both activities are text oriented and make use of words and word combinations, or lexical patterns, within the confines of discourse” (p.294). On the other hand, it is also true that a very limited number of studies have addressed the use

of corpora with respect to reading. A greater number of corpus studies have focused on how to use corpus technology in L2 writing research and pedagogy (Namaziandost, Sabzevari, & Hashemifardnia, 2018).

The works of Russian researchers of the late XVIII and early XX centuries are of great interest considering the enrichment of knowledge about the historical and cultural heritage of the Forest Volga region Finno-Ugric population.

Soon it became clear though that the settlements and the remains of residential structures play an important role in discovering many unknown aspects of the life of ancient people. In 1877 the IV Archaeological Congress was held in Kazan. The results of it were the following: the success of Kazan scholars of antiquity was recognized and the creation of a special society to study the antiquities of Kazan province by the field research method was suggested. The results were decided to be reported on the pages of the Proceedings of the Society of Archeology, History and Ethnography (IOAIE). Since then archaeologists have not only created collections of ancient tools, but also studied the sites of ancient man's settlements. The archeologists collected expressive products from flint and bronze and fragments of dishes, animal bones as well drawing up plans for sites, sketching sections of geological strata.

It was not an easy task to pass to this new stage of archaeological work, since Russian scientists had no experience of such studies. Much could be taken from the practice of foreign archaeologists, but not all the monuments characteristic of Western Europe is typical of our territory. At first, these guidelines were sometimes more confusing than helpful. Thus, outlining the issues to be discussed at the IV Archaeological Congress, D.P. Sontzov wrote: "Pile structures found in all parts of Europe ..., must be also found in our regions." (Proceedings of the IV Archaeological Congress, 1884).

In summary, several gaps from the existing literature motivated this study. Most important of all, further research needs to be carried out to examine how corpus⁸ technology affects students' development of competence as L2 writers as well as their writing experiences. The development of learner-specific descriptions in corpus use and its evaluation would also be valuable. Also, there is little research on the teacher's integration of a corpus component into L2 writing instruction. In order to understand how this new technology affects and contributes to students' L2 writing, we need to look at the phenomenon in depth in teaching and learning practice. Thus, while the existing body of literature enhances our understanding of the usefulness of corpora in L2 writing as composite parts, a triangulation of multiple perspectives from teachers and students is needed, which in turn requires a triangulation of multiple methods (Namaziandost, Shatalebi, & Nasri, 2019).

While it is generally accepted that corpus linguistics has made contribution to descriptive linguistics, more controversies have been raised as to the usefulness of

corpora and the extent to which corpus findings should influence language teaching. This section discusses the controversies about using the corpus approach in language teaching and learning. First it reviews arguments in favor of the corpus approach in language learning: how the basic premises of the corpus approach can help students with their L2 learning. Four basic tenets of corpus studies are reviewed: the focus on lexis, the use of frequencies of words, the connection of form and function, and the emphasis on discovery learning. The arguments against using the corpus approach in language teaching then follow (Namaziandost, Neisi, Kheryadi, & Nasri, 2019). The cons mainly address three aspects: extrapolation from linguistics to pedagogy, authenticity of language data, and learner autonomy in the learning process. Several counter-arguments to the cons are also presented. While it is generally accepted that corpus linguistics has made contribution to descriptive linguistics, more controversies have been raised as to the usefulness of corpora and the extent to which corpus findings should influence language teaching. This section discusses the controversies about using the corpus approach in language teaching and learning. First it reviews arguments in favor of the corpus approach in language learning: how the basic premises of the corpus approach can help students with their L2 learning. Four basic tenets of corpus studies are reviewed: the focus on lexis, the use of frequencies of words, the connection of form and function, and the emphasis on discovery learning. The arguments against using the corpus approach in language teaching then follow. The cons mainly address three aspects: extrapolation from linguistics to pedagogy, authenticity of language data, and learner autonomy in the learning process. Several counter-arguments to the cons are also presented.

First of all, the corpus approach has brought to light the importance of lexis in language learning. In the past few decades, mainly due to Chomskyan grammar, vocabulary has not been considered as central to language teaching, while an overemphasis is placed on the structure of language. In contrast, as the corpus approach places lexis at the center of language description, its approach has emphasized the significance of lexis in language learning. Willis (1990) goes on to argue for a lexical syllabus where instruction is constructed around vocabulary in language teaching. He contends that identifying the most frequent words by use of the corpus approach also enables learners to identify the most frequent notions and functions.

4. Conclusion

Corpus linguistics is a method of carrying out linguistic analyses. As it can be used for the investigation of many kinds of linguistic questions and as it has been shown to have the potential to yield highly interesting, fundamental, and often surprising new insights about language, it has become one of the most wide-spread methods of linguistic investigation in recent years. Furthermore, it may be that corpus linguistics is resilient to clear definition because, as Gries (2006b: 4) suggests, it “seems to be a category with a proto-type structure: there are a few criteria that are –

though not individually necessary – shared by much, if not most, work within corpus linguistics, and there is a variety of criteria which are less central to the work of many corpus linguists” (see also Williams 2006). Studies of modern scholars, journalists, local historians are based on the essays, ethnographic notes, travelling notes of the writers of the second half of the XIX century. The material of the article is recommended for specialists whose object of study is the problems of settlement archeology when using archeological and ethnographic parallels, as well as for those who study and teach the history of the archaeological science development, the oldest history of the Volga Forest population.

Acknowledgements

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

References

- Abedi, P., Keshmirshakan, M. H., & Namaziandost, E. (2019). The comparative effect of flipped classroom instruction versus traditional instruction on Iranian intermediate EFL learners' English composition writing. *Journal of Applied Linguistics and Language Research*, 6(4), 43-56.
- Abedi, P., Namaziandost, E., & Akbari, S. (2019). The impact of flipped classroom instruction on Iranian upper-intermediate EFL learners' writing skill. *English Literature and Language Review*, 5(9), 164-172.
- Biber, D., & Conrad, S. (2001). Quantitative corpus-based research: Much more than bean counting. *TESOL Quarterly*, 35(2), 331-336.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly*, 34, 548-560.
- Flowerdew, J. (Ed.). (2002). *Academic discourse*. Harlow, England: Longman.
- Gries, S. T. (2006b). *Introduction to S. Gries and A. Stefanowitsch (eds.). Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis*, 1–18. Berlin, Heidelberg, New York: Mouton de Gruyter.
- Jabour, G. (2001). Lexis and grammar in second language reading and writing. In D. Belcher, & Hirvela, A. (Eds.), *Linking literacies: Perspectives on L2 reading and writing connections* (pp. 291-308). Ann Arbor: University of Michigan Press.
- McCarthy, M. (2001). *Issues in applied linguistics*. Cambridge: Cambridge University Press.
- McCarthy, M., & Carter, R. (2001). Size isn't everything: Spoken English, corpus, and the classroom. *TESOL Quarterly*, 35(2), 337-340.

- Namaziandost, E., Neisi, L., Kheryadi, & Nasri, M. (2019). Enhancing oral proficiency through cooperative learning among intermediate EFL learners: English learning motivation in focus. *Cogent Education*, 6(1), 1-15. <https://doi.org/10.1080/2331186X.2019.1683933>
- Namaziandost, E., Neisi, L., Mahdavidrad, F., & Nasri, M. (2019). The relationship between listening comprehension problems and strategy usage among advance EFL learners. *Cogent Psychology*, 6(1), 1-19. DOI: 10.1080/23311908.2019.1691338.
- Namaziandost, E., Sabzevari, A., & Hashemifardnia, A. (2018). The effect of cultural materials on listening comprehension among Iranian upper-intermediate EFL learners: In reference to gender. *Cogent Education*, 5(1), 1-27. <https://doi.org/10.1080/2331186X.2018.1560601>.
- Namaziandost, E., Shatalebi, V., & Nasri, M. (2019). The impact of cooperative learning on developing speaking ability and motivation toward learning English. *Journal of Language and Education*, 5(3), 83-101. <https://doi.org/10.17323/jle.2019.9809>.
- Nasri, M., & Namaziandost, E. (2019). The impact of topic choice on descriptive writing ability among Iranian advanced EFL learners. *International Journal of English Language Studies (IJELS)*, 1(1), 1-9.
- Nasri, M., Namaziandost, E., & Akbari, S. (2019). Impact of pictorial cues on speaking fluency and accuracy among Iranian pre-intermediate EF learners. *International Journal of English Language and Literature Studies*, 8(3), 99-109
- Shakibaei, G., Shahamat, F., & Namaziandost, E. (2019). The effect of using authentic texts on Iranian EFL learners' incidental vocabulary learning: The case of English newspaper. *International Journal of Linguistics, Literature and Translation (IJLLT)*, 2(5), 422-432
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1992). The automatic analysis of corpora, In Svartvik, J. (ed.), *Directions in corpus linguistics (Proceedings of Nobel Symposium 82)*. Berlin: Mouton de Gruyter.
- Wallis, S. (2007). Annotation, Retrieval and Experimentation. In Meurman-Solin, A. & Nurmi, A.A. (eds.), *Annotating variation and change*. Helsinki: Varieng, University of Helsinki.
- Williams, G. (2006). La linguistique de corpus: Une affaire préposition-nelle. In F. Rastier & M. Ballabriga (eds.). *Corpus en lettres et sciences sociales: Des documents numériques à l'interprétation. Actes du colloque international d'Albi, juillet 2006* (151–158). Paris: Texto.
- Willis, D. (1990). *The lexical syllabus*. London: Harper Collins.