

**The National Russian Corpus as a teaching tool - general observations and case studies  
from the perspective of Russian language teaching in Poland**

Опубликовано в: Национальный корпус русского языка и проблемы гуманитарного образования. Под редакцией Н.Р. Добрушиной. Москва 2007. С. 52-59.

Цель этой статьи – представить некоторые теоретические соображения, а также примеры исследований отдельных проблем, иллюстрирующие применение Национального корпуса русского языка в преподавании общего курса русского языка, курса русского языка для делового общения и курса польско-русского перевода польским студентам в институте Восточно-славянских исследований в университете Ополе (Польша). Одной из причин, подвигших нас на написание этой статьи, стал постоянно усиливающийся дефицит современных учебных материалов по русскому языку в сегодняшней Польше; все острее ощущается на польском книжном рынке нехватка толковых словарей русского языка. Кроме того, растущее применение коммуникативных и лексических методов в преподавании русского языка и связанный с этим акцент на преподавании живого языка, используемого в повседневной коммуникации, усиливают роль НКРЯ как важного вспомогательного инструмента при обучении русскому языку в современной Польше.

There have been numerous journal articles or edited volumes published in recent years on the topic of the relationship between second or foreign language teaching and language corpora, where corpus methodology is introduced into the state-of-the-art teaching methods and strategies (Wichmann et al. 1997; McEnergy at all 2006, Gajek 2002). According to Leech (1997), the conspicuous convergence between teaching and language corpora may be observable at three interrelated levels, namely 1) the direct use of corpora in teaching (teaching about language corpora, teaching to browse the corpus, and browsing the corpus to teach); 2) indirect use of corpora (development of language teaching materials, development of testing and assessment tools; development of supplementary language teaching materials and mini-lexicons); 3) teaching-oriented corpus development (compiling learners' corpora, both general and specialist; compiling developmental L1 and L2 corpora; compiling parallel corpora etc). These three areas of convergence, as proposed by Leech (1997), provide ready-made and general typology of the role of corpora in foreign language teaching, which may serve as a methodological framework for the application of language corpora in foreign or second language teaching. Nevertheless, the specificity of teaching and learning environment as well as different languages taught to students require that the above typology be readapted and modified in order to meet the requirements of particular teaching situation. Due to the fact that majority of reference materials on the relationship between language corpora and language teaching have been written in English and that corpus linguistics as an academic subject is not commonly offered in departments or

institutes of Slavonic studies in Poland, the typology of using the NRC for teaching purposes shall be more elaborated than the one proposed by Leech (1997) and shall include the following:

Table 1. Framework of using the NRC in Russian language teaching (based on Leech (1997))

<b>Direct use of corpora in Russian language teaching</b>	<ul style="list-style-type: none"> <li>- <b>general introduction to corpus linguistics as branch of linguistics and methodology of research*</b></li> <li>- <b>specific introduction to the NRC (its size, structure and search options)</b></li> </ul>
<b>Indirect use of corpora in Russian language teaching</b>	<ul style="list-style-type: none"> <li>- development of Russian language teaching materials</li> <li>- development of testing and assessment tools</li> <li>- <b>syllabus design with reference to the application of the NRC across the curriculum</b> (general and domain-specific academic subjects) in order to facilitate effective use of the NRC during practical classes in <b>General Russian, Business Russian, translation</b>, lexicography, lexicology, methodology of teaching, language pedagogy etc)</li> <li>- conducting research on the effects of using the NRC in Russian language teaching</li> <li>- conducting research on the learners' feedback on using the NRC as a teaching tool</li> <li>- enhancement of awareness among Russian language teachers in Poland of the existence of the NRC and other Russian corpora</li> </ul>
<b>Teaching-oriented corpus development</b>	<ul style="list-style-type: none"> <li>- compilation of learners' corpora (both general and domain-specific) at various competence levels</li> <li>- compilation of parallel Russian-Polish and Polish-Russian language corpora (both general and domain-specific)</li> <li>- development of professional language teaching materials based on the corpus resources (course books as well as practical guides on how to use the NRC in foreign language teaching, targeted at teachers of Russian in Poland)</li> </ul>

\* The cases studies presented in this article, which focus on the application of the NRC in teaching General and Business Russian as well as Russian-to-Polish or Polish-to-Russian translation have been printed in bold.

The table above presents not only the typology of how the NRC may be used in Russian language teaching but also signals the needs and requirements of corpus-based and corpus-driven Russian language teaching in Poland. As a matter of fact, academic curriculum which does not include even the introductory course to corpus linguistics unavoidably results in low awareness of the NRC as a teaching tool, which further leads to the lack of corpus-based teaching materials and practical guides for teachers, lack of learners' corpora as well as comparable and parallel corpora. As a result, if the direct and indirect use of the NRC is neglected, then teaching-oriented corpus development is almost non-existent, which is unfortunately the challenge to be faced in the near future in Poland in order to make Russian language teaching more innovative and up-to-

date with state-of-the art developments in foreign language teaching, corpus linguistics and, in particular, with Russian corpus linguistics. This becomes even more important when one notices an increasingly difficult access to contemporary Russian language teaching materials in Poland as well as the scarcity of monolingual dictionaries of contemporary Russian on the Polish market.

Having presented the above theoretical observations, I will elaborate the points printed in bold in the table 1, which are the focal points of this article.

It is postulated that the prerequisite for application of the NRC in Russian language teaching is the introductory course targeted at both instructors as well as students. Such a course should cover the rudiments of corpus linguistics and familiarize the students with the most popular corpora of Russian and Polish (Grabowski 2006: 29-33). Finally, it should focus on the characteristics and options available in the NRC when used as a concordancing agent and language teaching tool. It should focus, among others, on the following aspects: range of tagging, search options (exact search as well as lexico-grammatical search) and logical operators, structure and contents of the NRC. Since the existence of the NRC is not widely-known among Polish teachers of Russian, the idea of such introductory course appears to be *sine qua non* condition before putting the NRC directly into teaching practice. In this context, it is obvious that such classes should take place in a computer laboratory with the access to the Internet.

It is paramount that the NRC is made an integral element of academic curricula. In this article I present selected case studies featuring application of the NRC during practical classes in General Russian, Business Russian as well as Russian-to-Polish Translation.

As for General and Business Russian teaching, I postulate that the NRC may appear to be an invaluable tool in development of professional (or specialist) vocabulary in full context (i.e. meeting the requirements of lexical approach), which is illustrated by the following exercise. The students' task is the following one:

**'Using lexico-grammatical search option, prepare a list of types of companies which operate on the Russian market. Try to provide functional Polish equivalents'.**

Students browse the NRC by means of lexico-grammatical search where they key in the following search command: adjective (A) + *предприятие* (N); adjective (A) + *компания* (N). The number of concordances displayed as results is exorbitant, namely over 12,000 and 15,000 contexts for *предприятие* and *компания*, respectively. Selected collocations are presented below:

- a) *предприятие – унитарное, совместное, полузакрытое, кооперативное; венчурное, казённое; малое, мелкое, крупное, сверхкрупное*
- b) *компания - лизинговая, аутсорсинговая, инвестиционная, экспедиторская, экспорт ориентированная, частная, дочерняя, кэптивная, материнская, холдинговая, агентская, консультационная, зарубежная, оффшорная*

Such types of exercises may be extended depending on the domain-specific vocabulary taught to students, e.g. a focus of the exercise may be types of banks operating on the Russian market or types of loans these banks offer. The latter exercise may look the following way:

**‘Browse the NRC using lexico-grammatical search and find the types of loans which are offered on the Russian market. Group the results according to the following criteria: 1) maturity of the loan; 2) purpose of the loan; 3) method of financing the loan.’**

Selected search results may be the following: 1) *однодневный* (one-day), *трехдневный* (three-day), *двухнедельный* (two-week), *тридцатидневный* (30-day), *сверхкороткий* (short-term), *долгосрочный* (long-term)...; 2) *ипотечный* (mortgage loan), *синдицированный* (syndicate loan), *экспортный* (export loan), *потребительский* (consumer loan) ... 3) *овердрафтный* (overdraft), *межбанковский* (interbank loan), *правительственный* (subsidized loan), *синдицированный* (syndicated loan), *рублевый* (in roubles), *валютный* (in foreign currency)...

Having completed the search, the students are asked to provide Polish equivalents of Russian expressions on the basis of available resources, such as Polish-Russian dictionaries, thematic dictionaries, internet websites and vertical portals devoted to relevant economic topics. Obviously enough, some of the expressions will overlap in both languages (especially in case of loanwords).

Summing up, such exercises enable students to access up-to-date professional vocabulary as well as acquire it in full contextual environment. Moreover, the use of the NRC enriches lexicographic material available in dictionaries which is by definition limited to a handful of selected examples and concordances.

In the case of General Russian and Business Russian teaching, the NRC may facilitate acquisition of grammatical relations between Russian verbs and the prepositions they govern. Moreover, the NRC may also provide the data on frequency of verbs, which is neglected in the majority of Russian dictionaries available on the Polish market. The advanced students’ task is the following one:

**'Using lexico-grammatical search option, determine the case of prepositional phrase which is governed by the following verbs belonging to the semantic field 'complaining about sth/dissatisfaction with sth': *жаловаться, сетовать, хныкать, роптать, брюзжать*. Find out which verb is the most frequently used and which one is the least frequently used'.**

Using relevant search option, students key in the following combination: a verb (one of the above, e.g. *сетовать*) (V) + preposition (PR), and analyze displayed concordances in terms of sought-after government relations. Eventually, the students compare frequency data on the verbs in question.

The NRC may be successfully applied during practical classes in translation. In one of illustrative exercises, the students are preoccupied with translation of specialist text about information technologies (IT) into Russian. Scarcity of relevant thematic resources implies the use of the NRC as a reference tool whereby the search is narrowed down only to texts related to the topic of IT. The students task is thus to create the NRC subcorpus related to IT by using available meta-situational information (thematic tagging) the NRC has been equipped with. Only then is the search limited to 50 texts about IT which form the NRC subcorpus.

The next exercise provides more specific example on how the NRC may facilitate solving the problem of equivalence on stylistic and collocational levels. The rationale behind it is the ambiguity and limitation of lexicographic description in some monolingual dictionaries of Russian available to Polish students. This problem was encountered when students were supposed to translate from Polish into Russian a text about modern tendencies in interior design and they were not sure which Russian equivalent would be the best one in translation of the Polish phrase: *dekorator wewnętrz* (interior designer) *дизайнер* or *декоратор*. The example illustrating their doubts are definitions of two semantically-related lexemes extracted from *Explanatory Dictionary of Russian Language* by S. Ozhegov and N. Shvedova [Ожегов С., Шведова Н. (2000) *Толковый Словарь русского языка. Москва: Русский язык. 4-е издание, дополненное*]:

- a) *дизайнер* - художник-конструктор, специалист по дизайну
- b) *декоратор* -
  - (1) художник, пишущий декорации, оформляющий сцену, съёмочную площадку
  - (2) специалист по декорированию помещений.

With the above ambiguity in mind, the students' task was to find the best Russian translation in the NRC. Moreover, they were asked if they could determine the difference in meaning on the basis of corpus data and what other adjectives collocated with the words *дизайнер* and *декоратор*?'. Having used lexico-grammatical search with relevant search

command, the students discovered that *интерьерный дизайнер* (not: *интерьерный декоратор*) was the best translation of the aforementioned Polish phrase. Moreover, on the basis of collocational data, students are able to determine semantic differences in the case of two Russian nouns as well as their frequency of occurrence in contemporary Russian:

- a) *дизайнер* – (400 occurrences in 320 texts) *ландшафтный, парковый, садовый, интерьерный, графический, фэшн, кабинетный, промышленный, автомобильный*
- b) *декоратор* – (41 occurrences in 32 texts) *оперный, театральный*

The practical exercise presented above shows that the NRC may be a useful supplement to traditional dictionaries used during translation classes and may provide help in selecting correct and appropriate stylistic and lexical equivalents which the students want to use when translating domain-specific texts. Moreover, the NRC may partly substitute native speakers of Russian (i.e. teachers or instructors), especially when the students have no contact and thus no opportunity to brainstorm their ideas with instructors who are not native speakers. On top of that, in such exercises the NRC offers more detailed analysis of collocations and phraseologisms characteristic for particular language registers (Uzar 2006: 160). Therefore language corpora are bound to become a useful and important teaching and reference tool during practical classes in translation (Gajek 2006: 153).

Summing up the contents of this article, it appears that the NRC may be used as a valuable teaching tool in the process of Russian language teaching to Polish students. Nevertheless, there has been no research conducted in Poland with the aim to study the reaction of the Polish students of Russian and their feedback related to the use of the NRC. Moreover, the research on effectiveness of using the NRC in Russian language teaching practice in Poland would also provide valuable insights and outline potential areas for improvement as far as a didactic process is concerned. With the above in mind, the case studies presented in this article provide individual and selected teaching situations which can not be regarded as the basis for proper generalizations (as there is no hard quantitative data supporting the assumptions presented herein).

Nonetheless, it is paramount to remember that the NRC is still not widely used in the Russian language teaching at Polish institutions of higher education and its application and popularity as a supplementary teaching tool will grow in the future only if the following obstacles are surmounted:

- a) technical and financial (i.e. learning) infrastructure – computer laboratory equipped with broadband access to the Internet, which enables instructors and students fast browsing the resources of the NRC

- b) organizational – corpus linguistics shall be offered as either separate academic subject or included into the existing syllabi of the courses offered at university level (e.g. General Russian, Business Russian, Translation etc.)
- c) pedagogical – instructors shall be trained on the rudiments of corpus linguistics and on how to use the NRC as a reference material and language teaching tool; they shall receive relevant know-how on how the use of the NRC may be integrated with contemporary methods of foreign language teaching
- d) personal (psychological) – which requires that the students are encouraged to change their learning habits (which would involve more analytical, bottom-up approach to the analysis of linguistic data) in order to use the NRC more effectively.

If the proper attention is paid to the above requirements, then the National Russian Corpus may serve as a valuable and popular innovation into Russian language teaching in Poland. To this end, the case studies presented in this article account for the step in the right direction.

## References

- Gajek, E. (2002). *Komputery w nauczaniu języka*. Warszawa: PWN.
- Gajek, E. (2006). „Korpusy w dydaktyce – opis innowacji dydaktycznej, studium przypadku”. In: *Korpusy w angielsko-polskim językoznawstwie kontrastywnym*. Teoria i praktyka. Kraków, pp. 143-155.
- Grabowski, Ł. (2006). „Comparative Characteristics of Polish and Russian Language Corpora”. In: B. Німчук (ed.) *Лексикографічний бюллетень: Збірник наукових праць*. Kiev. pp. 29-33.
- Leech, G. (1997). “Teaching and language corpora: a convergence”. In: A. Wichmann, S. Fligelstone, T. McEnery, G. Knowles, (eds). *Teaching and Language Corpora*: London: Longman. pp. 1-23.
- McEnery T., Xiao R., Tono Y. (2006). *Corpus-based Language Studies. An advanced resource book*. London.
- Wichmann, A, Fligelstone, S., McEnery, T., Knowles, G. (eds) (1997). *Teaching and Language Corpora*: London: Longman.
- Uzar R.(2006). „Korpusy w nauczaniu tłumaczenia i w pracy tłumacza”. In: A. Duszak, E. Gajek, U. Okulska, *Korpusy w angielsko-polskim językoznawstwie kontrastywnym*. Teoria i praktyka. Kraków, pp. 156-177.