JOANNA SATOŁA-STĄSKOWIAK

Institute of Slavic Studies, Polish Academy of Science, Warsaw
joannasat@op.pl

CONTEMPORARY CONTRASTIVE STUDIES OF POLISH, BULGARIAN AND RUSSIAN NEOLOGISMS VERSUS LANGUAGE CORPORA

Abstract

In the field of Slavonic linguistics **contrastive studies of neologisms** occupy little place, the newest words are insufficiently described and classified. The aim of this article is to draw attention to the need for contrastive description of the newest lexis and checking exclusively one of many possibilities of obtaining Polish, Bulgarian and Russian neologisms. Language corpora, as this possibility is in question, are not the only source from which the author obtains her research material, yet a growing interest in corpora has inspired her to also use this method. The author wants to show the reader to what degree language corpora can help in building the thesaurus of Polish, Bulgarian and Russian neologisms. Making an attempt to confront a collection of neologisms of contemporary Polish, Bulgarian and Russian language, the author points out the need to standardize the description (identical for each of the analysed languages^a), which she intends to propose in another publications on neologisms in Polish, Bulgarian and Russian language. The application of contrastive method to three different but related languages from the Slavonic group will help, in her opinion, to discover more mechanisms of new words coming into existence and examine the newest derivative processes and their productivity.

Keywords: neologisms, neosemantisms, internationalisms, language corpora, parallel corpora, contrastive studies, Polish language, Bulgarian language, Russian language.

^aMore on the same subject in other articles by the author — cf. e.g. Satoła-Stąskowiak (2012, 2013).

1. Introduction

We encounter the newest words every day, irrespective of the level of education or participation in culture. They appear in all the styles and concern every language

carrier. Recently part of my research has been connected with the newest Slavonic lexis, with special consideration given to neologisms in three languages listed in the title of this article which represent well the western, eastern and southern Slavism. Each of them has a rich literary tradition and clearly formulated norms. Despite the social consciousness of the existence of newer and newer lexical units called neologisms, I have not come across any publications which would systematize the issues of contrastive description of the newest lexemes in the languages: Polish, Bulgarian and Russian. I can see here for me and other linguists scope for search and studies which, I hope, will fill the current gap in working out this issue.

In this article I describe only a chosen fragment of the issues concerning neologisms. I deliberately do not focus on the origin of individual words in the analysed languages, on their history.

I do not indicate the directions of development of contemporary Slavonic languages (Polish, Bulgarian and Russian languages). I also realise the fact that language corpora are only one of many other sources of excerpts of the newest lexis but a very interesting one. That is why it is worth presenting as an essential but not the only contemporary research tool.

2. Neologisms

2.1. In the article apart from the terms *neologism*, *neologisms* I use expressions such as: *the newest lexis*, *the newest words*. One should accept that for the author these terms are exchangeable, describing the same language phenomenon.

2.2. Neologism¹ is: a new word created in a language in order to name the previously unknown object, phenomenon, situation or feature (*Pol. przeglądarka (internetowa)*, *Bulg. браузер*, *Russ. браузер*; *Pol. zmywarka*; *Bulg. съдомиялна машина*; *Russ. посудомоечная машина*, *Pol. dresiarz*, *Bulg. мутра*, *Russ. зонник*), a new word created in order to achieve a poetic effect (in this article I intentionally omit such examples as they are not within the range of my interest in neologisms); a lexical unit borrowed from another language because of a lack of a suitable word in a natural language (*Pol. bankomat*, *Bulg. банкомат*, *Russ. банкомат*; *Pol. email*, *Bulg. имейл*, *ros. электронная почта*; *Pol. monitoring*, *Bulg. мониторинг*, *Russ. мониторинг*; *Pol. laptop*, *Bulg. лаптоп*, *Russ. лаптоп*; *Pol. spa*, *Bulg. спа*, *Russ. спа*; *Pol. blog*, *Bulg. блог*, *Russ. блог*; *Pol. czat*, *Bulg. чат*, *Russ. чат*; *Pol. weekend*, *Bulg. уикенд*, *Russ. уик-энд*; *Pol. klonowanie*, *Bulg. клониране*, *Russ. клонирование*); internationalism² (*Pol. audyt*, *Bulg. одит*, *Russ. аудит*; *Pol. kiler*, *Bulg. килър*, *Russ. киллер*; *Pol. fashion*,

¹Cf. (Jadacka, 2010, p. 39–52).

²As a result of the strongly noticeable elevation of contemporary colloquial language, manifesting itself in the use of the growing number of common terms, mainly, as it is the case in the Polish, Bulgarian and Russian language, adapted English borrowings, we can observe a large growth in the new lexis. We should therefore distinguish the third group which is under a strong influence of the phenomenon of using internationalisms, which is more connected with fashion rather than a factual lack of a word in a given language. On the same subject also in (Satoła-Staškowiak, 2013).

Bulg. *фелушн*, Russ. —, Pol. *hit*, Bulg. *xum*, Russ. *xum*; Pol. *peeling*, Bulg. *ну-линг*, Russ. *пилинг*, although in the system of individual languages there are (not always) words such as (one by one according to examples): Pol. *kontrola*, Bulg. *проверка*, Russ. *проверка*; Pol. *zabójca*, Bulg. *убиец*, Russ. *убийца*, also Pol. *moda*, Bulg. *мода*, Russ. *мода*; Pol. *przebieg/szlagier*, Bulg. —, Russ. *шлягер* etc.); there is still another and a very interesting group of neologisms: neosemantisms, lexemes, which have been present in the system of the language for a long time but have only recently been used to describe a new meaning (Pol. *myszka*, Bulg. *мишка*, Russ. *мышь*; Pol. *wirus*, Bulg. *вирус*, Russ. *вирус*; Pol. *sieć*, Bulg. *мрежа*, Russ. *сеть*). The main source of neosemantisms in the Polish language is a general language (about 2500 words)³, next colloquial (800 lexical units).

2.3. The words analysed here come from the end of 20th century at the earliest, although a particular emphasis is put on the year 2000 and the following years.

In the year 2001 S. Grabias (2001, p. 247) indicated that neosemantisms constitute as much as 50% of all the words in *the Dictionary of Students' Slang*, which is the evidence of young people's exceptional formative abilities and their need to depict the world in a vivid way. It is also proved by the popularity of dictionaries such as *the Cool Dictionary of the Youngest Polish Language* (*Wyczesany słownik najmłodszej polszczyzny*) by B. Chaciński (2001) or *the Smashing Dictionary of the Youngest Polish Language* (*Wypasiony słownik najmłodszej polszczyzny*) (2003) by the same author.

In the Bulgarian language⁴ according to the authors of a dictionary entitled *Речник на новите думи в българския език* published in 2010 the newest lexis constitutes about 5000 words (that is the number of entries that the above mentioned dictionary contains) out of which 4300 are completely new lexical units, 700 are neosemantisms whereas 600 neologisms are specialist terms connected with specific fields of life and 150 are idioms. A characteristic feature of the newest Bulgarian language is (as in the case of the Polish and Russian language) a large number of words coming from youth language and an accumulation of technical terminology.

In the Russian language⁵, similarly as in Polish and Bulgarian, after the year 2000 we observe a fast growth of lexis, mainly, as mentioned above, specialist terminology and the language of subcultures. We can also observe a process of assimilating new lexemes to the system and replacing with them the ones which already exist in the language. This is the topic on which the linguists describing an influence e.g. of anglicisms on the contemporary Russian language focus. Different factors affect the formation of neologisms: expression also has great significance and is the reason for creating new words as well as the development of technology and the need for naming situations, events and objects which have not been named before. Small dictionaries of Russian neologisms are very popular on the Internet (cf. <http://www.tutoronline.ru/blog/kratkij-slovar-neologizmov.aspx>).

³(Grabias, 2001, p. 235–253).

⁴*Речник на новите думи в българския език* 2010.

⁵Cf. (Nadel-Czerwińska, 2006; Козырев, Черняк, 2000; Наумова, 2004; *Новые слова и значения*, 1997; *Толковый словарь русского языка конца XX века*, 1998).

2.4. The origin of the largest number of neologism is known. Thanks to the linguists' research⁶ we know that many of them are formed in technical and medical fields (that is the names of equipment and tools), in economics, but the most in youth language and the language of subcultures.

One of the most important factors in the formation of the newest lexis is **expression** whose examples can be found not only among the above mentioned environmental groups but also in groups connected with artistic activity.

3. Language corpora

What is a language corpus? At present many promoters of the newest linguistic technologies express their opinions about it, among others Piotrowski (2003, p. 146) who defines a corpus as: " a collection of texts recorded in a digital form which constitute a representative sample of a given language. What makes a corpus different from any given collection of texts is the fact that it constitutes a certain well-thought-out whole. A distinctive feature of computer corpora is also their size".

Among the creators and users of corpora two quite different views are evident. Some of them claim that corpora have to contain a representative sample of a language (that is a large amount of text in its different language variants) in electronic form, others believe that even the smallest collection of texts, properly prepared, has the right to become a corpus (cf. McEnery & Wilson, 1996, p. 2; Podhajecka, 2006, p. 338–339).

A number of the newest specialist texts define the notion of a corpus much more precisely, giving details necessary to build such a collection of texts called "language corpus". These are, among other things (according to John Sinclair (1996)): quantity, quality, simplicity, documentation and (according to McEnery & Wilson (1996)): representative character, finite quantity, electronic format and standard reference. **The author considers the representativeness of the texts collected in a corpus as its most important feature, which is of great significance in seeking neologisms.**

4.0. Neologisms and language corpora

4.1. Language corpora

Language corpora, due to the rapid development of corpus linguistics and computer technologies, have become important research material, used also in applied linguistics (cf. Hunston, 2002) as well as in lexicography (cf. Ooi, 1998).

Thanks to language corpora a given lexical unit is represented in a wider context, which enables the researchers to perform a linguistic analysis, assess the kind of text, its quality and, what is extremely important on account of research on neologisms, the number a given lexem is used to be able to state if it really exists in the

⁶Cf. also (Благоева, 2001, p. 187–189, 2003, p. 133–140, 2004, p. 45–52, 2005, p. 117–118, 2005, p. 102–108, 2006, p. 16–25; Колковска, 2010, p. 21–30, 2007, p. 303–310, 2011, p. 519–532; Buttler, 1962, 1977, 1978, 1979; Skubalanka, 1990; Smółkova, 1976, 1978, 1997, 2001; Dąbrowska, 1957; Chruścińska, 1975, 1976, 1977; Satkiewicz, 1976, 1981; Orylska, 1983; Markowski, 1994, 1988; Jadacka, 2001, 2010; Брагина, 1973; Земская, 1992; Козырев, Черняк, 2000).

system of a given language and is a fully accepted lexical unit or if it is a language phenomenon with a very restricted time of existence, as e.g. occasionalisms⁷.

Occasionalisms constitute an exceptional group of new words formed in special contextual conditions. They are not as significant as neologisms because the time of their existence is most often strongly restricted by a specific communicative situation and the time of the ending of the emergent phenomenon which has been a source of inspiration for the emergence of a neologism. Among the latest interesting Polish examples we can count new lexems which came into existence in recent months. They are all connected with the European Football championships Euro 2012 which took place in Poland and the Ukraine, e.g. *piłkoszał* (the football craze) 'delight at everything connected with football and the championships, collective football euphoria' (From television and radio advertisement broadcast in Poland during the championships Euro 2012); *eurogedon* 'the prediction of a difficult time for the Poles during the championships Euro 2012, among other things, transport problems' (From a broadcast of Polish Radio Programme Three, just before the championships Euro 2012); *eurooszołomstwo* (eurofanaticism) 'going crazy about only one topic and a lack of contact with the people who are not interested in it' (From a statement by a journalist Jarosław Kuźniar published on his blog) (19th June 2012) .

4.1.1. The most popular corpora can be helpful in search of the newest Polish lexis:

The Corpus of Polish IPI PAN (300 m. segments), The Corpus of Polish PWN (full net version 40 m. segments, demonstration version 7 250 000 segments), which is a part of The National Corpus of Polish (over one and a half billion words), The Reference Corpus of Polish 'PELCRA' (100 m. words)

Let us then use the PWN corpus to look for a word *bankomat* (a cash machine) and the surrounding context:

- Bank nie odpowiada za zatrzymanie karty w **bankomacie** z przyczyn technicznych lub w wyniku wadliwej obsługi bankomatu przez Posiadacza karty. (The bank is not responsible for keeping the card in the **cash machine** for technical reasons or as a result of the incorrect operation of the cash machine by the card holder.)
- Jednocześnie dla kart PBK START został wprowadzony dzienny Limit Wypłat Gotówki w **bankomacie**, co oznacza, iż kwota dokonanych przez Państwa wypłat gotówki w **bankomacie/bankomatach** w ciągu jednego dnia może wynosić maksymalnie: (900 zł) (At the same time for PKB START cards a daily Limit of Cash Withdrawal from a **cash machine** has been introduced, which means that a total amount of cash withdrawals made from a **cash machine / cash machines** within one day can come to maximally: (900 złoty))
- Dzięki Karcie Emerytalnej Commercial Union, jako członek naszego funduszu, będziesz mógł w **bankomacie** sprawdzić wysokość wpłacanej przez

⁷Cf. (Satoła-Staśkowiak, 2013).

pracodawcę składki. (Thanks to the Commercial Union Pension Card, as a member of our fund, you will be able to check in the **cash machine** the amount of an insurance fee paid by the employer)

- Każda wypłata w **bankomacie** lub zakupy mogą dojść do skutku tylko wówczas, gdy na koncie posiadacza karty są pieniądze. (Every withdrawal from a **cash machine** or purchases can take place only when there is money in the holder's account.)

4.1.2. When studying Bulgarian neologisms it is worth using The National Corpus of Bulgarian⁸.

Below some examples with a Bulgarian word *одит*.

- Изваил старио троа пари, та му дал на сиромашо и го испратил назат да си **одит**: „ай со здравје ой ми си пријателе, му ракол старио, и коа ќе му кажиш на старио за рајо што ќе го вардит, ако нешто неќит и се налутит, тој ќе бидит за газерите [2] от казаните, што се во вечна;
- Тутракан 2, ет. 4, ап. 13, с предмет на дейност: независим финансов **одит**, финансово-счетоводна и консултантска дейност, покупка на стоки или други вещи с цел препродажба в първоначален, преработен или обработен вид, продажба на стоки от собствено производство, хотелиерски, туристически, рекламни, ресторантърски и информационни услуги, комисионни, спедиционни, превозни, складови и лицензионни сделки, лизинг, сделки с интелектуална собственост, транспортни, таксиметрови, външнотърговски и бартерни дейности, покупка, строеж или обзавеждане на недвижими имоти с цел продажба.
- Людмил Спасов, експерт от дирекция "Одит на средствата от Европейския съюз", Людмила Рангелова, директор на дирекция "Вътрешен **одит**" към Министерството на финансите, Тодор Янкулов, началник на отдел "Предприсъединителни фондове и програми на ЕС" към дирекция "Одит на средствата от Европейския съюз" ", и Дилянка Желязкова, директор на дирекция "Одит на средствата от Европейския съюз".
- Въпросът е за кои, при положение че за единия от проектите (Разлог Банско Гоце Делчев Садово) в **одит** на КРМГ се твърди, че липсва важен анекс към договора, за друг (тунела при околоръстния път на Дупница), че няма писмени доказателства за основанията да бъде избрана фирмата изпълнител, за трети (участъкът от Ново село до магистрала „Марица“), че са изчезнали 3 тома от договор, и т. н.

4.1.3. In case of research on the newest Russian lexis The National Corpus of Russian, open for general use, is one of the corpora that can facilitate work. (<http://www.ruscorpora.ru/search-main.html>).

Let us look at a Russian term *клонирование*:

⁸On website: <http://search.dcl.bas.bg/>.

- Самой затратной оказалась Москва — здесь на **клонирование** образа Беспалова ушло около \$ 400 тысяч.
- Отвечая на вопрос журналиста, думал ли он, что **клонирование** как научное открытие может быть использовано во вред, С. Антинори ответил отрицательно.
- В то же время болезнетворные бактерии делятся вдесятеро быстрее, и **клонирование** никак не сможет поспеть за их размножением.
- Потом просто пошло **клонирование**, примерно как в сериале «Солдаты», одни ушли, другие пришли и т.д.

4.1.4. The use of language corpora with resources which include exclusively one language to obtain neologisms and the examples of their usage is very interesting. Nonetheless, in contrastive linguistics such a method can cause many problems and become time-consuming despite the modern way of obtaining material. After all, the knowledge of a language presented in a corpus and the knowledge of a word one is looking for are required. Therefore those who just want to expand their linguistic skills and find out about the existence of a specific lexical unit in language B have a more difficult task. Language corpora can only confirm the use of a given lexem in a wider context. They are therefore intended for those who know what they are looking for. The collections of contexts increase the possibilities of research into neologisms. They indicate clearly the scope of the occurrence of neologisms, frequency and their stylistic plane.

4.2. Parallel corpora

Parallel corpora are an important tool of contemporary comparative linguistics. They allow a quick and simple way of obtaining rich and reliable material for research in the field of the theory of translation and thanks to the diversity of the collected material also for research in other fields: cultural, sociological, sociolinguistic and concerning political science.

Because of the diversity of the material they are an excellent source of obtaining neologisms. People learning a language can find parallel corpora enormously helpful as they enable them to compare the language they know (A) with the language / languages they are just learning

(B) / (C). The number of parallel corpora have been increasing in the consecutive years, which is a sign of their great popularity and some of them also include Slavonic languages. Such corpora can facilitate linguists' work and present in a quick way the parallel material in two, three and a larger number of languages.

4.2.1. Below there are examples of neologisms excerpted from 'Parallel Polish-Bulgarian-Russian Corpus' which is currently in preparation by: V. Koseska-Tosze-wa, J. Satoła-Staśkowiak, W. Sosnowski and A. Kisiel and K. Staśkowiak as part of the European Project Clarin⁹.

⁹(Common Language Resources and Technology Infrastructure) — a scientific project which achieved legal status of ERIC (European Research Infrastructure Consortium) of the European

Polish language:

Nie wolno dotykać **ekranu**, ponieważ może to doprowadzić do uszkodzenia powłoki albo niektórych **pikseli** służących do **generowania** obrazu. (Touching the **screen** is not allowed as it may lead to damaging the coating or some pixels used for **generating** the picture.)

Bulgarian language:

Винаги избягвайте да докосвате **екрана**, тъй като това може да доведе до повреда на **екрана** или в някои **пиксели**, които служат за изграждане на изображенията.

Russian language:

Всегда избегать касания **экрана**, так как это может привести к повреждению всего **экрана** или отдельных **пикселей**

Polish language:

Gdy **urządzenie pamięci masowej USB** jest podłączone do telewizora, nie wolno wyjmować go ani wyłączać telewizora, ponieważ może to spowodować uszkodzenie **plików** albo samego urządzenia.

Należy często **wykonywać kopie zapasowe plików** zapisanych w **urządzeniu pamięci masowej USB**, ponieważ ich utrata lub uszkodzenie nie są objęte gwarancją. (When the **device of mass memory USB** is connected to the computer it cannot be taken out and the TV set cannot be switched off as it can lead to damaging the **files** or the device itself. One should often **make spare copies of the files** recorded in **the device of mass memory USB** because their loss or damage is not under guarantee.)

Bulgarian language:

Не изключвайте телевизора и не изваждайте **USB устройството за съхранение**, докато то е включено в телевизора, тъй като това може да доведе до загуба на **файлове** или повреда на **USB устройството за съхранение**. **Архивирайте** често **файловете**, записани на **USB устройство за съхранение**, тъй като е възможно да ги загубите и този проблем да не се покрива от гаранцията ви.

Russian language:

Не выключайте телевизор и не извлекайте **устройство хранения USB**, когда оно подключено. В противном случае это может привести к удалению **файлов** или повреждению **устройства хранения USB**.

Commission in February 2012. The founders of Clarin ERIC are Austria, Bulgaria, the Czech Republic, Denmark, Estonia, Germany, Holland and Poland.

CLARIN is a project from so-called roadmap ESFRI (European Roadmap for Research Infrastructures, European Strategy Forum on Research Infrastructures). The main aim of the project is to combine the resources and language tools for European languages into one common standardized network which is to become an important tool of work for researchers from very broadly understood humanistic branches of science.

Bulgarian language:

Часто **создавайте резервные копии файлов**, сохраненных в памяти **устройства хранения USB**, так как файлы могут быть удалены или повреждены, но на данное повреждение гарантия может не распространяться.

Polish language:

Nieobsługiwane **pliki** są wyświetlane w podglądzie tylko w postaci **ikon**. (The **files** which are not supported are displayed in preview only in the form of **icons**.)

Nietypowe **pliki** są wyświetlane w postaci **map bitowych**. (Untypical **files** are displayed in the form of **bit maps**.)

Bulgarian language:

Неподдържаните **файлове** се показват само чрез предварително определена **икона**. **Файловете** с необичаен формат се показват във вид на **растерно изображение**

Russian language:

Неподдерживаемые **файлы** отображаются при предварительном просмотре только в виде **значка**.

Неправильные **файлы** отображаются в виде **растрового изображения**.

4.2.2. Apart from the advantages connected with the use of resources of parallel corpora, I can also observe some technical problems. The main one is the fact that a substantial part of submitted parallel corpora¹⁰ (two- or multilingual) is under preparation and the language material they include is not known to a wider audience since at this stage it is not available in the Network. It poses a real obstacle for the users of corpora.

Parallel corpora can only be helpful in contrastive studies on neologisms when their resources include the newest texts from the 21st century — one of the corpora which fulfils such criteria is The Bulgarian-Polish Corpus which is being prepared by L. Dimitrowa and V. Koseska-Toszewa and which includes, e.g. the documents of the EU.

5. Briefly on the methodology of excerption of neologisms by means of corpora.

Not all the new words we hear can be called neologisms. The process of searching the corpora resources should be based on certain rules of ‘checking’ new lexems to make sure that they are rooted in the system of a given natural language. The

¹⁰An example of a corpus with a strictly restricted group of recipients is: French-Bulgarian Parallel Corpus, realized as a part of Clarin: <http://www.clarin.eu/french-bulgarian-parallel-corpus>; an example of a corpus in which there are Polish, Bulgarian and Russian texts among other materials is Intercorp: <http://www.korpus.cz/intercorp/?req=page:info&lang=en>. The difficulty may lie in the fact that the superior language in Intercorp is Czech. Under preparation, due to the project Clarin, there is the Parallel Polish-Bulgarian-Russian Corpus which is currently being worked on by V. Koseska-Toszewa, J. Satoła-Staškowiak, W. Sosnowski and A. Kisiel.

following facts can show whether a given lexical unit has entered permanently into the grammatical system of the languages compared in my research:

5.1. In Polish these are inflectional endings added to a new lexical unit, inflection of new lexems and multiplication of words within the range of formative family. A Polish example of a borrowing which has entered the language system is *blog*:

- Otóż ukończywszy w pocie czoła olbrzymią kampanię dla naszych graczy, która przeprowadzi ich przez dwadzieścia poziomów podziemi i pół znanego w danym systemie świata, zakładamy *blog*, który będzie nas wspomagał. (So, having completed by the sweat of our brows a huge campaign for our players, which will lead them through twenty levels of the underworld and half of the world known in a given system, we're setting up a **blog**, which will support us.)
- Zajrzałem na jego *bloga*. (I took a look through his **blog**.)
- Są dwa sposoby: a) jednostkowy, b) wspólnotowy. Pierwszy to założenie własnego *bloga* lub strony internetowej i zamieszczenie tam napisanego dzieła. (There are two ways: a) individual, b) joint. The first one is to set up one's own **blog** or Internet website and placing there the written work.)
- Ja jako uważny obserwator Waszych poczynąń o samego początku na tym portalu pragnę stwierdzić, że jesteście najczęściej odwiedzanym, *blogiem* na tym portalu wśród publicystów i dziennikarzy piszących tutaj swoje *blogi*. (As an attentive observer of your doings from the very beginning on this portal I wish to state that you are the most frequently visited **blog** among the columnists and journalists writing their **blogs** here.)
- Antoni Bielewicz nie tylko będzie *blogował* o tym, co ciekawego dzieje się na SAP TechEd, lecz także przeprowadzi szereg podcastowych wywiadów z wybranymi uczestnikami konferencji — dodaje Michał Kreczmar, wydawca serwisów online w International Data Group Poland S.A. Konferencja SAP TechEd odbędzie się w dniach 18–20 października. (Antoni Bielewicz will not only **blog** about what's interesting at SAP TechEd but he will also conduct a number of podcast interviews with the chosen participants of the conference — adds Michał Kreczmar, an editor of online services in the International Data Group Poland S.A. The SAP TechEd conference will take place on 18–20th October.)
- Najbardziej wpływowy *bloger* 2012 roku. (The most influential blogger 2012.)

Another example of a word 'assimilated' on the Polish ground can be *laptop* or *hit*:

- Włączyłem *laptop*. Wpatrywałem się w monitor, jakby to z niego, a nie z mojej głowy wyłaniały się pytania. (I switched on my **laptop**. I stared at the monitor, as if the questions were emerging from it and not from my head.)

- Włączył laptopa, poczekał, aż uruchomią się programy, i kliknął na ikonę wyszukiwarki. (He switched on his **laptop**, waited until the programs started and clicked on the icon of a search engine.)
- W laptopie włączył bostoński koncert zespołu Fleetwood Mac, z czasów gdy Peter Green nie popadł jeszcze w obłęd i był geniuszem gitary. (He turned on the Boston concert of Fleetwood Mac in his **laptop**, from the times when Peter Green hadn't gone mad yet and was a guitar genius.)
- Wiadomo, że baterie litowo-jonowe Sony są używane także w laptopach Lenovo, ale inżynierzy tej firmy są pewni, że nie dojdzie do wypadków dzięki odpowiedniej konfiguracji. (It is known that Sony lithium-ionic batteries are also used in Lenovo laptops but the engineers of this company are sure that accidents will not happen thanks to appropriate configuration.)
- Alfabetyczne **hity** ostatniego roku. (Alphabetical hits last year.)
- Wiele teatrów przyjeżdża z teatralnymi **hitami**. (Many theaters come from theatrical hits.)
- Co miesiąc proponujemy Ci Płytę Miesiąca — prawdziwy **hit**. (Every month, we offer you the disc Month — a real hit.)
- Nie, zaczął opowiadać, że to listy **hitów**. (He began to tell, it's a list of hits.)

5.2. In the Bulgarian language a newly created or newly borrowed and accepted by the users word receives articles and often, though not always, plural form. Let us examine the same example:

- Блогът няма стриктно предписана структура но един типичен **блог** би могъл да съдържа заглавие, главна част, връзка с цялата статия и дата на публикуване.
- Политическият **блогър** от този вид наистина има по-скоро самочувствие-то на хакер и със сигурност стои на страната на торентите, но той действа позитивно и там, където обичайните средства са изчерпани.
- Форумът в **блога** му още не е доразработен, но той щял да е едно от най-важните неща в **блога**.

Or another:

- **Хит** в програмата е услугата "Лазерно вагинално подмладяване", което струва 3500 евро и се ползва от хиляди американки.
- Десетте най-големи летни **хита**, избрани от слушателите на БГ радио.
- SMS и имейл ще са **хитът** в агитацията и зарибяване на електората на тези избори, твърдят запознати.

5.3. In the **Russian language**, similarly to Polish, full acceptance of a lexeme by the grammatical system of a natural language takes place when inflectional endings develop as well as when a newly created word undergoes declensional process as other units of the same language. One can also observe the multiplication of words within the created derivational family and case inflection of new lexemes.

- За это **блогер** хочет не столько похвалы, сколько признания.
- Попробуй завести свой личный **блог** и наполнять его регулярно!
- Журналы для детей, о детях и родителях, для родителей, порталы, **блоги** и форумы — всего этого так много, что просто разбегаются глаза, и невольно встает вопрос, а нужно ли их столько, и в чем их предназначение?
- Можно найти в инете порядка двух десятков **блогов** и форумов пропагандистского содержания.

Or another:

- Пшеница — **хит** сезона на продовольственном рынке Рынки сельскохозяйственной продукции сезонны и высокочувствительны к изменениям погоды.
- Толпа талантливейших исполнителей (ни одного случайного или бездарного) перепевает сочинения Ли Хэзлвуда, знаменитого поп-композитора, автора, в частности, многих **хитов** Нэнси Синатры.
- Год назад я старалась держаться подальше от евреев, а фраза "Понаехали" и другие московские **хиты** в адрес приезжих часто звучали из моих уст.
- Этакий антибойз-бэнд: три совершенно разных и почти не умеющих петь человека исполняют реально зажигательные песни: **хит** на **хите**...

6. Summary

6.1. In some countries (France can be the best example of such conduct.) *neologisms* are created in the most important circles of language councils which out of concern for the clarity of their language look for new terms to avoid borrowings from foreign languages. In the three Slavonic languages (Polish, Bulgarian and Russian) which I analyse in the article a different situation takes place where new lexical units begin to exist in the language system as a result of inventing a new word and its frequent use, thanks to borrowing words from foreign languages or as a result of using an existing word with a new meaning.

One can observe many ways of enriching and diversifying vocabulary within the range of the languages compared in the article. Among other things, the most

popular are: derivation, borrowings and neosemantization. The processes of creating new words or the processes of neosemantization, observed among other things thanks to the corpora, are dependent today on the strong influences of the English language. Native words are driven out by internationalisms (whose numerous examples can also be found in this text) which facilitate communication on the Internet forums, using computer software, modern mobile phones, etc. It is one of many noticeable globalisation processes of contemporary times. These processes are unavoidable and they result from a need for faster interpersonal, social, etc. communication.

6.2. Thanks to corpora the users receive reliable information about the frequency and range of use of a given language element. In order to obtain reliable contrastive material, all kinds of language corpora are useful. Certainly, parallel corpora are more practical, 'handy', because in one place, as it were next to each other, we have units which are of interest to us. What is more, they are in the same or similar contexts. However, there is no scientific evidence that parallel corpora are better than the ones which are 'based' on one language. Thanks to the latter ones researchers will not omit these aspects of units (neologisms) which do not have a parallel character (e.g. a different degree of semantic assimilation of polysemic units).

Diverse limitations are a big problem: the necessity to pay subscription for access online to the whole corpus or severely restricted access to corpus on account of works it is undergoing (as it is the case with parallel corpora with Slavonic languages described in this article) or a foreign language of operation of a given corpus, as in the case of the corpus described above (cf. Intercorp). Very often, though lack of other materials, the Internet resources, available thanks to search engines, have to be used in the role of a corpus.

6.3. The corpora which will support my research on neologisms have been presented here as the collections of contexts. The linguistic context confirms the suitability of contrastive meanings of lexemes. It is of particular importance in situations where a new word is not an anglosemantism and is not overused and yet it is present in the system of the examined language and the verification of its meaning and frequency of its use are necessary.

6.4. In the Polish, Bulgarian and Russian languages, as the language corpora used for comparison confirm, **the areas** (specialised vocabulary: technical, computer, youth vocabulary, etc.), **the ways** (semantic borrowings, metonymy, metaphor, opposites, abbreviating, generalising) and **the reasons for creating new meanings** language fashion, a need for naming new things, phenomena or a need for adding variety to the language of communication (e.g. expressiveness of a language), a desire to exist in virtual or media reality **often coincide**. All the three languages described in the article assimilate many new words from the field of economics, law, medicine, technical sciences and popular colloquial expressions. The quantity of words created every year and the number of lexemes borrowed from English in each of the languages is similar although not identical. In spite of a huge level

of activity in creating and borrowing new words by all the languages described in the article, the users of Bulgarian (the result of the confrontation of a bigger number of neologisms) seem to be a little less resistant to English-language influences whereas the users of Russian display less interest in linguistic borrowings. As the preliminary contrastive research shows Polish is situated between the two above mentioned Slavonic languages although it has many features similar to Bulgarian. An attentive observer can notice in the area of **computer terminology** of the **Bulgarian language** an exceptional popularization e.g. of words such as: *даунлоад*, *даунлоадам*, *атачмѐнт*, *блутуут*, etc., whereas in the **Polish language** words which do not have their Polish equivalent are more frequently assimilated, most often when we describe new devices, such as: *laptop*, *tablet* and *palmtop*. In other cases words are used which have already existed in the system of the Polish language, cf. *pobierz* (download), *pobierać*, *załącznik* (attachment), but also *blue-tooth* (although with their original spelling preserved, which may prove a lack of acceptance in the system of the Polish language) etc. In the **Russian language**. Equivalents such as: *скачай*, *скачивать*, *вложение* (*присоединенный файл*) and *блютуз* show tendencies similar to the Polish language.

English-language **sports terminology** is accepted and used in the **Bulgarian language** more than in Polish and Russian, e.g. Bulg. *байк*, *байкѐт*, *байкове*; *даунхил*. In the **Polish language** sports terminology, if it is possible, remain native, cf. Pol. *rower* (*górski*) (mountain bike), *góral* (climber); *kolarstwo górskie* (mountain biking); Russ. *горный велосипед*; *горный велоспорт*; *маунтинбайкин*

The increased assimilation of anglicisms in the Bulgarian and Russian languages is noticeable in **technical terminology**. The Polish language is characterized by 'higher frequency of native conventional lexis and abstinence from borrowings', (cf. J. Mazurkiewicz-Sułkowska, 2013, p.6) e.g.: Pol. *wirówka* (centrifuge), Russ. *центрифу́га*, Bul. *центрифу́га* (examples of technical terminology after Mazurkiewicz-Sułkowska)

This regularity is confirmed by the neologisms I am studying in the language corpora.

6.5. The summary of the article concerning neologisms in three different, although related Slavonic languages and their excerpt thanks to the newest linguistic tools should be considered as preliminary contrastive deliberations which will become the origin of a bigger dissertation on neologisms. The source of obtaining the newest lexis will become not only the corpora but first of all book and electronic dictionaries, the Internet, subject literature and field research. I believe that intensive research of scholars on language corpora will enrich my studies with a lot of interesting and reliable language materials which, through confrontation, will most probably allow me to show interesting conclusions.

References

- Blagoeva, D. (2001). Д. Благоева, Интернационализация и неологизация (наблюдения върху неологизмите с компонент *пост-/post-* в българския и чешкия език). In: Арnaudов сборник. Доклади и съобщения. vol. 2. Русе, "Лени Ан", p. 187–189.

- Blagoeva, D. (2003). Д. Благоева, Интернационални и национални елементи в речниковия състав (аспекти на калкирането в съвременните славянски езици). In: *Internacionalizmy v nově slovní zásobě*, Praha, UJČ AV ČR, p. 133–140.
- Blagoeva, D. (2004). Д. Благоева, Един тип неолотизми универбати в българската разговорна реч (в съпоставка с руски). *Проблеми на българската разговорна реч*, 6. Велико Търново, УИ "Св. Св. Кирил и Методий", p. 45–52.
- Blagoeva, D. (2005). Д. Благоева, За някои неправилни или ненужни калки в компютърната терминология. *Български език*, 1, София, Академично издателство "Проф. Марин Дринов", p. 117–118.
- Blagoeva, D. (2005). Д. Благоева, За някои нови интерсемантизми в славянските езици,. In: *Юбилеен славистичен сборник*. Благоевград, УИ "Неофит Рилски", p. 102–108.
- Blagoeva, D. (2006). Д. Благоева, Неолотизмите в съвременния български език. *Български език и литература*, 6, p. 16–25.
- Sotirov, P., Mostowska, M., Mokrzycka, A. (2011). *Bułgarsko-polski słownik leksyki potocznej*, vol. 1. Lublin, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, p. 1–280.
- Buttler, D. (1962). Neologizm i terminy pokrewne. *Poradnik Językowy*, vol. 5, 6. Kraków, Państwowe Wydawnictwo Naukowe, p. 239–251.
- Buttler, D. (1977). Neologizmy potoczne powojennej polszczyzny. *Przegląd Humanistyczny*, 12. Warszawa, Państwowe Wydawnictwo Naukowe, p. 99–105.
- Buttler, D. (1978). Powojenne innowacje w polskim zasobie słownym. *Przegląd Humanistyczny*, 5. Warszawa, Państwowe Wydawnictwo Naukowe, p. 55–67.
- Buttler, D. (1979). *Powojenne ekspresywizmy polskie*. *Prace Filologiczne*, XXIX, Warszawa, Wydawnictwo Uniwersytetu Warszawskiego, p. 85–90.
- Брагина, А. А. (1973). *Неологизмы в русском языке*. Москва.
- Chruścińska, K. (1975). Rzeczownikowe neologizmy słowotwórcze w języku prasy (derywaty proste). *Poradnik Językowy*, 10, Warszawa, Państwowe Wydawnictwo Naukowe, p. 560–567.
- Chruścińska, K. (1976). Nowe wyrazy złożone w języku polskiej prasy. *Poradnik Językowy*, 11, Warszawa, Państwowe Wydawnictwo Naukowe, p. 438–443.
- Chruścińska, K. (1977). Nowe wyrazy złożone w języku polskiej prasy. *Poradnik Językowy*, 1, Warszawa, Państwowe Wydawnictwo Naukowe, p. 19–27.
- Chruścińska, K. (1978). O formacjach potencjalnych i okazjonalizmach. In: M. Szymczak (Ed.), *Z zagadnień słownictwa współczesnego języka polskiego*, Wrocław–Warszawa–Kraków–Gdańsk, Zakład Narodowy im. Ossolińskich, p. 69–79.
- Dąbrowska, K. (1957). Neologizmy słowotwórcze w publicystyce polskiej okresu Odnowy, *Zeszyty Naukowe WSP w Opolu. Językoznawstwo*, 1, Warszawa, Państwowe Wydawnictwo Naukowe, p. 39–99.
- Dimitrova, L., Koseska-Toszeva, V., Satola-Staškowiak, J. (2012). About Neologisms in Bilingual Dictionaries (on example of Bulgarian-Polish Dictionary). *Cognitives Studies / Etudes cognitives*, 12, p. 107–114
- Grabias, S. (2001). Środowiskowe i zawodowe odmiany języka — socjolekty. J. Bartmiński (Ed.), *Współczesny język polski*, Lublin, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, p. 235–253.
- Jadacka, H. (2001). *System słowotwórczy polszczyzny (1945–2000)*, Warszawa, Państwowe Wydawnictwo Naukowe, p. 1–201.
- Jadacka, H. (2009). Nieporozumienia wokół nazw żeńskich. In: M. Skarżyński, M. Szpiczawska (Eds.), *W kręgu języka, Materiały konferencji „Słowotwórstwo — słownictwo — polszczyzna kresowa” poświęconej pamięci Profesor Zofii Kurzowej, Kraków 16—17 maja 2008*, Kraków, p. 277–284.

- Jadacka, H. (2010). Neologizmy jako przedmiot opisu lingwistycznego. *LingVaria*, Rok V, 1 (9) http://www2.polonistyka.uj.edu.pl/LingVaria/archiwa/LV_1_2010_pdf/004_Jadacka.pdf
- Kirova, L. (2004). Л. Кирова, Особенности на чат-разговора. In: *Проблеми на българската разговорна реч, книга шеста*. Велико Търново, Университетско издателство "Св. св. Кирил и Методий", p. 135–144.
- Kolkovska, S. С. Колковска, *Динамика и утвърждаване на новите лексикални значения в български език*, www.balgarskieszik.org
- Kolkovska, S. (2007). С. Колковска, Нови икономически названия в български език, възникнали от медицински термини. In: *Лексикогрофията и лексикологията в съвременния свят. Материали от IV национална конференция с международно участие на Българското лексикографско дружество в чест на проф. Кристилина Чолакова, София, 21–22 октомври 2006*. Велико Търново, ИК "Знак'94", p. 303–310.
- Kolkovska, S. (2008). С. Колковска, Автоматично разпознаване на концептуално-семантични отношения в електронен корпус от биологически текстове и приложението му в лексикографска практика. In: *Lexikografie v kontextu informační společnosti*. Praha, UJČ AV ČR, p. 181–192.
- Kolkovska, S. (2010). С. Колковска, Лексикални неологизми в български език, възникнали от чужди инициални абривиатури. *Български език*, 4, p. 21–30.
- Kolkovska, S. (2011). С. Колковска, Семантични неологизми в съвременната българска терминология, възникнали чрез ретерминологизация. In: *Езиковедски изследвания в чест на проф. Сийка Спасова-Михайлова*. София, АИ "Проф. Марин Дринов", p. 519–532
- Koseska-Toszewa, V., Satoła-Staškowiak, J., Sosnowski, W. (2013). From the Problems of Dictionaries and Multi-lingual Corpora. *Cognitive Studies / Études cognitives*, 13, (in this volume).
- Kozyrev, V. A., Czerniak, V. D. (2000). В. А. Козырев, В. Д. Черняк, *Вселенная в алфавитном порядке*. In: *Очерки о словарях русского языка*. СПб.
- Lopatin, V. V. (1973). В. В. Лопатин, *Рождение слова. Неологизмы и окказиональные образования*. Москва.
- Markowski, A. (1988). Neologizmy z formantem *-arka* w ogłoszeniach prasowych. *Poradnik Językowy*, 3, p. 232–234.
- Markowski, A. (1994). Nowsze neologizmy polszczyzny potocznej. *Poradnik Językowy*, 1–2, p. 73–76.
- Markowski, A. (2006). *Kultura języka polskiego. Teoria. Zagadnienia leksykalne*, Warszawa, Wydawnictwo Naukowe PWN, p. 1–292.
- Mazurkiewicz-Sułkowska, J. (2013) Współczesna terminologia techniczna w językach słowiańskich (na materiale polskim, rosyjskim i bułgarskim). In: *Rozprawy Komisji Językowej ŁTN* (in print)
- Nadel-Czerwińska, M. (2006) *Nowa leksyka w słownikach rosyjskich na końcu XX i początku XXI wieku*, ISBN 83-226-1528-0, Białystok, Wydawnictwo Uniwersytetu w Białymstoku, p. 1–172.
- Orylska, H. (1983). Neologizmy słownikowe w dokumentach partyjnych z lat 1949–1953, „Acta Universitatis Nicolai Copernici”. *Filologia Polska*, 23, p. 65–71.
- Ożóg, K. (2001). *Poliszczyna przelotu XX i XXI wieku. Wybrane zagadnienia*, Rzeszów. Biblioteka Frazy.
- Percova, N. N. (1995). Н. Н. Перцова, *Словарь неологизмов Велимира Хлебникова*. Вена–Москва.

- Perniška, E. (2005). E. Пернишка, Съвременни явления и отношения в огледалото на българския език. In: *Езиковедски приноси в чест на чл.-кор. Михаил Виденов*. Велико Търново, Университетско издателство "Св. св. Кирил и Методий", p. 541–551.
- Perniška, E., Blagoeva, D., Kolkovska, S. (2010). E. Пернишка, Д. Благоева, С. Колковска, *Речник на новите думи в българския език*, първо издание, София, Института за Български Език "Проф. Л. Андрейчин", Наука и Изкуство, p. 1–515.
- Podhajecka, M. *Korpusy językowe — jak mogą pomóc w nauce języka obcego?*, <http://spnjo.polsl.pl/konferencja/materialy/referaty/podhajecka.pdf>.
- Satkiewicz, H. (1976). O niektórych innowacjach słotwórczych w polszczyźnie XX wieku. In: J. Magnuszewski (Ed.), *Z problemów współczesnych języków i literatur słowiańskich*. Warszawa *Materiały z Konferencji Uniwersytetu Warszawskiego i Uniwersytetu Karola dla uczczenia trzydziestolecia PRL i trzydziestej rocznicy wyzwolenia Czechosłowacji*, p. 111–121.
- Satkiewicz, H. (1981). Innowacje słotwórcze w powojennym trzydziestoleciu. In: H. Kurkowska (Ed.), *Współczesna polszczyzna*, Warszawa, Państwowe Wydawnictwo Naukowe, p. 130–155.
- Satoła-Staśkowiak, J. (2009a). Translating into something that does not exist... — literary ways of translating Polish sentences with uninflected perfect participles into the Bulgarian language. *Cognitive Studies / Études cognitives*, 9, p. 211–221.
- Satoła-Staśkowiak, J. (2009b). Zdania z zawieszoną asercją na podstawie polskich i bułgarskich zdań praeterytalnych. In: V. Koseska-Toszeva, R. Roszko (Eds.), *Semantyka a konfrontacja językowa*. vol. 4, Warsaw, Instytut Sławistyki PAN, Sławistyczny Ośrodek Wydawniczy, p. 101–118.
- Satoła-Staśkowiak, J. (2010). *Polsko-bułgarskie odpowiedniości przekładowe czasów przeszłych*, Warszawa, Instytut Sławistyki PAN, Sławistyczny Ośrodek Wydawniczy, p. 1–144.
- Skubulanka, T. (1990). Problematyka gramatyczno-stylistyczna neologizmu. *Studia Linguistica Polono-Slovaca*, 3, p. 31–37.
- Smółkowa, T. (2001). *Neologizmy we współczesnej leksyce polskiej*, Kraków, Wydawnictwo Instytutu Języka Polskiego Polskiej Akademii Nauk, p. 194.
- Smółkowa, T. (2001). Nowe słownictwo polskie. In: J. Bartmiński (Ed.) *Współczesny język polski*, Lublin, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, p. 397–404.
- Sosnowski, W. (2013). Forms of address and their meanings in Polish and Russian languages in contrast. *Cognitive Studies / Études cognitives*, 13, (in this volume).
- Tokarski, R. (2001). Słownictwo jak interpretacja świata. In: J. Bartmiński (Ed.) *Współczesny język polski*, Lublin, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, p. 343–370.
- Waszakowa, K. (2009). Derywowane anglicyzmy jako wyraz ekspansji słownictwa potocznego i środowiskowego w języku mediów ostatniego ćwierćwiecza. *Slavia Meridionalis*, 9, Warszawa, Sławistyczny Ośrodek Wydawniczy, p. 1–14.
- Waszakowa, K. (2012). Drogi do słownika neologizmów współczesnej polszczyzny. *Poradnik Językowy*, 3, Warszawa, Państwowe Wydawnictwo Naukowe, p. 5–21.

Corpora

Corpus Encoding Standard: <http://www.cs.vassar.edu/CES/CES1-0.html>
Corpus Typology: <http://webdeptos.uma.es/filifa/personal/amoreno/teaching/cl/corpus-typ.html>
EAGLES Guidelines: <http://www.ilc.pi.cnr.it/EAGLES96/browse.html>
Korpus IPI PAN: <http://korpus.pl>
Korpus PWN: www.korpus.pwn.pl
Korpus referencyjny języka polskiego PELCRA: <http://korpus.ia.uni.lodz.pl/index.php>
Korpus CLARIN: <http://www.clarin.eu/french-bulgarian-parallel-corpus>
Korpus Intercorp: <http://www.korpus.cz/intercorp/?req=page:info&lang=en>
Narodowy Korpus Języka Rosyjskiego: <http://www.ruscorpora.ru/search-main.html>
Korpus Języka Bułgarskiego IBE BAN: <http://search.dcl.bas.bg/>