

**AUTOMATYZACJA WIZUALIZACJI GRAFOWEJ SYNONIMÓW  
DLA POTRZEB DYDAKTYKI JĘZYKÓW OBCYCH  
(NA PRZYKŁADZIE PRZYMIOTNIKÓW ROSYJSKICH  
Z PREFIKSEM БЕЗ-/БЕС-)**

**АВТОМАТИЗАЦИЯ ГРАФОВОЙ ВИЗУАЛИЗАЦИИ  
СИНОНИМОВ В ЛИНГВОДИДАКТИЧЕСКИХ ЦЕЛЯХ  
(НА ПРИМЕРЕ ИМЕН ПРИЛАГАТЕЛЬНЫХ РУССКОГО  
ЯЗЫКА С ПРЕФИКСОМ БЕЗ-/БЕС-)**

**THE AUTOMATED GRAPH VISUALIZATION OF SYNONYMS  
FOR THE PURPOSES OF THE DIDACTICS  
OF FOREIGN LANGUAGES  
(ILLUSTRATED WITH AN EXAMPLE OF RUSSIAN  
ADJECTIVES WITH БЕЗ-/БЕС- PREFIXES)**

**Daniel Dzienisiewicz, Yury Fedorushkov**

Uniwersytet im. Adama Mickiewicza w Poznaniu, Poznań – Polska

**Abstract:** The article presents methods for application of lexicographic data for the purposes of foreign language teaching and linguistic studies. The automated graph visualization technique is introduced. Both the retrieval and the visualization of linguistic data in form of directed graphs have been conducted with the use of available computer software. The material on which the above-mentioned operations have been performed comprises the group of Russian adjectives with *без-/бес-* prefixes and the synonymic chains initiated by them. The automated extraction with the use of regular expressions, the steps leading to the edition of the generated database and the import of the final data base into a visualization software are described in the article. The grammatical and semantic characteristics of the sample group of lexemes are not the focus of this article, although further analysis of the selected vocabulary units is planned to be conducted in future studies.

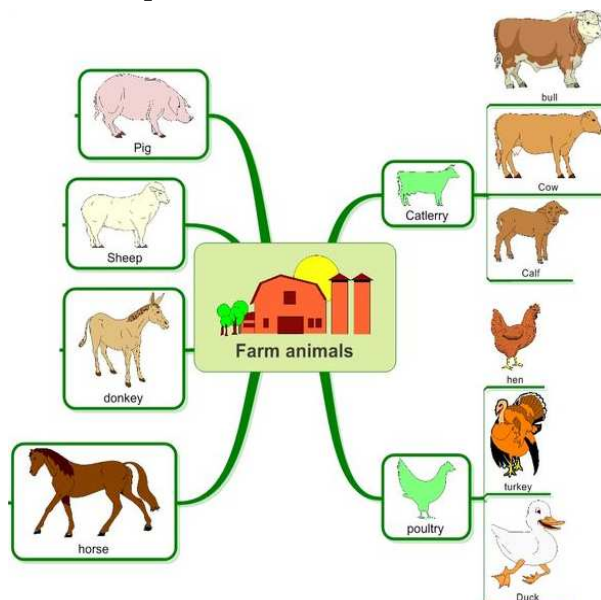
**Słowa kluczowe:** dydaktyka języków obcych, synonimia, leksykografia komputerowa, techniki ekscerpcyjne, wyrażenia regularne, wizualizacja grafowa

**Ключевые слова:** лингводидактика, синонимия, компьютерная лексикография, техники экскерпции, регулярные выражения, графовая визуализация

**Keywords:** foreign language teaching, synonyms, computational lexicography, retrieval techniques, regular expressions, graph visualization

Graficzne przedstawianie zależności językowych znajduje szerokie zastosowanie w dydaktyce języków obcych (DJO). Zgodnie z założeniami metodyki nauczania języków obcych odpowiednio skonstruowane

pomoce wizualne uatrakcyjnają lekcję i koncentrują uwagę uczącego się na jej przedmiocie. Ponadto stosowanie materiałów wizualnych pobudza procesy pamięci wzrokowej i ośrodki obu półkul mózgowych, ułatwiając zapamiętywanie materiału leksykalnego<sup>1</sup>. Dysponując ową metodologią z zakresu DJO, skłaniamy się ku wykorzystaniu wizualizacji również jako narzędzia pomocniczego w szeroko pojmowanych badaniach językoznawczych. We współczesnej DJO powszechnie używane jest określenie „mapa myślowa”<sup>2</sup>. Prezentacja określonych grup leksykalnych (np. rzeczowników konkretnych powiązanych odpowiednią relacją) może odbyć się za pomocą grafu, w którym oprócz wyrazów zastosowane są obrazki, np.:



Ryc. 1. Przykład tzw. mapy myśli w metodzie Callana<sup>3</sup>

<sup>1</sup> Por. H. Komorowska, *Metodyka nauczania języków obcych*, Warszawa 2004, s. 56.

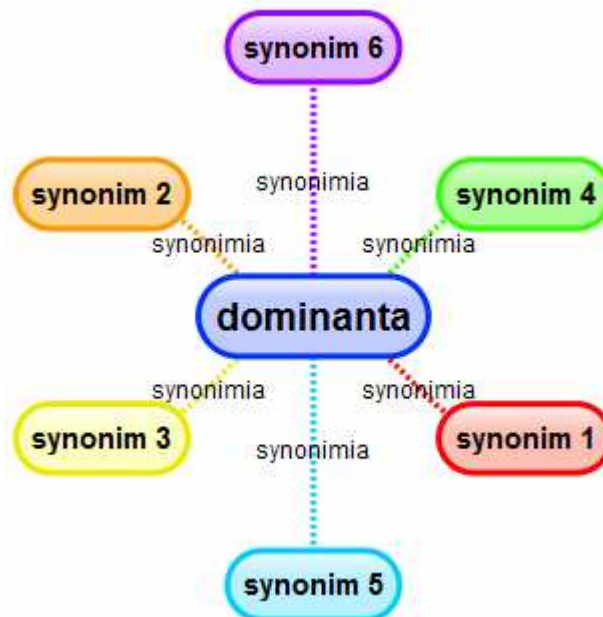
<sup>2</sup> Szczegółowy opis – por. m. in. T. Buzan, B. Buzan, *Mapy Twoich Myśli*, Łódź 2003.

Wykorzystanie map myśli w dydaktyce oraz ogólnie o mapach myśli jako nieliniarnej metodzie notowania podczas wykładów, lekcji itp., por. – M. Janowicz, *Wykorzystanie map myśli w dydaktyce*, „Oeconomica: Folia Pomeranae Universitatis Technologiae Stetinensis” 2011, no. 287 (63), s. 65–72, [w:] źródło elektroniczne: <http://www.wydawnictwo.zut.edu.pl/fileadmin/pliki/wydawnictwo/Folia/Oeconomica/287/Janowicz.pdf> (03.12.2013).

Wykorzystanie map myśli w nauczaniu innych dyscyplin, np. matematyki – por.: A. Brinkmann, *Graphical Knowledge Display – Mind Mapping and Concept Mapping as Efficient Tools in Mathematics Education*, „Mathematics Education Review” 2003, no. 16, s. 35–48, [w:] źródło elektroniczne: [http://www.marccouture.com/formations/cartes\\_mentales/LinkedDocuments/mereview-16-Apr-2003-4.pdf](http://www.marccouture.com/formations/cartes_mentales/LinkedDocuments/mereview-16-Apr-2003-4.pdf) (23.12.2013).

<sup>3</sup> Źródło elektroniczne zdjęcia: <http://www.nauka-jezykow.net/images/nauka-jezykow-mapy-mysli.jpg> (23.12.2013).

Pojęcia abstrakcyjne mieszczą się w nie mniej licznych grupach leksykalno-tematycznych, kategoryalnych, szeregach synonimicznych. Tak na przykład, wyraz *szybki* w najnowszym słowniku synonimów Wojciecha Broniarka (<http://www.synonimy.pl/>) liczy 75 synonimów, natomiast wyraz *ostrzy* – 121 słów oraz zwrotów bliskoznacznych. Rzecz jasna, dodawanie obrazków do wyrazów *zaostrzony*, *kłujący*, *mocny* graniczy z niemożliwością, a manualne tworzenie poszczególnych map myślowych cechuje duża czasochłonność. W niektórych oprogramowaniach do poszczególnych obiektów grafu można dołączać linki, pliki audiowizualne, tj. „obciążać” obiekty dodatkowymi informacjami. Jest to dużym plusem dla procesu dydaktycznego oraz przyswajania wiedzy. Oprogramowania służące do manualnego tworzenia prezentacji w postaci grafów – to m. in. *SimpleMind Desktop*<sup>4</sup>, a także np. *Skitch*, *Prezi Desktop* i in., np.:



Ryc. 2. Tworzenie sieci relacji synonimicznej w środowisku *SimpleMind Desktop*

Jednak przy 121 synonimach musielibyśmy dodawać każdy z obiektów do takiego edytora grafowego wyłącznie manualnie.

W związku z tym chcielibyśmy zaproponować metody konstruowania minimapy myślowej w sposób automatyczny, przewidujący niezasochłonny dobór materiału badawczego, tworzenie arkusza dla dobra-

<sup>4</sup> Wersja „trial” oraz wersja odpłatna, [w:] źródło elektroniczne: <http://www.simpleapps.eu/simplemind/desktop#trial> (14.10.2013).

nych obiektów, umieszczenie arkusza w bazie danych odpowiedniego oprogramowania, za pomocą którego z kolei dokonujemy rozmieszczenia tych obiektów. Swoje metody przedstawimy na przykładzie synonimów języka rosyjskiego. Zademonstrujemy mianowicie metody automatyzacji wizualizacji grafowej szeregów synonimicznych przymiotników języka rosyjskiego z prefiksem *без-/без-*.

Etapy eksperymentu:

- 1) Automatyczna ekscerpacja przymiotnikowych artykułów hasłowych słownika z prefiksem *без-/без-* za pomocą tzw. *filtrów wyrażeń regularnych* (por. niżej);
- 2) Eliminacja nieprzydatnych informacji z treści hasłowej;
- 3) Zapis otrzymanej listy w formacie pliku CSV;
- 4) Import listy do programu Gephi (por. opis niżej) i automatyczne generowanie grafów na podstawie zaimportowanych danych;
- 5) Edycja otrzymanej siatki grafu w celu nadania jej odpowiedniej przejrzystości i czytelności.

Przeprowadziliśmy serię eksperymentów polegających na tzw. *wizualizacji grafowej*<sup>5</sup> wybranych grup wyrazowych powiązanych relacją synonimii. Na bazie dostępnych słowników elektronicznych podjęliśmy próbę stworzenia automatycznej wizualizacji wybranych szeregów synonimicznych przy wykorzystaniu dostępnych programów komputerowych. Aby tego dokonać, należało przeprowadzić dobór oraz edycję materiału słownikowego, a także – ostatecznie – samą wizualizację, tj. operacje z grafiką grafów. Gotowy produkt eksperymentu przedstawia rycina 8 (por. niżej).

W sprawie opisu grupy przymiotników należy wspomnieć, że ich właściwości semantyczne i charakterystyka gramatyczna nie są przedmiotem badawczym w niniejszym artykule. Demonstrujemy jedynie metody wykorzystania gotowych danych leksykograficznych dla potrzeb dydaktyki. Zdajemy sobie sprawę, że niektóre treści akademickie (np. *Русская грамматика*, pod red. N. J. Szwedowej<sup>6</sup>) lub takie prace, które powtarzają akademickie uogólnione sformułowania, nie uwzględniają badań szczegółowych. Niemniej jednak należy zaznaczyć, że grupę dobranych leksemów w opisywanym przez nas eksperymencie stanowią przymiotniki języka rosyjskiego zawierające w swojej strukturze słowotwórczej prefiks *без-/без-*, tj. przymiotniki ze znaczeniem prywatywności (por. niżej), oraz wyrazy tworzące szeregi synonimiczne dla każdego przymiotnika<sup>7</sup>. W języku rosyjskim semantyka prywatywności (rozu-

<sup>5</sup> Zob. podstawowe wiadomości z dziedziny teorii grafów, [w:] źródło elektroniczne: <http://www.mini.pw.edu.pl/MiNIWyklady/grafy/grafy.html> (16.12.2013).

<sup>6</sup> *Русская грамматика*, ред. Н. Ю. Шведова, т. 1-2, Москва 1980.

<sup>7</sup> Np. dla wyrazu *беспокойный* synonimami są *бурный, бурливый, буйный, неугомонный, бесшабашный, вертлявый, игривый, баловливый, суетливый, резвый, непоседливый, подвижной,*

miana jako 'brak cechy nazwanej wyrazem podstawowym') wyrażana jest za pomocą przedrostków *a-*, *без-/бес-* oraz – bardzo rzadko – *вне-* i *не-*. Wyrazy z prefiksem *a-* są mało produktywne i motywowane są z reguły słowami o pochodzeniu greckim lub łacińskim. Drugi przedrostek (*без-/бес-*), będący przedmiotem naszego zainteresowania w niniejszym tekście, najczęściej<sup>8</sup> łączy się w procesie słowotwórczym z bezprefiksalnymi przymiotnikami motywowanymi rzeczownikami zarówno rdzennie rosyjskimi, jak i stanowiącymi grupę zapożyczeń. Najczęściej również w funkcji podstawy słowotwórczej dla omawianych przymiotników występują przymiotniki z sufiksami *-и-* oraz *-ов-/ев-*. Należy jednak wspomnieć w tym miejscu, że wśród danej listy znajdują się także przymiotniki z sufiksem zerowym (typu *безусый, беззубый, безголовый*), które tworzone są od rzeczowników z sufiksem zerowym. Niektóre wyrazy z tej grupy opierają się na model przenośni, tracą zatem znaczenie prywatności, np. wyrazy *безмозглый, бессердечный*.

БЕЗМОЗГЛЫЙ, -ая, -ое. Презрит. Очень глупый, бестолковый. Б. старик. < Безмозгло, нареч. Безмозглость, -и; ж.

БЕССЕРДЕЧНЫЙ, -ая, -ое; -чен, -чна, -чно.

Лишённый чуткости; бездушный. // Проникнутый бездушием. Б. поступок. Б-ое отношение к людям. < Бессердечие, -я; ср. Бессердечно, нареч. Б. гать. Бессердечность, -и; ж.<sup>9</sup>

Przymiotniki z prefiksem *без-/бес-* opierają się o wyrazy ze znaczeniem konkretnym (np.: *безавтомобильный, безвизовый*), abstrakcyjnym (np.: *бездефектный, бесстрессовый*), a także materiałowym (np.: *безвольфрамовый, бестабачный*)<sup>10</sup>.

*подвижный, шустрый, юркий, живой, шаловливый, егозистый, егозливый, вожены, егоза.* Na marginesie należy zaznaczyć, że znalezienie się rzeczownika *егоза* w powyższej grupie jest sporne. Ze względu na charakter pracy jedynie sygnalizujemy dany problem, który powinien być rozpatrzony z osobna. Niestety, przy opisie metody z zakresu leksykografii komputerowej nie możemy „dyskutować” z autorem słownika synonimów. W związku z tym, należy zaznaczyć, iż dane językowe oraz relacje między nimi należy pobierać z reprezentatywnych źródeł oraz poddawać wnikliwej weryfikacji według szczegółowego modelu teoretycznego.

<sup>8</sup> Wśród przymiotników z prefiksem *бес-* badacz T. N. Bucewa przedstawia grupy różniące się pod względem semantyki – zob. Т. Н. Буцева, *О принципах разграничения префиксального и префиксально-суффиксального способов образования прилагательных*, [w:] *Новые слова и словари новых слов*, ред. Н. З. Котелова, Ленинград 1983, s. 123.

<sup>9</sup> *Большой толковый словарь русского языка*, С. А. Кузнецов (red.), Санкт-Petersburg 1998.

<sup>10</sup> Zob. szczegółowe informacje na temat opisywanej grupy przymiotników oraz ich semantyki: В. Nowicka, *Префиксальные прилагательные в новой русской лексике*, Poznań 2007, s. 17–22. Termin *znaczenie materiałowe* (ew. *substancjalne*) w w/w pozycji nosi nazwę „вещественное значение”.

Podczas pracy użyliśmy przede wszystkim dwóch programów: edytora tekstu o nazwie *EmEditor*<sup>11</sup>, obsługującego składnię wyrażeń regularnych, oraz programu *Gephi 0.8.2-beta*, służącego do graficznej wizualizacji grafów<sup>12</sup>. *EmEditor* posłużył nam do skonfigurowania materiału badawczego, aby uzyskać formę danych obsługiwaną przez format baz danych programu *Gephi*. Dopiero w ten sposób uzyskaliśmy produkt w postaci grafu przedstawiającego wybrane szeregi synonimiczne. Materiał, którym posłużymy się w eksperymencie pochodzi z wersji elektronicznej<sup>13</sup> słownika pt. *Словарь синонимов и сходных по смыслу выражений*<sup>14</sup> (format \*.txt).

W celu uzyskania wizualizacji w postaci grafów przymiotników wraz z ich synonimami konieczne jest stworzenie odpowiednio skonstruowanego pliku zawierającego szeregi synonimiczne.

Do utworzenia listy interesującego nas słownictwa podstawową czynnością jest automatyczna ekscerpca<sup>15</sup> przymiotnikowych wyrazów hasłowych z prefiksem *без-/без-* wraz z ich treścią hasłową. W tym celu należy użyć filtrów wyrażeń regularnych<sup>16</sup> umożliwiających ekscerpcję interesującej nas grupy wyrazów. Na przykład z rekordów zawierających przymiotniki (por. niżej – 2 rekord) oraz przysłówki (1 rekord) należy wyekscerpować tylko przymiotniki (2 rekord):

**1 rekord:** Безвозмездно, безмездно, бесплатно, бескорыстно, даром. Ср. <Даром>. См. бесплатно, даром

**2 rekord:** Безрассудный, неразумный, бестолковый, безголовый, безмозглый, недогадливый, несмышленный, сумасбродный, безумный. Безрассудно: сдуру, сглупа, по глупости, без толку. Прот. <Рассудительный>. Ср. <Глупый>. См. глупый

<sup>11</sup> Por. stronę: <http://www.emeditor.com> (16.12.2013).

<sup>12</sup> Por. stronę: <https://gephi.org> (16.12.2013).

<sup>13</sup> Źródło e-wersji słownika: <http://speakrus.ru/dict> (16.12.2013).

<sup>14</sup> Н. Абрамов, *Словарь синонимов и сходных по смыслу выражений*, Москва 1999. E-wersja słownika ma rozmiar 450 KB.

<sup>15</sup> Por. np. P. Wierzchoń, *Automatyczne metody ekscerpacji neologizmów, czyli słowotwórstwo faktograficzne*, „Scripta Neofilologia Posnaniensia” 2005, nr VII; Y. Fedorushkov, *Methods for electronic excerption of new words in Russian*, [w:] PALC 2007: *Practical Applications in Language and Computers. Papers from the International Conference at the University of Łódź, 19–22 April 2007*, pod red. B. Lewandowskiej-Tomaszczyk, Frankfurt am Main 2007.

<sup>16</sup> Por. szczegółowe informacje na temat wyrażeń regularnych: <http://www.regular-expressions.info>; J. Friedl, *Mastering Regular Expressions*, Sebastopol 2006. Wykorzystanie RE-składni w badaniach korpusowych – por. Y. Fedorushkov, *Ekscerpca wyrazów cudzojęzycznych z tekstów prasy rosyjskiej*, „Semiosis Lexicographica” 2009, nr LII.

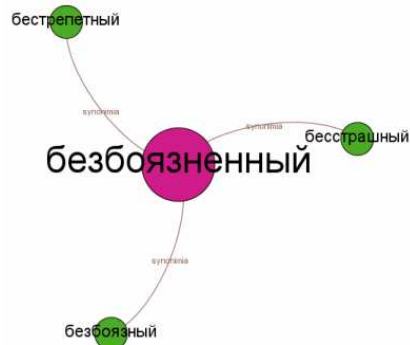
Do tego posłuży filtr o następującej RE-składni:

**^без.\*ый#.\*\$|^бес.\*ый#.\*\$<sup>17</sup>**

„Ostateczna”, tj. gotowa do importu lista szeregów synonimicznych zawiera następujące rekordy:

- 1 rekord:** безбоязненный;безбоязный;бесстрашный;бестрепетный; неустрашимый [...];
- 2 rekord:** безвозвратный;безнадежный;бесповоротный;невозвратимый; невозвратный [...];
- 3 rekord:** безвозмездный;безмездный;бесплатный;даровой;бескорыстный [...].

Struktura tekstu w rekordzie (np. brak spacji oraz średniki służące jako separatory) jest dostosowana do prawidłowego importu pliku do bazy danych w Gephi. Baza danych w Gephi składa się z dwóch Arkuszy: 1) z arkusza *Nodes*, tj. bazy wierzchołków: dominant i obiektów-satelit (synonimów) oraz 2) z arkusza *Edges*, tj. bazy krawędzi – linii pomiędzy wierzchołkami – por. ryc. 4. W niniejszej metodzie linie te wyrażają relację synonimii, np. dla czterech wierzchołków (*безбоязненный*, *безбоязный*, *бесстрашный*, *бестрепетный*) uzyskujemy trzy krawędzie:



**Рис. 3.** Cztery wierzchołki, trzy krawędzie

<sup>17</sup> Dany RE-filtr zadaje wyszukiwarce edytora następujące zapytanie: „potrzebny jest taki rekord, który na początku (tj. od stanu zerowego) zawierałby ciąg znaków *б, е, з* lub *б, е, с* i kończący się ciągiem znaków *ы, ъ* z całą dalszą treścią należącą do danego rekordu”. Przymiotniki z morfemem fleksyjnym *-ий* nie są brane pod uwagę ze względu na ich liczebność w słowniku Abramowa oraz brak pokąźnego szeregu synonimicznego dla wyrazu wyjściowego. Słownik zawiera tylko dwa artykuły hasłowe z następującymi wyrazami wyjściowymi:

*безногий* см. калека

*безрукий* см. калека

Odpowiednia struktura rekordów (por. wyżej) umożliwia generowanie<sup>18</sup> sieci krawędzi (w arkuszu *Edges*) w sposób automatyczny:

Source	Target	Type	ID	Label	Weight
безбоязненный	безбоязный	Directed	1	synonimia	1.0
безбоязненный	бесстрашный	Directed	2	synonimia	1.0
безбоязненный	бестрепетный	Directed	3	synonimia	1.0

Gotową listę należy zapisać w formacie \*.csv<sup>19</sup>, gdyż w takim formacie możliwy jest eksport danych do programu Gephi<sup>20</sup> (por. niżej – ryc. 4).

<sup>18</sup> Możliwe są jednak inne, bardziej czasochłonne techniki formatowania listy wejściowej zawierające dane językowe. Przy różnorodnych typach relacji leksykalnych ekscerptor może wygenerować bazę danych dla krawędzi (arkusz *Edges*) samodzielnie. Jeśli w arkuszu *Nodes* (wierzchołki) wyraz *безбоязненный* posiada numer identyfikatora ID=1, a jego synonimy mają ID=2 (*безбоязный*), ID=3 (*бесстрашный*) etc., np.:

ID (dla wierzchołków)	Label (zawartość wierzchołka, tzw. etykieta)
1	безбоязненный
2	безбоязный
3	бесстрашный
4	бестрепетный

to listę krawędzi (arkusz *Edges*) można wygenerować za pomocą edytora tekstowego w sposób następujący:

Source (numer ID dla wierzchołka <i>безбоязненный</i> )	Target (numery ID dla wierzchołka <i>безбоязный, бесстрашный, бестрепетный</i> )	ID (dla krawędzi)
1	2	1
1	3	2
1	4	3

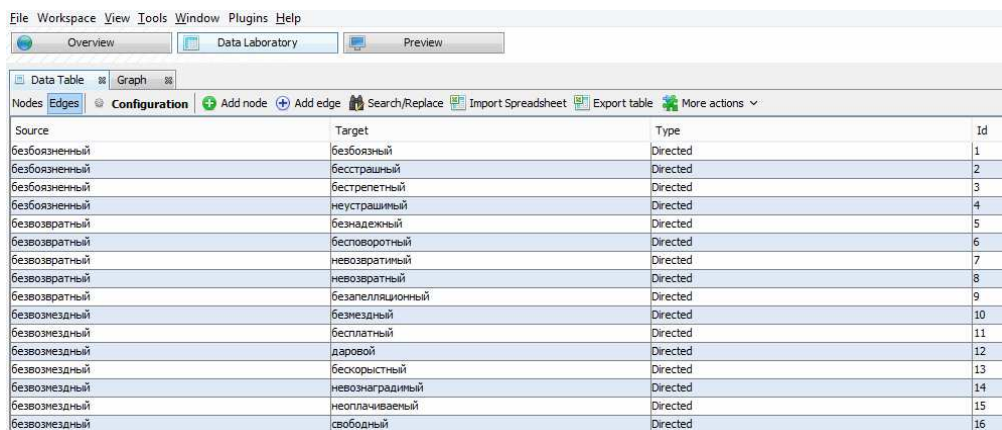
Warto w tym miejscu zaznaczyć, że każda relacja pomiędzy wierzchołkami posiada osobne ID. Innymi słowy, ID dla wierzchołków oraz ID dla krawędzi są funkcjonalnie różne: każda relacja pomiędzy ID wierzchołków (baza *Nodes*), np.: 1↔2, 1↔3, 1↔4 ma indywidualny ID: tj. identyfikator dla relacji pomiędzy wierzchołkiem 1 oraz wierzchołkiem 2 (1↔2) ma numer 1, ID dla 1↔3=2, ID dla 1↔4=3 (por. powyższe tabele) itd. Identyfikatory dla wierzchołków (*Nodes*) oraz dla krawędzi (*Edges*) są bytami osobnymi.

<sup>19</sup> Format CSV jest obsługiwany m. in. przez MS Office Excel.

<sup>20</sup> Prawidłowe zastosowanie filtrów sprawia, że program Gephi automatycznie generuje listę przymiotników i wyrazów tworzących wraz z nimi szeregi synonimiczne. Lista zawiera też pozostałe informacje, zbędne przy tworzeniu przejrzystej wizualizacji grafowej. Dlatego należy usunąć znaki interpunkcyjne, przykłady użycia, kwalifikatory itd., tak, aby z całej zawartości treści hasłowej pozostały wyłącznie synonimy wyrazu hasłowego. Usunięcie kwalifikatorów (np. *Прот.* – „противоположное” – jako antonim) oraz odsyłaczy (np. *См.*, *Ср.* – „смотри”, „сравни” – tj., „zobacz” inne hasło lub „porównaj” z innym hasłem wyjściowym



Środowisko Gephi składa się z trzech modułów: 1). *Data Laboratory* (moduł do wprowadzenia – importu – bazy danych w formacie CSV); 2). *Overview* (moduł do opracowania graficznego grafu za pomocą odpowiednich filtrów ustalających proporcje pomiędzy obiektami oraz relacjami) - por. ryc. 5; 3). *Preview* (moduł przeglądu grafu, służący także dla ostatecznych optyczno-wizualizacyjnych poprawek) – por. ryc. 6. Instrukcja wprowadzenia danych znajduje się na oficjalnej stronie środowiska Gephi w prezentacji pt. „Gephi Quick Start Tutorial”<sup>21</sup>. Dana prezentacja ukazuje szereg czynności dotyczących graficznej obróbki grafu w modułach *Overview* oraz *Preview*. W związku z tym nie opisujemy etapu graficznej obróbki z należytą dokładnością.



The screenshot shows the 'Data Laboratory' module in Gephi. It displays a table with columns for 'Source', 'Target', 'Type', and 'Id'. The table contains 16 rows of data, each representing a directed edge between two nodes. The nodes are labeled with Russian adjectives, and the edges are all of type 'Directed'.

Source	Target	Type	Id
безбоязненный	безбоязний	Directed	1
безбоязненный	бесстрашный	Directed	2
безбоязненный	бестрепетный	Directed	3
безбоязненный	неустрашаемый	Directed	4
безвозвратный	безнадежный	Directed	5
безвозвратный	бесповоротный	Directed	6
безвозвратный	невозвратный	Directed	7
безвозвратный	невозвратный	Directed	8
безвозвратный	безапелляционный	Directed	9
безвозмездный	бездельный	Directed	10
безвозмездный	бесплатный	Directed	11
безвозмездный	даровой	Directed	12
безвозмездный	бескорыстный	Directed	13
безвозмездный	невознаградный	Directed	14
безвозмездный	неоплачиваемый	Directed	15
безвозмездный	свободный	Directed	16

Ryc. 4. Moduł *Data Laboratory*: importowane wierzchołki oraz krawędzie

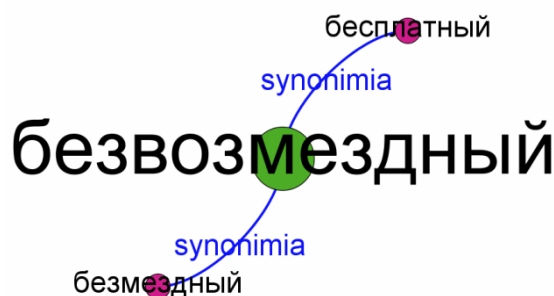
Pierwszy wyraz z rekordów będzie dominantą w grafie. Przykładowo, dla rekordów 1–3 otrzymujemy trzy dominanty (wyrazy inicjujące szeregi synonimiczne: *безбоязненный*, *безвозвратный*, *безвозмездный*). Dominanta *безвозмездный*, posiadając największą liczbę synonimów, automatycznie staje się najbardziej uwypuklonym obiektem grafu.

dla szeregu synonimicznego), tj. metadanych niezbędnych do nawigacji w klasycznym słowniku jest zasadniczo minusem omawianej metody. Wiąże się to z trudnościami technicznymi, które dotyczą przejrzystości matrycy danych oraz szeregiem licznych dodatkowych operacji manualno-ekscerpcyjnych, ograniczających automatyczność poszczególnych czynności w omawianej metodzie. Metody pozwalające na zostawienie wspomnianych metadanych oraz opracowanie systemu odsyłaczy jako potencjalnych hiperłączy w środowisku grafowym zamierzamy opisać w kolejnych pracach.

<sup>21</sup> Prezentacja dla osób rozpoczynających pracę z Gephi, [w:] źródło elektroniczne: <https://gephi.org/2010/quick-start-tutorial/> (12.10.2013).

Рис. 5. Umieszczenie gotowej listy w modułach *Overview*Рис. 6. Umieszczenie gotowej listy w modułach *Preview*

Powyższe ryciny przedstawiają *de facto* trzy grafy – po jednym dla każdego z rekordów (1–3). Niemniej jednak Gephi uwzględnia proporcjonalność pomiędzy obiektami (wierzchołkami) oraz relacjami (krawędziami) w każdym z grafów. Grafy te względem siebie ostatecznie „zachowują się” jak jeden „megagraf”<sup>22</sup>. Zależność pomiędzy poszczególnymi grafami uwidacznia się także w przypadku, gdy któryś z obiektów zawiera się w kilku rekordach jednocześnie. W takiej sytuacji Gephi nie powtarza dwa razy tego samego wyrazu, tj. „nie mnoży bytów”, a kwalifikuje taki obiekt jako element wspólny dla kilku grafów (por. też niżej – ryc. 9: obiekt *слабый* dla grafów z dominantami *бессильный* oraz *бесхарактерный*). Dane do modułu *Data Laboratory* można także wprowadzać ręcznie: obiekty-wierzchołki do arkusza *Nodes* oraz relacje-krawędzie – do arkusza *Edges*. Jak już wspomniano, poszczególnym wierzchołkom w *Nodes* zostaje „przypisany” ID (numer identyfikacyjny) w sposób automatyczny. Przykładowo, jeżeli wierzchołek o zawartości „безвозмездный” ma ID=1, to pierwszy synonim w szeregu synonimicznym, np. *бесплатный*, będzie miał ID=2, drugi – *бесплатный* – ID=3 itd. Relacja synonimii w tabeli *Edges* będzie zatem reprezentowana poprzez graf o krawędziach skierowanych: 1→2, oraz 1→3 itp., np.:



Ryc. 7. Przykład grafu skierowanego w module *Preview*

Podsumujmy: po wykonaniu powyższych czynności otrzymujemy bazę danych zawierającą listę czterdziestu szeregów synonimicznych. Gotową listę należy zapisać w formacie pliku CSV. Jak już wspomniano wyżej, uzyskaną w ten sposób listę można importować do programu Gephi i na jej bazie utworzyć sieć relacji synonimii.

Program Gephi automatycznie generuje grafy na podstawie zaimportowanych danych. Ważne jest, aby w importowanej bazie danych wyraz-dominanta (wyraz inicjujący szereg synonimiczny) znajdował się

<sup>22</sup> W innej terminologii tzw. *multigraph* (por. źródło elektroniczne: <http://www.edmondchow.com/pubs/SS105ChowE.pdf> [22.12.2013]).

na początku każdego wiersza, ponieważ posiadając takie położenie, w grafie będzie on figurował jako wierzchołek centralny. Korzystając z narzędzi programu, definiujemy rozmieszczenie grafów, ich kolorystykę, opcje napisów, wprowadzamy dodatkowe modyfikacje estetyczne itd. Po takiej edycji otrzymujemy przejrzystą sieć grafów.

Za pomocą narzędzi programu jesteśmy w stanie wizualnie wyodrębnić najbardziej produktywne szeregi synonimiczne (por. ryc. 8.). W ten sposób można zilustrować grupy wyrazów bliskoznacznych posiadające najwięcej komponentów. Ponadto program oferuje możliwość wykorzystania filtrów, pozwalających wyodrębnić konkretne grafy z interesującą nas ilością wierzchołków. Obsługa dodatkowych modułów, np. *Force Atlas*, *Yifan Hu Proportional*, jest intuicyjna. Moduły oraz wybrane filtry pozwalają uzyskać niezbędny efekt: kolorystykę, proporcjonalność, nienakładalność obiektów na siebie, grawitację<sup>23</sup> (filtr *Gravity*) obiektów-wierzchołków wokół dominanty i in.

Kolejną zaletą programu jest fakt, iż domyślnie łączy on ze sobą krawędzie grafów posiadających wspólne wierzchołki (por. ryc. 9.). Oznacza to, iż jeśli określony leksem jest częścią składową więcej niż jednego szeregu synonimicznego, owa zależność zostanie automatycznie przedstawiona przez program w tworzonej sieci grafowej. Taka przejrzysta wizualizacja sprzyja lepszemu postrzeganiu relacji zachodzących między analizowanymi leksemami i zaoszczędza czas potrzebny na poszukiwanie elementów wspólnych dla kilku szeregów synonimicznych. Wizualizacja materiału leksykalnego pozwala na płynną „podróż” po mapie leksemów, co nie byłoby możliwe przy użyciu tradycyjnych, manualnie tworzonych metod wizualizacyjnych. Na rycinie 8 (por. niżej) przedstawiona została kompletna (tj. zawierająca 40 szeregów synonimicznych) sieć relacji synonimicznych analizowanej grupy wyrazów. Krótko podsumowując, należy podkreślić, że dzięki ustawieniu wielkości wierzchołków (dominant oraz obiektów-synonimów wokół nich) proporcjonalnie do ilości tworzonych przez nie krawędzi (relacji synonimii) jesteśmy w stanie w krótkim czasie wyodrębnić najbardziej produktywne szeregi synonimiczne.

---

<sup>23</sup> Opis danego filtra jest następujący: “Attracts nodes to the center. Prevents islands from drifting away” (środowisko *Gephi 0.8.2-beta*).

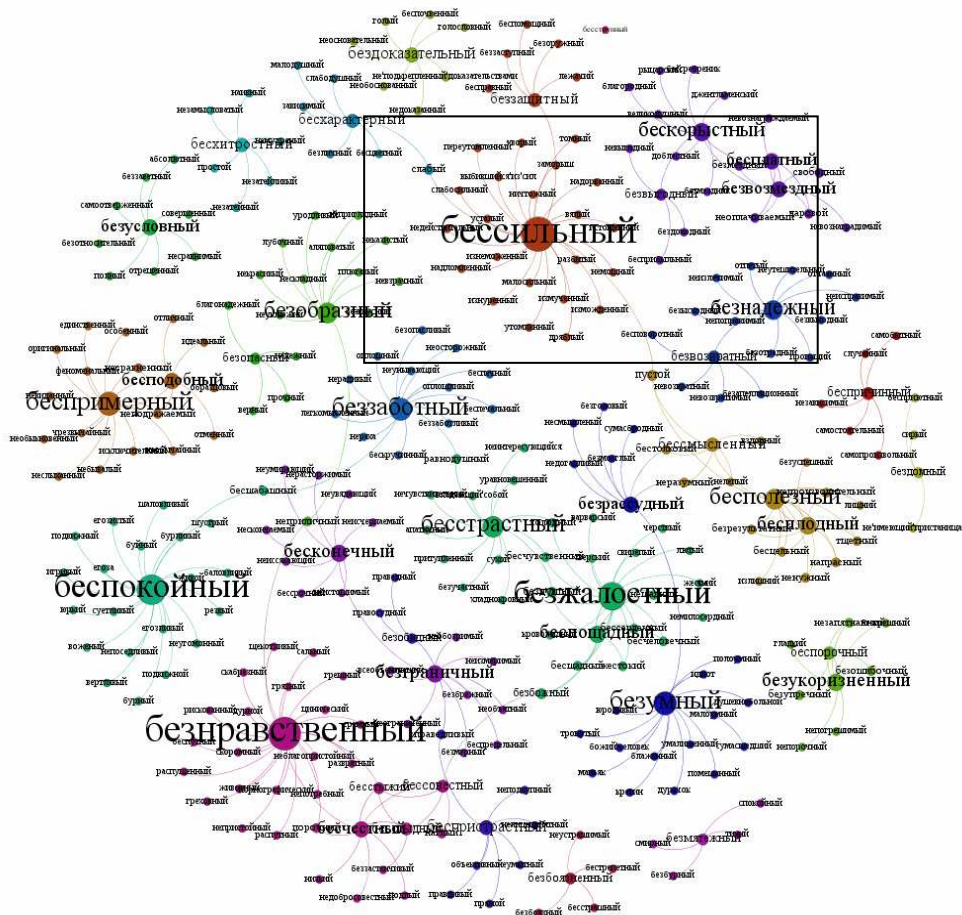


Рис. 8. Sieć relacji synonimicznych przedstawionych w postaci grafów.  
Obszar zaznaczony prostokątem – to rys. 9

Na rycinach 8, 9 widoczne są dominanty połączone relacją synonimii oraz obiekty-synonimy należące do kilku dominant, tj. będące synonimami jednocześnie dla kilku leksemów rozpoczynających szeregi synonimiczne. Należy w tym miejscu podkreślić, że moduł *Overview* wykonuje operacje łączenia wierzchołków automatycznie, bez konieczności ingerencji użytkownika.

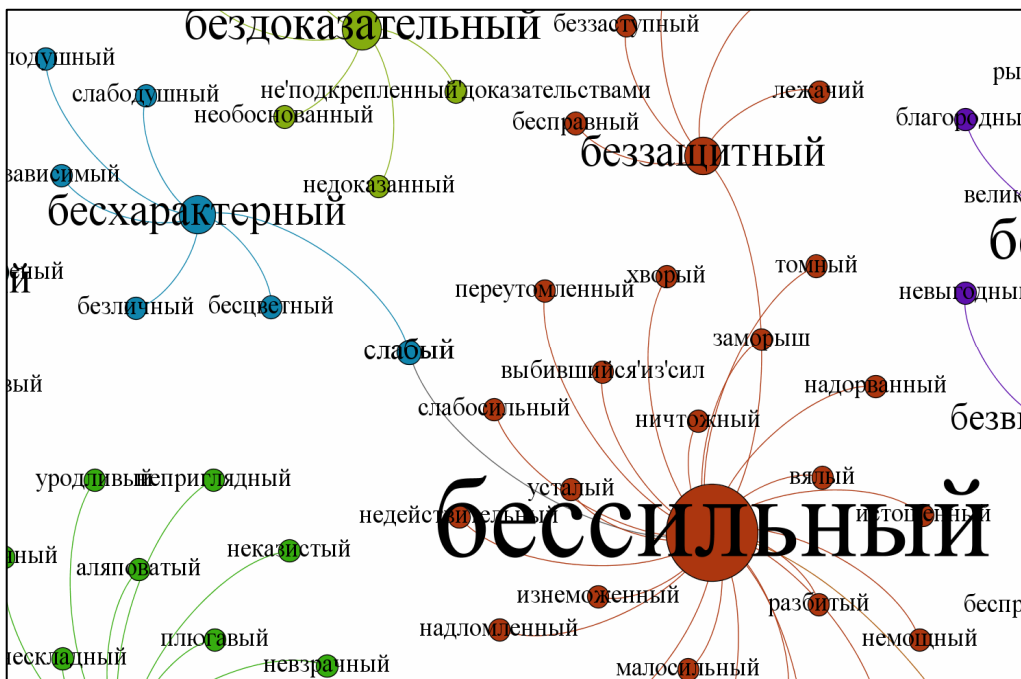


Рис. 9. Графы połączone wspólnymi wierzchołkami. Fragment wizualizacji z rys. 8. Obiekt *слабый*

W wyniku przeprowadzonej analizy uzyskujemy listę produktywności poszczególnych szeregów synonimicznych. Poniżej przedstawiona została lista wyrazów hasłowych uszeregowanych pod względem tworzonej przez nie liczby relacji. Obok każdej dominanty podana została liczba jej synonimów (tabela niżej):

бессильный	25	бесплатный	8
безнравственный	24	бесподобный	8
беспокойный	21	бессовестный	8
безжалостный	19	бесчестный	8
беспримерный	15	бездокладательный	7
безумный	15	беспристрастный	7
беззаботный	14	беззащитный	6
бесстрастный	13	бессмысленный	6
безобразный	12	бесхарактерный	6
бесполезный	12	бесхитростный	6
безнадёжный	11	безвозвратный	5
бескорыстный	10	беспорочный	5
безграничный	9	беспричинный	5

безукоризненный	9	бесстыдный	5
бесконечный	9	безбоязненный	4
бесплодный	9	безвыгодный	4
беспощадный	9	безмятежный	4
безвозмездный	8	безопасный	4
безрассудный	8	бездомный	3
безусловный	8	безобидный	3

Ogółem wśród analizowanej grupy leksemów znajdują się 93 wyrazy z przedrostkiem *без-*, lecz tylko 42 z nich zostały umieszczone przez autorów słownika w pozycji wyrazów hasłowych. Jednakże w skład analizowanego materiału weszło 40 szeregów synonimicznych, ponieważ dwa wyrazy hasłowe (*безапелляционный*, *бесструнный*) nie posiadają synonimów w słowniku. Dużą produktywnością odznaczają się synonimy z prefiksem *на-* (np. dla wyrazu *безрассудный* są to: *неразумный*, *недогадливый*, *несмышленый*), bowiem w całym materiale znaleźliśmy aż 73 przykłady owych leksemów. W przyszłości przewidujemy analizę zależności pomiędzy semantyką oraz modelem słowotwórczym leksemów.

Automatyzacja ekscerpcji, jakkolwiek posiadająca liczne zalety (zwłaszcza w kwestii szybkości edycji dużych ilości materiału leksykalnego), charakteryzuje się również pewnymi brakami, których należy być świadomym podczas pracy z wykorzystaniem omawianej metody<sup>24</sup>. Elektroniczna wersja słownika posiada wady dotyczące niekonsekwencji w użyciu znaków interpunkcyjnych, których usunięcie często może

<sup>24</sup> Jednym z mankamentów pracy ekscerptora jest problem umieszczenia w grafie wielowyrazowców, tj. obiektów składających się z kilku leksemów. Matryca bazy danych w środowisku Gephi przyjmuje tylko tak zwane słowa – ciągi znaków oddzielone separatorami, np. spacjami. Po przeprowadzeniu ekscerpcji wśród otrzymanych jednostek-synonimów należących do poszczególnych szeregów synonimicznych znajdujemy m.in. struktury wielowyrazowe, np.

не подкрепленный доказательствами;  
не имеющий пристанища;  
божий человек;  
выбившийся из сил;  
владеющий собой.

Aby w procesie wizualizacji każda z powyższych struktur została przedstawiona jako niepodzielna całość (tzn. jako jeden wierzchołek), należy wszystkie spacje oddzielające poszczególne wyrazy zamienić na inny znak, np. apostrof ('). W końcowym rezultacie otrzymują one poniższą postać:

не' подкрепленный' доказательствами;  
не' имеющий' пристанища;  
божий' человек;  
выбившийся' из' сил;  
владеющий' собой.

wiązać się z eliminacją części interesującego nas materiału. Niedogodności te, które ze swej natury nie mogą być w pełni rozwiązane przez żadną z metod zautomatyzowanych, zmuszają nas każdorazowo do przeprowadzenia manualnej weryfikacji materiału, będącej niezastąpionym narzędziem w pracy ekscerptora.

Pomimo wspomnianych wyżej mankamentów, uniemożliwiających operacje na większej ilości danych, przedstawiona metoda może z powodzeniem zostać wykorzystana przy wizualizacji małych partii materiału leksykalnego. Co więcej, ukazywanie relacji leksykalnych z wykorzystaniem tradycyjnych pomocy dydaktycznych może zostać zastąpione przez wygenerowane (w sposób zautomatyzowany) grafy, których tworzenie jest czynnością szybką i wydajną, a ponadto w znaczący sposób mogą one urozmaicić, uatrakcyjnić oraz wzmocnić proces dydaktyczny.

Przedstawiona metoda znacznie skraca czas formułowania ogólnych wniosków językoznawczych na podstawie optycznie uwypuklonych gradacji, hierarchii, relacji pomiędzy poszczególnymi obiektami grafu i jego obszarami: ściśle dane statystyczne nie są wówczas absolutnie konieczne. Ponadto, architektura środowiska Gephi jest dostosowana do umieszczenia poszczególnych projektów grafowych na serwerze. Serwer platformy środowiska *Neo4j* dodatkowo pozwala na zdalną edycję projektu grafu stworzonego w Gephi. W konsekwencji taki stan rzeczy daje możliwość samodzielnego wykonania przez ucznia zadań związanych z wizualizacją poszczególnych map myślowych bądź map dotyczących relacji systemowych w leksyce. Innymi słowy, powszechna dostępność narzędzi umożliwiających wizualizację danych w Internecie może zostać wykorzystana jako pomoc dydaktyczna w nauce języków obcych dla uczniów i studentów w ramach zdalnego nauczania (e-learningu).

Mamy nadzieję, że przedstawione techniki mają szansę zapoczątkować nowe badania w wielu dziedzinach nauki o języku. Zgodnie z koncepcją tzw. *definicji otwartej* Ryszarda Tokarskiego<sup>25</sup>, istnieje możliwość zastosowania opisaney metody także w leksykografii, bowiem uzyskiwana forma wizualizacji grafowej może posłużyć jako wyjściowy model do stworzenia oryginalnego słownika albo mapy konceptu językowego. Kolejną dziedziną, w której wizualizacja mogłaby okazać się pomocna, jest translatoryka. Jednak centralnym obszarem badań, w którym opisana metoda wizualizacji danych językowych mogłaby się rozwijać, jest, naszym zdaniem, semantyka.

<sup>25</sup> R. Tokarski, *Światy za słowami. Wykłady z semantyki leksykalnej*, Lublin 2013; J. Bartmiński, R. Tokarski, *Definicja semantyczna: czego i dla kogo?*, [w:] *O definicjach i definiowaniu*, pod red. J. Bartmińskiego, R. Tokarskiego, Lublin 1993; R. Tokarski, *Semantyka barw we współczesnej polszczyźnie*, Lublin 2004.