

NEURAL NETWORK PATTERN FOR ENHANCING FUNCTIONALITY OF ELECTRONIC DICTIONARIES

Serhii Fokin

Taras Shevchenko National University of Kyiv, Ukraine

sergiyborysovykh@ukr.net

The value of a dictionary is traditionally considered to be proportional to its physical volume, measured in the number of entries. However, the amount of useful data varies depending on existing hypertextual links across a dictionary. Therefore, its utility might also be calculated as proportional to the number of useful links among its structural parts which can interact in a similar way as neurons do via synapse links, provided that the number of links turns out to be exponentially greater than the number of entries. Today's lexicographic practice, as well as an experiment held by the author with his own developed onomasiological electronic dictionary of phraseological synonyms "IdeoPhrase", appears to demonstrate that the main criterion for establishing links automatically is the repetition of each kind of signs (stylistic labels, graphical word, metalinguistic comments). Automatically generated hypertextual links can be used for finding out semantic relations of different types among lexemes (synonymic, antonymic and others), semantic equivalence or similarity among lexemes in different languages (which is close to automatic translation), as well as compiling a new dictionary. The fact that generated relation established by a computer constitute new useful knowledge which has not been directly input by the compiler, qualifies this algorithm as artificial intelligence engine.

Keywords: dictionary; onomasiology; hypertextuality; computational lexicography; translation; semantics; phraseology.

Introduction

The advantages of electronic dictionaries over traditional ones are well known: they provide the possibility of a real-time interactivity (Selegey, 2003; Schryver, 2004) along with a much wider search engine in combination with the Internet resources. Nevertheless, a key feature of electronic dictionaries is a higher potentiality for hypertextuality, i.e. a reference to another text (Shepherd, 1998), has not been explored widely enough. Another essential feature of electronic dictionaries, which is seldom mentioned, is generating automatically new dictionaries on the basis of those which are already available.

At present most electronic dictionaries contain useful information which constitutes a piece of knowledge processed and directly input by their compilers in the format: "question" – "answer" or, in terms of lexicography: "lemma" – "gloss", or, in terms of informatics: "key" – "value". In other words, a user is supposed to introduce a key for which the programme should search for the correspondent value as is shown in the example below (Ukrainian Linguistic Resources, 2013):

Key:

Star

Value:

[sta:]

1. n

1) зоря, зірка; світило

fixed stars – нерухомі зорі

shooting star астр. – падаюча зірка

2) друк. Зірочка

3) об'єкт, що нагадує зірку; зірочка (на лобі тварини)

4) доля

5) зірка, провідний актор (актриса); відома особа, видатна особистість

film star – кінозірка

stars and stripes – державний прапор США

I saw stars – у мене іскри посипалися з очей

2. adj

1) зоряний

2) видатний; прекрасний; провідний

star witness – головний свідок

3. v

1) прикрашати зірками

2) позначати зірочкою

3) театр. бути зіркою, грати головні ролі

4) надавати головну роль

This dictionary entry illustrates a typical set of keys and values (or lemmas and glosses) which are to be contained in a database. In this dictionary, the gloss is not certainly generated by the computer, but is directly input by the compiler. In fact, still a really narrow part of electronic dictionaries are capable of generating glosses by themselves “on the fly”. Let us make a brief review of some possibilities of this process treated theoretically.

On the one hand, there is a bunch of research on automatic extraction of equivalents from parallel corpora, process, that might be qualified as artificial intelligence involving an automatic search of equivalents (Taljad, Schryver, 2002; Manning, 2008; Fokin, 2017). However, this application is barely a top of the iceberg among the possibilities of generating knowledge that electronic dictionaries possess. We understand the term “knowledge” here as searching and relation of different pieces of information which may satisfy a potential query by the user. The searching results constitute intellectual novelty which has not been directly input to the database by the compiler. As a prototype of computer-generated knowledge, we can consider the so called “Lullian circles” which produce statements from randomly aligned premises (Gardner, 1958, p. 119).

Thus, the objective of the article is to explore a hypertextual feature of an electronic dictionary, namely repetition, as well as the ways of its application for generating new knowledge by applying neural network pattern to the macro- and microstructure of the dictionary.

Semasiological VS onomasiological dictionaries

In modern lexicography, there are two basic modes of structuring dictionaries: semasiological and the onomasiological. Each of them needs a specific approach for generating automatic knowledge. Semasiological dictionaries, that constitute the overwhelming majority of the dictionaries, are usually based on the alphabetic order of graphic words, therefore their useful volume is traditionally measured up by a linear sequence of alphabetically ordered lemmas or entries. The explanatory dictionary of Spanish DRAE is semasiological, and a developed system of special grammar, stylistic and other markers allows the user to generate subdictionaries on its basis.

By contrast, the onomasiological dictionaries (commonly comparable to thesauri), are not very popular due to their intuitive, non-linear organisation, based on ideas. The dictionary “Dirae”, an inverted Spanish language explanatory dictionary (DIRAE, n.d), is one of the simplest ways of organising an onomasiological dictionary in a digital form. Instead of finding a gloss for a lemma (or a value for a key), it searches for the keys corresponding to a certain value. For example, by searching the term “mueble” (“furniture”), the programme finds 379 lemmas which are potential hyponyms of the concept “furniture” (DIRAE, n.d.). The big dilemma is that there is not, if any, a universal alphabet nor ordering system for ideas. This gap is bridged by different so-called synoptic schemes developed in onomasiological dictionaries, starting with P.M. Roget’s Thesaurus (first published in 1852), passing by “Diccionario ideológico de la lengua española” by J. Casares (Casares, 1942), and ending with current editions like O.S. Baranov’s ideographic dictionary of the Russian Language (“Идеографический словарь русского языка”, 1995). Those dictionaries are based on schemes containing specific semantic metalanguage which is difficult for ordinary users to master and constitute dramatically unknown domain even for experienced philologist. That is the reason why it is just in this very field where artificial intelligence would be of great help. One of the first attempts to take advantages of computational technologies for compiling onomasiological dictionaries was AUTOFRAS (Pamies Bertrán, Iñesta Mena, Balmacz, Kaloustova, 1998), which provides the user with the possibility to find phrases related to a certain idea in 10 languages. In the whole, the dictionary contains 12.000 phraseological entries. Unfortunately, works like the above mentioned are rather rare findings in modern computational lexicography.

Google search console can be widely used as a primitive onomasiological dictionary where a user can find a term by describing it approximately in the query line. A developed onomasiological dictionary, however, is more complicated. In those schemes, the user should start selecting from a set of wider categories and proceed to several levels of subcategories/subconcepts. This task requires a serious intellectual effort along with intuitive skills, while with modern electronic search engines it becomes easier and faster, and modern databases of terminology make a good use of this feature. For instance, Multilingual Thesaurus of the European Union renders 8 fields of knowledge with 11 respective subcategories to which refers the word “plant”:

006 EU institutions and European civil service
EU office or agency
3606 natural and applied sciences
life sciences
5211 natural environment

natural resources
plant resources
physical environment
 harmful **plant**
wildlife
plant life
 5606 *agricultural policy*
agricultural policy
 5621 *cultivation of agricultural land*
cultivation techniques
 5626 *means of agricultural production*
means of agricultural production
 5631 *agricultural activity*
crop production (EuroVoc, n.d.).

This example illustrates how a repeatedly used sign within a dictionary, particularly, the graphic word “plant”, serves to create a network of relations among the indicated fields. However, this kind of relation is quite artificial, because the relations turn out to be rather formal, where “industrial plant” turns out to be related to a biological one. But the grid of fields, categories, subcategories and concepts do create a considerable number of non-formal, certainly meaningful links between terms. These repetitions go far beyond just the list of terms pertaining to a category, but also every imaginable multiple relations, including those terms which are related to two or more fields at the same time: biology and chemistry; biology and human rights, or other kinds of relations which make some sense. These results are possible due to a thorough metalinguistic description performed by experts. Such description is supposed to be human resource work involving a big intellectual effort, but at the same time compensated by immense possibilities of creating new knowledge automatically. As H.-J.C. Long-Sheng observes: “In practice, the semantic orientation indexes have poor classification performance. Although machine learning techniques have better classification abilities, they need additional learning time and the information of classes (class labels) which should be determined by domain experts have to be provided before training” (Long-Sheng, 2009).

Unlike traditional hypertextual links, repeated signs themselves do not explicitly address each other, but, we do maintain, they are often implicitly related not just formally, but at a functional level. By this, we do not affirm that every casual link between two or more randomly chosen words in a system (a dictionary, a literary work) should be useful, but the examples adduced above do clearly show that repeated signs are more likely to conceal a useful piece of potential knowledge for a user.

Experiment and methods

Generation of new links based on repeated signs could be used for different practical purposes. For instance, finding synonyms or pseudosynonyms might be easily formalised by comparing the number of identical or even similar wordforms within the glosses: similar values within the glosses may indicate that their respective keys (lemmas) are likely to be synonymous or even antonymic. Another possible practical usage of the dictionary hypertextual character, which is widely explored, is making the so-called frequency lists for a corpus. Without knowing elementary morphological rules, a machine, by finding identical combinations of letters, is capable of building up entire lexical families and/or semantic fields, as well as producing more interesting practical effects. So, the repetition of a sign is, beyond any doubt, a parameter which is not to be deprecated in electronic dictionaries, both, when developing the interface with its possible functionality, and when calculating its useful volume.

In order to make a good use of hypertextuality in a dictionary for finding a synonym, we decided to apply the above described property in order to create an electronic multilingual dictionary of phraseological synonyms.

While the original idea consisted in creating an onomasiological dictionary exceptionally for personal use, step by step, a simple programme has been transformed into a multilingual onomasiological dictionary “IdeoPhrase” (IdeoPhrase, n.d.). Its algorithm consists in finding repetitions, particularly, the repeated semes. As a programme does not understand semes, and they do not necessarily appear in definitions, we decided to explicitly indicate the categories related to each phrase covering their respective semes along with a set of equivalent phrases in different languages where it was possible. The current version covers phraseological units in Ukrainian, Russian, Spanish, English, Italian, French, Latin and Hebrew. Here is a snippet of the database (Table 1).

Table 1. Fragment from the database used in “IdeoPhrase”

Spanish	French	English	Ukrainian	semes
CUERNO DE ABUNDANCIA	CORNE D'ABONDANCE	CORNUCOPIA	РІГ ДОСТАТКУ	Richness, money, gold, dream, treasure, unreality
VELLOCINO DE ORO	TOISON D'OR	GOLDEN FLEECE	ЗОЛОТЕ РУХО	Fortune, treasure, unreality, dream, richness
TRABAJOS DE HÉRCULES	TRAVAUX D'HÉRACLÈS	LABOURS OF HERCULES	ПОДВИГИ ГЕРАКЛІА	Efforts, feat, result, achievement, braveness, miracle

The fifth column reflects a metalinguistic description of the entries. As we can see in the table, both “CORNUCAPIA” and “GOLDEN FLEECE” could be accessed from the descriptor “richness”. By observing this phenomenon we have proceeded to profit from existing descriptors to “teach” the machine to find out phraseological synonyms. The machine considers potential synonyms to phrases which have at least two descriptors in common and outputs them in orange colour; if a phrase has three or more descriptors in common with the original phrase, the result is output in green as a potential synonym of greater probability. The programme is provided with a “filter” which turns the search results more precise (though less abundant), that is to say, the criteria are set to three common descriptors, a threshold to consider potential synonyms output in orange and four descriptors for green colour. The results, in their turn, suggest possible corrections to the descriptors.

Results and Discussion

Each randomly selected phrase among 10 “IdeoPhrase” generates links with averagely 11 potential synonyms. We acknowledge that this number is not enough for statistic evaluation, though it may appear to prove its wide practical use. Supposing that a descriptor can be repeated up to 20 or 30 times across the dictionary, containing each entry on average 5 descriptors, introduction of a new entry generates at once up to 100 new useful links. The given numbers are too approximate, the exact calculation requires a separate research, but this example clearly shows that the value and its useful volume do not necessarily correlate with the number of its entries. In other words, the number of links will logarithmically grow along with the growth of general dictionary volume. Here is an example of the dictionary output which demonstrates a search of phraseological synonym for “TO BE IN A RUSH” in the Appendix 1.

The way of finding proper descriptors have been suggested and discussed in our previous work. In short, a descriptor is not only supposed to answer the question “What is the phrase about”, but also “When does it usually take place”, “What are the consequences”, “Who participates in that process”, though this process is not at all exempt from subjectivity (Fokin, 2015, p. 72). Despite the time-consuming process of typing in descriptors, the big advantage is that by introducing just one entry we gain up to 100 useful singular links, because the descriptors on average are repeated 10-25 times (on increasing the database, the number of repetitions will grow, too). Double and triple links also matter as they uncover potential strong semantic relations such as synonymy or antonymy.

The possibility of generating new knowledge is widely explored by the compilers of electronic dictionaries. For illustration let us use the explanatory dictionary of the Spanish Language “Diccionario de la Real Academia Española” (DRAE, n.d.). The major value of its CD version, in our judgement, is its capability of building up more specific dictionaries on its own basis. For instance, the user can make a query in order to compile the list of all transitive verbs contained in the dictionary, or all hellenisms, all vulgar lexemes, etc. The DRAE can compile a list of entries containing certain wordforms, which in fact results in a new dictionary. This process is simple due to its easy formalisation, which is based on a well distinctive criterion: repetition of a stylistic/grammar label or a certain lexeme throughout the dictionary.

It turns out that a linear volume of the electronic dictionary does not tell us much about the practical use of the immense bytes array it covers. As a matter of fact, a good organisation of that volume makes it possible to considerably augment the number of useful links within a dictionary. The analogy could be that of a famous trick from our childhood of passing through a sheet of paper: by cutting out the paper in a specific way in the form of a curved ribbon, we can augment considerably the perimeter, though the area remains unchangeable, and perfectly pass through it. But the most eloquent analogy is the human brain, which does not seem to be huge by its physical dimensions, though the longitude of all the taxons put consecutively one after another could overcome a distance of 100,000 miles which is enough for wrapping the Earth four times.

The volume of a brain or even the number of neurons does not matter to evaluate nor the brain's potentiality nor its level of development. What does matter is the number of neural connections. In a similar way, the potentiality of a dictionary could be much larger than a simple number of entries and be proportional to the number of hypertextual links. The easiest way of establishing hyperlinks is by finding equal signs and interpreting the meaning of these coincidences. As Aboitiz (1996) concludes in his article "Does Bigger Mean Better? Evolutionary Determinants of Brain Size and Structure" emphasising "the role of connectional modifications in increasing brain capacity and contrasts with current ideas of a unitary process of phylogenetic brain growth, where a larger brain size *per se* produces better processing capacity, regardless of the causal factor behind it" (p. 225).

The programme's output contains noticeably more useful information than the compiler could possibly have foreseen. For example, the phrases "to try ones best", "to bend over backwards"; in Spanish: "partirse los cuernos", "a Dios rogando y con el mazo dando"; in Italian: "aver fare trenta e fare trentuno", "cavarsela per il rotto della cuffia"; in Ukrainian: "дістати зірку з неба", "зі шкіри пнутися", "не покладати рук"; in Russian: "сдвинуть гору с места" are referred to as synonyms thought they are situated in different rows of the database.

The aforementioned algorithm unveils an important property which electronic dictionaries possess and the paper dictionaries do not: curiously, the distinction between "lemma/key", "gloss/value" in this dictionary blurs away, because a lemma can potentially become a "gloss" for another lemma, that is to say: depending on the format of user's query, each component might be either key, either lemma. Using a data set as a not-at-all-strict database makes it closer to a neural network behaviour. Let us observe several examples of that kind of reciprocal interchange between keys and values depending on the type of query (Table 2):

Table 2. Types of queries possible in "IdeoPhrase"

Query	Key	Value
Which are the ideas of the phrase "to be inventing the wheel"?	PHRASE: To be inventing the wheel	SEMES: Repetition, discovery...
Which are the phrases related to the idea of repetition?	SEME: Repetition	PHRASES: Ukr.: БЛІДА ПОДОБА Ukr.: БУТИ ВИЛИТОЮ КОПІЄЮ Ukr.: ВИЛИТА КОПІЯ Ukr.: ВІДКРИВАТИ АМЕРИКУ Ukr.: ВСІ ВУХА ПРОКВАСИТИ Ukr.: ЖУВАТИ КЛОЧЧЯ Ukr.: ЗА ЩО КУПИВ, ЗА ТЕ Й ПРОДАЮ Ukr.: ЗНОВ ЗА РИБУ ГРОШІ Ukr.: НАБИТИ ОСКОМУ Ukr.: НАВ'ЯЗНУТИ В ЗУБИ Ukr.: НАЧЕ КОБИЛА В СПАСІВКУ Ukr.: ОЧІ МУЛЯТИ Ukr.: ПОВЕРНЕМОСЬ ДО НАШИХ БАРАНІВ Ukr.: СЛОВО-ПАРАЗИТ Ukr.: ЧАСТО ГУСТО Eng.: A SPITTING IMAGE Eng.: BACK TO THE DRAWING-BOARD Eng.: BY THE SAME TOKEN Eng.: CAST IN THE SAME MOULD Eng.: COME FULL CIRCLE Eng.: DO A DOUBLE TAKE Eng.: JUNK WORD Eng.: TO BE INVENTING THE WHEEL Hebr.: אמריקה לגלות Hebr.: לאמת לחטוא It.: ACQUA PASSATA NON MACINA PIÙ It.: AVERE UN CHiodo FISSO It.: BATTERE SEMPRE LO STESSO TASTO It.: ESSERE TAGLIATO CON L'ACCETTA It.: ESSERE UNA BARBA Lat.: CÍRCULUS VITIOSUS

(to be continued)

**Table 2. Types of queries possible in “IdeoPhrase”
(continued)**

Query	Key	Value
Which phrases are similar to the phrase “all out of the blue”	PHRASE: “All out of the blue”	<p>Rus.: ЗА ЧТО КУПИЛ - ЗА ТО И ПРОДАЮ Rus.: ОТКРЫВАТЬ АМЕРИКУ Rus.: ПОРЧНЫЙ КРУГ Span.: CADA DOS POR TRES Span.: INVENTAR LA BICICLETA Span.: TENER UNA ESPINITA CLAVADA</p> <p>Ukr.: БУТИ НЕ В ДУСИ Ukr.: В ОДИН МЕНТ Ukr.: З НЕБА ВПАСТИ Ukr.: З ЯКОЇ РАДОСТІ? Ukr.: ЗАСКОЧИТИ ЗНЕНАЦЬКА Ukr.: І НА ТОБІ Ukr.: МОВ ГРИМ СЕРЕД ЯСНОГО НЕБА Ukr.: МОВ З ЗЕМЛІ РОДИТИСЬ Ukr.: МОВ КУЗЬМА З МАКУ Ukr.: НЕ ТИМ ВІТРОМ ПОВІЯЛО Ukr.: НІ З ТОГО НІ З СЬОГО Ukr.: ПРИШІЙ КОБИЛІ ХВИСТ Ukr.: ЧУДЕСА В РЕШЕТІ Ukr.: ШИЮ ЗІАМАТИ Ukr.: ЩОСЬ НАЙШЛО НА КОГОСЬ Ukr.: ЯК ПИЛИП З КОНОПЕЛЬ Ukr.: ЯК СНІГ НА ГОЛОВУ Ukr.: ЯКАСЬ МУХА ВКУСИЛА Ukr.: ЯКИЙСЬ ГЕДЗЬ ВКУСИВ Ukr.: ЯКИМ ВІТЕРЦЕМ ЗАНЕСЛО Eng.: CATCH SOMEONE ON THE HOP Eng.: CHANGE SOMEONE’S TUNE Eng.: COME OUT OF NOWHERE Eng.: בהיר ביום כרעם Hebr.: בהיר ביום כרעם Hebr.: גבול להשיג Hebr.: גבול להשיגת It.: ASPETTARE QUALCUNO AL VARCO It.: CAPITARE TRA CAPO E COLLO It.: ENTRARE PER IL BUCO DELLA CHIAVE Lat.: SINISTRO PEDE ASSURGERE PHRASES: Ukr.: НІ ДУМАНО НІ ГАДАНО Rus.: В ОДИН МОМЕНТ Rus.: КАК ЗДРАСЬТЕ Rus.: НИ С ТОГО НИ С СЕГО Span.: ¿A SANTO DE QUÉ? Span.: DE BUENAS A PRIMERAS Span.: PILLAR DE IMPROVISO</p>

Thus, both lemmas and descriptors (semes) can be used as a query in “IdeoPhrase”. Descriptors in onomasiological dictionaries are usually based on hierarchically organised synoptic schemes, sometimes tricky and difficult to manage: the concept of “fraud” is likely to appear under a wider concept, “crime”, but also it forms part of other categories, as “infracrime”, “deception”. That’s the reason why as a first step we have decisively opted for a one-level scheme in the first version of the programme. One-level scheme presupposes connection of a set of descriptors for each entry without the need of generalising them in hyperonyms and other types of global notions. The main inconvenience of this approach is that an extremely general descriptor, such as “contradiction” or “problem” would appear in many entries and, as a result, it would produce an exaggerated output of phrases related to the indicated descriptors. But it does not create inconveniences in small volume dictionaries, in this very case, it’s up to the user to easily reduce this amount by using a more specific noun.

As Pamies, Balmacz and Iñesta (1998) affirm, due to the problem of hyponymity-hyperonymity and synonymity between descriptors, the system would become unworkable in big volume data (p. 209). We assume that a system indeed fails when it requires more operation than the computer is able to perform in a reasonable period of time. However, in a dictionary with 100.000 entries, each containing up to 10

descriptors, an automatic revision of one million descriptors would be hardly a noticeable time for the user, as the time spent by the processor would be directly proportional to the number of entries and descriptors, in informatics, the so-called “big $O(n)$ ”. The real difficulty comes when the programme should deal with more complicated combinatorial tasks, for example, comparing entries with each other for creating a list of phraseological synonyms, either in the same language or in different languages. On each iteration, ten descriptors should be compared to those of another entry, which results in 100 operations. Logically, this cycle should be performed in combinations of “each-to-each” pair of entries, in our example, which would sum up 100.000×100.000 steps, that is, ten billion, corresponding to the order of the so called “big $O(n^2)$ ”. If we multiply ten billion steps by 100 required iterations, it would finally produce one trillion operations, an impressively huge number which, however, should take less than a second in modern computers. In any case, we do recognise that an optimisation could considerably reduce the electronic efforts, the problem which requires a separate case-study.

Conclusions

An electronic dictionary becomes a much more powerful tool if, apart from the traditional “key” – “value” interchange, it makes use of possible links between its part in order to generate new knowledge. This knowledge can satisfy a potential user’s query in the same way as in human communication: new knowledge is often based on statements that join well-known facts in a new relation.

As a first and very important step in finding out the useful links, an algorithm that selects repeated components (signs, abbreviations, words) should be implemented, since it is the most outstanding feature of their hypertextuality. In spite of the fact that the volume of current dictionaries is traditionally calculated in a linear way: the amount of lemmas or even the amount of wordforms, its usefulness should be measured by the amount of potentially valuable information which does not necessarily correlate with a simple linear amount. An introduction of several lemmas might generate by two or three orders more useful links around the dictionary. The repeated components are also useful for creating new dictionaries or databases (concordances, frequency lists, specialised dictionaries), finding synonyms and other semantically related lexemes. This property is particularly useful for onomasiological dictionaries due to their complicated organisation and, logically, difficult usage. Manually added semes in each entry should help with a better performance of the search. This idea has been implemented in the compilation of the onomasiological multilingual dictionary “IdeoPhrase” which provides the user with search of semantically related phrases, synonymical phrases as well as the translations of phrases. We should not disregard the possibility of automatic search of other semantic relations: hyponymity, hyperonymity, holonymity, cause/effect. Since the compiler need not introduce the translations and synonyms, automatically found by the programme, this programme performance might be considered artificial intelligence.

The perspectives of the research outlined above can be summarised in the following way: 1) testing and improving the set of descriptors used in electronic onomasiological dictionaries, including experiment, informants participation, database enhancement in order to make it more objective with the description of possible methodologies; 2) creating a hierarchy of descriptors in order to make the access to the metalinguistic descriptions more handy and manageable for the users; 3) calculating statistical parameters of the dictionary: the number of entries, descriptors, repetitions and different applicable relations among them in order to measure dictionary practical value for the users, which does not necessarily correlate with the number of entries; 4) finding out an inverse algorithm: to make machine learn to retrieve semes, descriptors on the basis of given relations between phrases.

References:

- Aboitiz, F. (1996). Does Bigger Mean Better? Evolutionary Determinants of Brain Size and Structure. *Brain, Behavior and Evolution*, 47(5), 225-245. <https://doi.org/10.1159/000113243>
- Baranov, O.S. (1995). *Ideograficheskiy slovar russkogo yazyka* [Ideographic Dictionary of Russian Language]. Moscow, Russian Federation: Izdatelstvo ETS.
- Casares, J. (1942). *Diccionario ideológico de la lengua española*. Barcelona, Spain: Ed. Gustavo Gili.
- DRAE (n.d.). Diccionario de la Real Academia Española. Retrieved May 30, 2018 from www.rae.es
- DIRAE (n.d.). Diccionario Inverso de la Real Academia Española. Retrieved May 30, 2018 from <https://dirae.es/palabras/?q=mueble>
- EuroVoc (n.d.), Multilingual Thesaurus of the European Union. Retrieved May 30, 2018 from: <http://eurovoc.europa.eu/drupal/>
- Fokin, S.B. (2015). Vybir deskriptoriv dlya onomasiolohichnoho slovnyka: pytannya metodologiyi [Descriptor selection for onomasiological dictionaries: methodological criteria]. In V.B. Bourbelo & M.M. Popovych (Eds.), *Pyata Vseukrainska naukova konferentsiya romanistiv. Strukturalna-semantychni i kohnityvno-dyskursyvnia paradyhmy suchasnoho romanskoho movoznavstva* (pp. 71-72). Odesa, Ukraine: KP OMD.
- Fokin, S.B. (2017). Optimisation de l'extraction automatique des équivalents des bases de données bilingues par un filtrage lié à la longueur des mots. In H. Kriuchkov & V. Bourbelo (Eds.), *Actes du 1er Colloque international francophone en Ukraine*

- 19-20 octobre 2017 "Langues, Sciences et Pratiques" (p. 44). Kyiv, Ukraine: Université Nationale Taras Chevtchenko de Kiev, Ambassade de France en Ukraine, Institut Français d'Ukraine, Agence Universitaire de la Francophonie.
- Gardner, M. (1958). *Logic machines and diagrams*. New-York-Toronto-London, USA-Canada-GB: McGRAW-HILLBOOK COMPANY, INC..
- IdeoPhrase (n.d.), Onomasiological Multilingual Dictionary of Phraseological Synonyms. Retrieved May 30, 2018 from <http://postup.zzz.com.ua/IdeoPhrase.html#>
- Long-Sheng, H.-J. C. (2009). Developing a Neural Network based Index for Sentiment Classification. In S.I. Ao, O. Castillo, F. Douglas, D.D. Feng, & J.A. Lee (Eds.), *Proceedings of the International MultiConference of Engineers and Computer Scientist* (pp. 744-749). Hong Kong: University of Hong Kong. Retrieved May 30, 2018 from: http://www.iaeng.org/publication/IMECS2009/IMECS2009_pp744-749.pdf
- Manning, Ch.D., Raghavan, P. & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge, Great Britain: Cambridge University Press. Retrieved May 30, 2018 from <https://nlp.stanford.edu/IR-book/>
- Pamies Bertrán, A., Iñesta Mena, E.M., Balmacz, M., & Kaloustova, O. (1998). *Multilingual Electronic Phraseological Dictionary "AUTOFRAS"*. Tempus Language Toolbox (CD-version).
- Pamies, A, Balmacz, M., & Iñesta, E.M. (1998). Criterios para una fraseología onomasiológica automatizada. In J.D. Luque Duran & A. Pamies Bertrán (Eds.), *Léxico y Fraseología* (pp. 207-217). Granada, Spain: Método Ediciones.
- Schryver UGent, G.-M. & Joffe, D. (2004). On How Electronic Dictionaries are Really Used. In G. Williams & S. Vessier (Eds.), *Proceedings of the 11th EURALEX International Congress* (pp. 187-196). Vannes: Université de Bretagne Sud. Retrieved May 30, 2018 from: <https://biblio.ugent.be/publication/299034/file/6778172>
- Selegey, V. (2003). *Elektronnye slovari i kompiuternaya leksikografiya* [On electronic dictionaries and computational lexicography]. Retrieved May 30, 2018 from: <https://studfiles.net/preview/1771482/page:28/>
- Shepherd, M. & Waters, C. (1998). The evolution of cybergenres. In P. Apers et al. (Eds.), *Proceedings of the 31st Hawaii Conference on System Science* (pp. 97-103). Los Alamitos: IEEE Press. Retrieved May 30, 2018 from: <https://www.cybermovia.com/cgi-bin/olenuapro.pl>
- Taljad, É. & Schryver, G.-M. (2002). Semi-automatic Term Extraction for the African Languages, with Special Reference to Northern Sotho. *Lexikos*, 12, 44-74. Retrieved May 30, 2018 from <https://doi.org/10.5788/12-0-760>
- Ukrainian Linguistic Resources (2013). English-Ukrainian Dictionary. Retrieved May 30, 2018 from: <https://www.cybermovia.com/cgi-bin/olenuapro.pl>

Received: June 04, 2018

Accepted: May 02, 2019

Appendix 1. Results of searching synonyms for the phraseological unit "To be in rush"

