

MODERN APPROACHES TO AUTOMATED IDENTIFICATION OF METAPHOR

Olena LEVCHENKO

*Lviv Polytechnic National University,
Department of Applied linguistics,
5, Mytropolyta Andriya Str., build. 4/131, Lviv, Ukraine,
tel.: (032) 258 21 38*

Nataliia ROMANYSHYN

*Lviv Polytechnic National University,
Department of Applied linguistics,
5, Mytropolyta Andriya Str., build. 4/131, Lviv, Ukraine,
tel.: (032) 258 21 38*

This article discusses the existing methods and techniques of metaphor identification in text corpora. It focuses on the advantages of combination of semantic and statistical approaches to metaphoric corpus annotations, the role of dictionary-based data in mechanisms of verification of language media figurative usage and automated identification of metaphors.

Keywords: metaphor, metaphor identification, text corpus, text corpus annotation.

Introduction. In pretechnological era metaphor was identified empirically on the ground of language competence, such identification being largely based on researcher's intuition. Lakoff and Johnson's (1980) conceptual metaphor theory [12] constituted the foundations of this practice, which has been further developed in modern linguistics. Nowadays the emergence of text corpora that accumulate huge data files raised the question of metaphor identification, especially automatic identification which requires the development of special algorithm, i.e. to explicate conscious and subconscious processes accompanying metaphor identification by experts or just make it easier for them to process data, reducing the factor of subjectivity. Thus, the authors of Amsterdam corpus of metaphors claim that "depending on the degree of semantic indivisibility and motivation MIPVU identifier may treat a phraseological unit as a single metaphoric expression or as a combination of metaphoric and non-metaphoric ones – depending on the way it recognizes these units" [19]. Notably, S. Glucksberg presents experimental data suggesting that the time we need to understand metaphor equals the time we need to understand literal expressions [8].

Traditionally metaphors are divided into live and dead. Lakoff and Johnson suggest to define metaphors according to the degree of their novelty as 1) trite (dead); 2) widely used, understandable to everybody (conventional); 3) original (novel) [12]. They define dead metaphor as a linguistic expression which is the product of live conceptual metaphor application in language. Then the conceptual projection ceased to exist and the current expression retained only the meaning of target domain [12]. According to Lakoff and Johnson the word *pedigree* is an example [12]. The scholars contend that the application of the term ‘dead metaphor’ is restricted [12].

For the research purposes nowadays there have been developed CorMet system [14], MetaNet repository [24]; Amsterdam Metaphor Corpus [19: 6], ATT-Meta [25: 8] and the corpus of verbs marked according to the source and target of metaphoric transference [22] are available.

Stages and strategies of metaphor identification. Thus, cormet is defined as corpus-based system used to discover metaphorical representations of concepts by determining “systematic variations in domain-specific selectional preferences, which are inferred from large, dynamically mined internet corpora” [14].

a. Stefanovitsch distinguishes specific techniques for extracting linguistic expressions representing metaphors from non-annotated corpora. A set of metaphors identification strategies includes: manual searching; sampling based on source domain vocabulary; on target domain vocabulary; extricating of sentences containing words from both the source domain and the target domain; sampling based on ‘metaphoric markers. The procedures also include corpus analysis annotated for semantic fields/domains [5]. However not all of the above listed strategies can be technically applied for the Ukrainian language considering the available resources and software.

The second type of strategies is represented by works of A. Deignan [6]. A. Deignan suggests the procedure of metaphor analysis from identification of possible list of language metaphors to processing conceptual metaphors [6]. This procedure is described as the way from bottom upwards [6]. The algorithm of metaphor analysis proceeds from the research of a small corpus or part of a large corpus manually. Then the determined metaphoric connections are verified on the basis of a great amount of empirical material [6]. Working with the British language corpus with the aim of metaphor selection A. Deignan suggests focusing on the vocabulary representing source domain of metaphorisation – ANIMALS that include nouns and verbs (pig, fox, wolf, monkey, rat, dog) [10]. A scholar individually determines specific vectors of metaphoric source domain research – either all elements of source domain of metaphorisation or its parts. If the analysis aims to determine the source domain of metaphorisation the following algorithm can be applied. It is required to define a large amount of representational monothematic texts related to the topic of target domain of metaphorisation. Then by means of key words it is possible to define the metaphoric source domains represented in the text. The application of this algorithm is effective when metaphoric target domain is represented by concepts ECONOMICS, SPORTS, POLITICS, etc., and less effective for concepts EMOTIONS, MENTAL ACTIVITY, PERCEPTION [10]. However not all approaches of metaphor identification elaborated for different languages can be applicable

to Ukrainian. For example, analyzing English metaphors A. Deignan concludes that a noun denoting animal transfers into another part of speech depending on whether it is used in direct or figurative meaning – metaphorically used words ‘hound’, ‘hare’, ‘weasel’, ‘squirrel’ become verbs [10]. Such processes are not characteristic of Ukrainian and, consequently, the morphologic marking of the text is ineffective.

The third type of strategies represented by the works of P. Koivisto-Alanko, H. Tissari [13] is aimed at identification of concepts that constitute the source domain of metaphorisation; analysis of the target domain concepts environment and identification of metaphoric projections [13]. In application of this approach the subjectivity is unavoidable and, according to Stefanovitsch, the described source-domain oriented approach is only partially effective, because preliminary assumptions about source and target domains are necessary [5].

The fourth type of strategies listed above is a combination of the following stages: selection of sentences (or suitable language units) that include source and target domain vocabulary, which can be applied for automatic annotation/extraction [9]. This procedure requires comprehensive source and target domain vocabulary lists compilation as well as corpora annotated for clause and/or sentence boundaries to ensure the validity of metaphoric expressions [5].

The next strategy based on ‘markers of metaphor discussed by Goatly [11] is an intriguing possibility for the automatic retrieval of metaphors based on a diversity of explicit linguistic means indicating the metaphor. They include a wide paradigm of figurative metalinguistic expressions, general metalanguage expressions indicating semantic variability and even orthographic devices like quotation marks [5].

Stefanowitsch argues that first, second and third strategies are adaptable to corpora annotated for semantic fields/domains [5]. This procedure includes obligatory comparison of words ascribed a source domain tag with the concordances to verify the results [19].

The procedure of metaphor identification MIPVU is based on comparison between basic and contextual meaning of a word. Basic meaning of a word is a meaning which a) has characteristic “materiality” – is the most objective, material of all existing meanings of the word: nominates perceptible physical objects, features or phenomena – everything that can be seen, felt, smelled, tasted, being thought over is mentally resented by a specific “image”; b) has characteristic “corporeality” – nominates objects, features or actions related to human body and bodily actions; c) has characteristic “concreteness” and “unambiguity” – is the most concrete and unambiguous of all word’s meanings [19].

G. Steen, a participant of MIPVU project, suggests the following empirical cycle of metaphor identification [20].

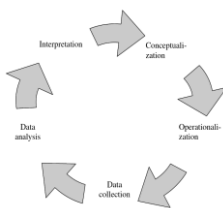


Figure 2
The empirical cycle

G. Steen believes that while identifying metaphor a researcher should define 1) metaphoric focus, 2) metaphoric idea 3) metaphoric comparison 4) metaphoric analogy for tenor interpretation 5) work out metaphoric projection [21].

L.V.Kulchytska arrives at a conclusion that all five stages of analysis is an integration of existing approaches to the nature of literalness as opposed to figurativeness in general and these approaches originated from traditional theory of metaphor. The first stage is the realization of M. Black's ideas about frame and focus, the second – is an analysis in terms of extensional semantic theories, the third – is related to the name of G. Miller application of formalized representation of metaphor, construed as an “open” comparison <...>, and the analysis of vehicle, the fourth – is the analysis of metaphorised notion – a tenor and introduction of analogies (in terms of ancient and post-ancient theories of metaphor), then fifth – is the procedure of metaphor's projections derivations in terms of G. Lakoff and M. Johnson [28].

However E. Shutova, L. Sun and A. Korhonen believes that automatic identification of metaphor may include two procedures: identification of metaphor and interpretation of metaphor [17].

S. Hahalova indicates that the difficulties with corpus data retrieval arise because linguistic forms do not represent semantic and conceptual identification. Despite the diversity of available types of semantic annotations (identification) the majority of researches of secondary nomination means cannot be based on annotated corpora [29].

There is an approach considered semantic in its essence. The method constitutes the application of semantic annotation to sample expectant metaphors from the source domain. It is based on the assumption that semantic tags ascribed to the words may refer to metaphoric source domains. The method implies the availability of dictionary based semantic annotation of the whole corpus [19].

Nowadays the process of metaphor identification in a text corpus by a human is facilitated by information technologies. One of the existing annotation devices is, for example, BRAT instrument. Shutova's research was dedicated to automated annotation based on the previous manual one. After the metaphoric expressions had been identified manually, 241 out of 761 sentences turned out to contain metaphor, while metaphoric meaning was expressed by a verb in 164 phrases [17]. The suggested methods rest on the assumption that target concepts associated with the definite source concepts should appear in similar lexico-syntactic environments as the manifestation of language usage. Thus, clustering concepts using grammatical and lexical parameters enables to trace their associative relations and pick up a great number of metaphorical expressions beyond our seed set. For example, any previous sentence as a part of the seed set enables the system to identify metaphors in all following ones. Consequently, the proceeds from a seed set of metaphors performs unsupervised noun clustering with the aim of identifying target concepts associated with the same source domain; then by means of unsupervised verb clustering it creates a source domain verb lexicon [17].

A group of scholars [26; 1] introduce semasiological approach to identification of conceptual metaphors in a certain discourse. The scholars focus on a number of methodological solutions concerning the starting points of the research, samples of the potential metaphors.

This method is created to account for all potential conceptual metaphors in corpora under investigation and is based on combination of qualitative with quantitative analysis.

It is believed that “the development of automated systems capable of metaphors processing requires manually created electronic resources containing samples of metaphoric expressions and information referring to their correlation with conceptual metaphors to facilitate metaphoric annotation of corpora [3]. However manual compilation of “metaphoric” resources is considered the only possible way of corpora analysis which reflects the current situation. Nevertheless, there has been collected a huge amount of dictionary-based data about metaphor (manually compiled) that can be useful in the procedures of automated metaphors annotations.

Procedures of linguistic metaphors identification MIPVU contain indistinct instructions concerning identification of phrases based on metaphor/metonymy/ metaphonymy and proved to be unsuccessful when scholars attempted to verify it on the basis of the Russian language [3].

MIPVU operates the following classes of metaphoric expressions: Indirect Metaphor (MIPVU considers phraseological units as indirect metaphors; ambiguous cases (WIDLII); Implicit Metaphor; Possible Personification; cases subject to analysis (DFMA); Direct Metaphor and direct metaphor marker (mFlag) [19]. It is worth mentioning that the inclusion of Personification to metaphoric expressions is rather dubious so far as the suggested classification is based on different parameters. Firstly, whether there are any cases of implicit personification; secondly, personification is not the only type of conceptual interrelation (scholars identify alongside anthropomorphic personification – zoomorphic, phitomorphic personifications and personification of artifacts, depending on the type of concepts engaged in the process of metaphorisation).

Another team of researchers extended this list by adding two more classes: phraseological units (Set Expression) and Proper Name [3]. MIPVU treats proper names as indirect metaphors [19].

It's worth attention that Y. Badrylova, N. Shehtman, R. Kerimov keep to a broad definition of metaphor – as a transference of both human features and characteristics of any animated creatures to unanimated objects and abstract notions. In a number of cases of source domain identification it is practically impossible to specify the animated creature – whether a human being or an animal is implied.

Statistical methods are applied either as a purely statistical approach or as a complement of other methods, in particular semantic one. The latter employs “the degree of coordination between annotators” – the statistic methods used to estimate the degree of coincidences in decisions made by annotators independent of one another in analogous situations [3].

Some researchers identify metaphor by statistical methods [23; 15; 22; 7]. They bring forward an idea that the analysis should rely on statistical patterns of language usage [4]. They argue that any previous level of metaphor analysis informs the following; which makes it possible to avoid recomputing already-computed results. All levels of analysis rely on corpus statistical analysis and a certain amount of predefined semantic knowledge [4].

The scholars apply the parameter of “abstractness – concreteness” claiming that the noun is used metaphorically when its features do not associate with attributive adjective [4]. After the potential metaphors having been defined authenticity of metaphorisation is estimated by statistical methods and the achieved results are manually verified by corpus annotators [4].

A. Caruso introduces the method of metaphors identification based on the sampling of collocations within the collection of monothematic texts. The aim of such investigations is to find out the tendency of certain words to co-occur as an indicator of their literal or metaphoric meaning [16]. The method implies selective identification of metaphors which belong to a specific target conceptual sphere [16]. Unfortunately the author did not mention the coefficient used to calculate the degree of semantic associations of collocation components. The author proceeds pointing out that the procedure implied generating of concordances and manual analysis of context [16]. This methodology is less effective than the methodology suggested by Gandy et. al [4], but it is active when any other powerful resources and instruments developed for some languages, including English, are unavailable.

Comparative approach has been suggested by T. Shimizu, M. Shimokura presenting computer software “T-Scope version 2.0” to sample collocations [27]. ‘Mental Distanse’ analysis, introduced by T. Shimizu [27] is the key notion of their approach.

Consequently, one of the preliminary stage in metaphor identification in text corpora implies automated annotation of potential metaphors with the help of metaphor databases (that include lexical and phraseological databases, databases of figuratively used proper names). Metaphoric databases should include as their component information about metaphoric collocations, identified by means of n-gram method and semantically annotated respectively. This will enable further automated annotation of metaphoric expressions based on determination whether a certain element of the collocation belongs to a specific semantic class. Semantic class of the element can be determined in the course of its dictionary definition analysis. For example, the metaphoric meaning of the following Ukrainian noun word combinations (*На синьому оксамиті футляра лежала вибаглива золота браслетка...; і поглядає на свій маленький золотий годинничок: час іти; Мій Ярослав був золотою дитиною до шостого класу, поки не перехворів...; А таки комплекую, золота моя дівчинко...*) can be identified depending whether they belong to the class animated/unanimated creatures.

Scholars make use of dictionaries while performing different analytical procedures. O. Belikova and G. Gurin developing operational criteria of metaphor conventionality evaluation point out that in this case the dictionary serves as an analogue of an “average native speaker”. This criterion, first of all, enables to differentiate innovative, occasional, fresh, rare metaphors – on the one hand, and usual metaphors, registered in the dictionary – on the other. Secondly, if the reference meaning is not registered in the dictionary this may be interpreted as an indicator of metaphoric expression development in the language, its transference to the area of etymological metaphors [29]. The employment of dictionary information makes it possible to avoid limitations such as language consciousness and language competence of “an average speaker”. The dictionary should be considered as a product of collective consciousness of native speakers devoid of consciousness/competence fragmentariness of a single language personality.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Al-Saggaf M. A. Semasiological corpus-based approach to identifying conceptual metaphor (SCAICM) / M. A. Al-Saggaf, M. S. Yasin, I. Ho-Abdullah // Cognitive Linguistic Studies. – 2015. – Jan 1; 2(1). – P.116–128.*
2. ATT-Meta Project [Elektronic resource]. – Access mode : <https://www.cs.bham.ac.uk/~jab/ATT-Meta/Databank/>
3. *Birke J. A clustering approach for nearly unsupervised recognition of nonliteral language / J. Birke, A. Sarkar // 11th Conference of the European Chapter of the Association for Computational Linguistics. – 2006. – P. 329–336.*
4. *Caruso A. A corpus-based metaphor analysis of news reports on the Middle East Road Map peace process / A. Caruso // Proceedings of the Corpus Linguistics 2011 conference, University of Birmingham . – UK, 2011. – P. 2–9.*
5. *Deignan A. Metaphor and Corpus Linguistics / A. Deignan. – Amsterdam, 2005. – 236 p.*
6. *Deignan A. The Grammar of Linguistic Metaphors / A. Deignan // Corpus-Based Approaches to Metaphor and Metonymy. – Berlin ; N. Y., 2006. – P. 106–122.*
7. *Gandy L., Automatic Identification of Conceptual Metaphors With Limited Knowledge / L. Gandy, N. Allan, M. Atallah // AAAI 2013. – Jun 30. – Vol. 2. – P. 1–1.*
8. *Glucksberg S. On understanding nonliteral speech: Can people ignore metaphors? / S. Glucksberg, P. Gildea, H. B. Bookin // Journal of verbal learning and verbal behavior. – 1982, Feb 1. – 21(1). – P. 85–98.*
9. *Goatly A. The language of metaphors / A. Goatly. – London : Routledge, 1997. – 107 p.*
10. *Koivisto-Alanko P. Sense and Sensibility: Rational Thought versus Emotion in Metaphorical Language / P. Koivisto-Alanko, H. Tissari // Corpus-Based Approaches to Metaphor and Metonymy. – Berlin ; N. Y., 2006. – P. 191–213.*
11. *Koller V. Using a semantic annotation tool for the analysis of metaphor in discours / V. Koller. [Electronic resource]. – Access mode : in www.mataphorik.de/15/koller*
12. *Lakoff G. Metaphors we live by / G. Lakoff, M. Johnson. – University of Chicago press, 1989. – 265 p.*
13. *Martin J. H. A Corpus-Based Analysis of Context Effects on Metaphor Comprehension. / J. H. Martin // Corpus-Based Approaches to Metaphor and Metonymy. – Berlin ; N. Y., 2006. – P. 214–236.*
14. *Mason Z. J. CorMet: a computational, corpus-based conventional metaphor extraction system. Computational linguistics / Z. J. Mason. – 2004, Mar; 30(1). – 23–44. MetaNet [Electronic resource]. – Access mode : <https://metanet.icsi.berkeley.edu/metanet/>*
15. *Nayak S, Mukerjee A. A Grounded Cognitive Model for Metaphor Acquisition / S. Nayak, A. Mukerjee // InAAAI. – 2012, Jul 12.*
16. *Shimizu T. Mental Distance'concept for chronological metaphor analysis of business executive speeches. Osaka Keidai Ronshu / T. Shimizu // The Journal of Osaka University of Economics). – 2010. – 60(6). – P. 245–268.*
17. *Shutova E. Metaphor identification using verb and noun clustering / E. Shutova, L. Sun, A. Korhonen // Proceedings of the 23rd International Conference on Computational Linguistics. – 2010, Aug 23. – P. 1002–1010. Association for Computational Linguistics.*
18. *Shutova E. Statistical metaphor processing / E. Shutova, S. Teufel, A. Korhonen // Computational Linguistics. – 2013. – Jun; 39(2). – P. 301–353.*
19. *Steen G. J. A method for linguistic metaphor identification: From MIP to MIPVU / G. J. Steen, A. G. Dorst, J. B. Herrmann, A. Kaal, T. Krennmayr, T. Pasma. – John Benjamins Publishing, 2010.*

20. Steen G. J. Identifying metaphor in language: A cognitive approach / G. J. Steen // *Style*. – 2002, Sep 1; 36(3) . – P. 386–406.
21. Steen G. J. Is metaphor always a matter of thought? / G. J. Steen // *Issues in collecting converging evidence. Converging Evidence in Linguistics*. – 2008. – P. 33–54.
22. Stefanowitsch A. *Corpus-Based Approaches to Metaphor and Metonymy* / A. Stefanowitsch // *Corpus-Based Approaches to Metaphor and Metonymy*. – Berlin, New York, 2006. – P. 1–16.
23. Turney P. D. Literal and metaphorical sense identification through concrete and abstract context / P. D. Turney, Y. Neuman, D. Assaf, Y. Cohen // *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. – 2011, Jul 27. – P. 680–690. Association for Computational Linguistics.
24. VU Amsterdam Metaphor Corpus. [Electronic resource]. – Access mode : <http://ota.ahds.ac.uk/headers/2541.xml>.
25. Wallington A. M. Metaphoricity signals: A corpus-based investigation / A. M. Wallington, J. A. Barnden, M. A. Barnden, F. J. Ferguson, S. R. Glasbey // *COGNITIVE SCIENCE RESEARCH PAPERS-UNIVERSITY OF BIRMINGHAM CSRP*. – 2003.
26. Бадрызлова Ю. Г. Правила применения процедуры лингвистической идентификации метафоры (мірву) в русскоязычном корпусе: лингвокогнитивный опыт (уточнения и дополнения) / Ю. Г. Бадрызлова, Е. В. Исаева, Р. Д. Керимов, Н. Г. Шехтман // *Гуманитарный вектор*. – Серия: Филология, востоковедение. – 2013. – 4 (36).
27. Беликова А. Е. Методика оценки конвенциональности метафорических выражений: от интуитивистских критериев к операциональным / А. Е. Беликова, Г. Б. Гурин // *Ученые записки Петрозаводского государственного университета*. – 2012(1). – С. 44–50.
28. Кульчицкая Л. В. К вопросу об идентификации лингвистической метафоры / Л. В. Кульчицкая // *Вестник Иркутского государственного лингвистического университета*. – 2012. – 1 (17).
29. Хахалова С. А. Современное состояние исследований метафоры в европейской лингвистике / С. А. Хахалова, Е. В. Третьякова // *Вестник Северного (Арктического) федерального университета. Серия: Гуманитарные и социальные науки*. – 2014(2).

REFERENCES

1. Al-Saggaf, M. A., Yasin, M. S., Ho-Abdullah, I. (2015). Semasiological corpus-based approach to identifying conceptual metaphor (SCAICM). *Cognitive Linguistic Studies, Jan 1; 2(1)*, 116–128.
2. *ATT-Meta Project*. URL <https://www.cs.bham.ac.uk/~jab/ATT-Meta/Databank/>
3. Birke, J, Sarkar, A. (2006). A clustering approach for nearly unsupervised recognition of nonliteral language. *11th Conference of the European Chapter of the Association for Computational Linguistics*, 329–336.
4. Caruso, A. (2011). A corpus-based metaphor analysis of news reports on the Middle East Road Map peace process. *Proceedings of the Corpus Linguistics 2011 conference, University of Birmingham*. UK, 2–9.
5. Deignan, A. (2005). *Metaphor and Corpus Linguistics*. Amsterdam.
6. Deignan, A. The Grammar of Linguistic Metaphors. *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin; N. Y., 106–122.
7. Gandy, L, Allan, N, Atallah, M, Frieder O, Howard, N, Kanareykin, S, ... Argamon, S. (2013). Automatic Identification of Conceptual Metaphors With Limited Knowledge. *AAAI, Jun 30, vol. 2*. P. 1–1.

8. Glucksberg, S., Gildea, P., Bookin, H. B. (1982). On understanding nonliteral speech: Can people ignore metaphors?// *Journal of verbal learning and verbal behavior*. Feb 1. 21(1). P. 85–98.
9. Goatly A. (1997). *The language of metaphors*. London: Routledge.
10. Koivisto-Alanko, P., Tissari, H. (2006). Sense and Sensibility: Rational Thought versus Emotion in Metaphorical Language. *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin; N. Y., 191–213.
11. Koller, V. *Using a semantic annotation tool for the analysis of metaphor in discours*. Retrieved from in www.mataphorik.de/15/koller
12. Lakoff, G., Johnson, M. (1989). *Metaphors we live by*. University of Chicago press.
13. Martin, J. H. (2006). A Corpus-Based Analysis of Context Effects on Metaphor Comprehension. *Corpus-Based Approaches to Metaphor and Metonymy*. Berlin; N. Y., 214–236.
14. Mason, Z. J. (2004). CorMet: a computational, corpus-based conventional metaphor extraction system. *Computational linguistics*. 2004, Mar; 30(1), 23–44. MetaNet. Retrieved from <https://metanet.icsi.berkeley.edu/metanet/>
15. Nayak, S., Mukerjee, A. (2012). A Grounded Cognitive Model for Metaphor Acquisition. *InAAAI, Jul 12*.
16. Shimizu, T. (2010). Mental Distance'concept for chronological metaphor analysis of business executive speeches. Osaka Keidai Ronshu. *The Journal of Osaka University of Economics*, 60(6). 245–268.
17. Shutova, E., Sun, L., Korhonen, A. (2010). Metaphor identification using verb and noun clustering. *Proceedings of the 23rd International Conference on Computational Linguistics, Aug 23*, 1002–1010. Association for Computational Linguistics.
18. Shutova E., Teufel S., Korhonen A. (2013). Statistical metaphor processing. *Computational Linguistics, Jun;39(2)*, 301–353.
19. Steen, G. J., Dorst, A. G., Herrmann, J. B., Kaal, A., Krennmayr, T., Pasma, T. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins Publishing.
20. Steen, G. J. (2002). Identifying metaphor in language: A cognitive approach. *Style, Sep 1; 36(3)*, 386–406.
21. Steen, G. J. (2008). Is metaphor always a matter of thought? *Issues in collecting converging evidence. Converging Evidence in Linguistics*, 33–54.
22. Stefanowitsch, A. (2006). *Corpus-Based Approaches to Metaphor and Metonymy. Corpus-Based Approaches to Metaphor and Metonymy*. Berlin, New York, 1–16.
23. Turney, P. D., Neuman, Y., Assaf, D., Cohen, Y. (2011). Literal and metaphorical sense identification through concrete and abstract context. *Proceedings of the Conference on Empirical Methods in Natural Language Processing, Jul 27*, 680–690. Association for Computational Linguistics.
24. *VU Amsterdam Metaphor Corpus*. Retrieved from <http://ota.ahds.ac.uk/headers/2541.xml>.
25. Wallington, A. M., Barnden, J. A., Barnden, M. A., Ferguson, F. J., Glasbey, S. R. (2003). Metaphoricity signals: A corpus-based investigation. *Cognitive science research papers-university of birmingham CSR*.
26. Badryzlova, U. G., Isayeva, E. V., Kerimov, R. D., Shehtman, N. G. (2013). Pravila procedury primienieniya lingvisticheskoi identifikatsii metafory v russkoyazychnom korhushie: lingvisticheskiy opyt primienieniya. *Gumanitarnyi vector. Philologiya: vostokovedeniye*, 4 (36).
27. Belikova, A. E., Gurin, G. B. (2012). Metodika otsenki konvetsionalnosti metaforicheskikh vyrazhenii: ot intuitivistskikh kriteriev k operatsyonalnym. *Ucheniye zahiski petrozavodskogo gosudarstvennogo universiteta*, (1), 44–50.

28. Kulchitskaya, L. V. (2012). K voprosu ob identifikacii lindvisticheskoy metaphory. *Vestnik irkutskogo universiteta*, 1 (17).
29. Nahalova, S. A., Trietiakova, E. V. (2014). Sovremennoye sostoyaniye issledovaniy metaphory v yevropeyskoy linguistike. *Vestnik Severnogo (Arkticheskogo) federalnogo universiteta. Seriya: Gumanitarniye nauki*, (2).

Стаття: надійшла до редакції 15.10.2018

прийнята до друку 13.11.2018

СУЧАСНІ ПІДХОДИ ДО АВТОМАТИЗОВАНОЇ ІДЕНТИФІКАЦІЇ МЕТАФОРИ

Олена ЛЕВЧЕНКО

*Національний університет “Львівська політехніка”,
кафедра прикладної лінгвістики,
вул. Митрополита Андрея, 5, корпус 4, кім. 131, Львів, Україна,
тел.: (032)258 21 38*

Наталія РОМАНИШИН

*Національний університет “Львівська політехніка”,
кафедра прикладної лінгвістики,
вул. Митрополита Андрея, 5, корпус 4, кім. 131, Львів, Україна,
тел.: (032)258 21 38*

Висвітлено сучасні методи та процедури ідентифікації метафор у корпусах текстів. Основну увагу приділено перевагам комбінованого семантичного та статистичного принципу метафоричного анування корпусу, ролі механізму верифікації переносного значення мовних одиниць, який заснований на використанні відповідних словникових ресурсів для автоматизованої ідентифікації метафор.

Основою дослідження є текстові корпуси як джерело великого обсягу даних та розгляд відповідних алгоритмів, спроможних пояснити свідомі чи підсвідомі процеси ідентифікації метафор дослідниками, а також можливості спрощення процедури аналізу даних, у яких (процедурах) фактор суб'єктивності зменшено чи усунуто. Текстуальний аналіз метафор засновано на наступних ресурсах: система CorMet, репозиторій MetaNet, Amsterdam Metaphor Corpus, АТТ-Meta та корпус дієслів, позначених відповідно до їхньої належності до мети чи джерела метафоричного переносу. Етапи та стратегії ідентифікації концептуальних метафор становлять специфічні прийоми вилучення з неанотованих корпусів лінгвальних виразів, що містять метафори, та визначають процедури аналізу метафор шляхом ідентифікації можливого списку мовних метафор. Автоматичне вилучення метафор засноване на аналізі різноманітних мовних виразів, що вказують на метафори. До них належить широка парадигма переносних металінгвальних виразів, загальних метамовних виразів, які вказують на неоднозначність семантики та невідповідність між основним та контекстуальним значенням слова, в тому числі й орфографічні прийоми, наприклад, лапки.

Незважаючи на різноманітність наявних способів семантичного анотування (ідентифікування) у переважній більшості досліджень засобів вторинної номінації анотовані корпуси не можуть становити об'єктивної основи аналізу. Різноманітні аналітичні процедури ідентифікації метафор залучають словникові ресурси як модель "середньостатистичного носія мови" та/або як продукту колективної свідомості мовців на протилежності суб'єктивній фрагментарності окремої мовної особистості.

Ключові слова: метафора, ідентифікація метафори, текстовий корпус, анотування текстового корпусу.