DESIGNING "HUMAN CHARACTERS" LEXICAL DATABASE

Lukashevich N. Ju. (natalukashevich@mail.ru), **Kobozeva I. M.** (kobozeva@list.ru)

Lomonosov Moscow State University, Moscow, Russia

The paper discusses a general layout of "Human Characters" lexical database specifically developed to study the meanings of words from the semantic field of human character traits. It is intended as a resource providing a format for a comprehensive analysis of character words usage in different languages. A database with contexts from large modern corpora is considered a convenient tool for semantic analysis which offers such advantages as facilitating data storage and presentation, and keeping the analysis consistent while making changes possible at the same time. It is shown how several issues which significantly influence the analysis procedures are resolved in the pilot database version. These include identifying relevant contexts, describing features of a typical situation in which the character trait in question is exhibited, and comparing contextual meanings of the studied words. The suggested technique provides a more flexible tool for capturing similarities and differences between contexts within one language on the one hand, and gives ground for comparing the usage of translation equivalents on the other.

Key words: cognitive semantics, database, semantic analysis, human character, corpora contexts, behaviour pattern

Our work is devoted to the analysis of words from the semantic field of human characters—words either naming a person in accordance with the person's character traits (e.g., greedy, avaricious, mean; meanie, cheapskate) or the traits themselves (greed, avarice), or their derivatives (e.g. meanly, greedily; Russian zhadnichat' 'to be greedy' derived from zhadny 'greedy').

The semantic field of human characters

In psychology where the phenomenon of human character is the subject of study, there exists a variety of theories of human character and personality. In Russian psychology human character is mostly regarded as a system of individual psychological features revealed in human behaviour (in various types of activity, communication and interaction with other people). A character trait is seen as a stereotype of behaviour which is realised with high probability in a relevant situation. It is generally accepted that volitional qualities constitute a part of human character, whereas intellectual qualities do not. The relationship between character and temperament is still a subject of discussion.

To identify the limits of the semantic field of character one should not rely on scientific notions because language reflects naïve psychology: we need to know what properties speakers of a language consider as pertaining to human character (if this language has a lexical item with a given meaning). To do it, a special experiment was designed in (Lukashevich 2004) to check whether clear-cut temperamental, intellectual and volitional qualities represent character for native Russian speakers. According to the results obtained, volitional qualities are, whereas intellectual qualities are not regarded as character traits by Russian speakers. As for temperament and character, there is no clear boundary between these notions in Russian (for more details see (Lukashevich 2004)).

The nature of semantics of words from this field has been previously discussed in (Lukashevich 2002, Lukashevich 2004). Character words pose problems for a lexicographer, which is why their meanings are poorly represented in existing dictionaries (Kobozeva, Lukashevich 2012). A more effective approach was introduced in (Lukashevich 2004) It takes into account the above mentioned link between a character trait and the situation in which it is usually triggered. According to this approach a character word meaning is represented with the help of behaviour pattern—a generalised implicative scheme which links the initial typical situation with the stereotyped behavioural response of a person with this trait of character. (The notion («shablon povedenija» in Russian) was first proposed in (Martemianov, Dorofeyev 1969) and discussed in detail in (Martemianov 1999, Lukashevich 2002, Lukashevich 2004)).

To obtain the information about the typical triggering situation and the actions performed by a person characterised by a character word a thorough and extensive study of this word's use is required. This may be done by analysing large numbers of contexts with the word (thus possibly repeating the work a native speaker does when acquiring such notions). A database containing contexts from large modern corpora can help to keep the analysis consistent, and provide a convenient way to store both the material for studying and the results of such analysis.

Words denoting human character in existing database projects

Taking a look at the existing DB projects one can say that they are mostly not well suited for the purpose of performing an in-depth analysis of character word meanings. An important feature of this field is that it is mostly represented by rows of near-synonyms. The way they are handled in the current versions of such projects as WordNet and FrameNet has already been discussed in (Lukashevich, Kobozeva 2011): though both resources provide valuable information (WordNet—on synonymic relations between lexemes in a language¹, FrameNet—on the roles played by participants of the described situation), they leave the differences between near-synonyms mostly unclear. FrameBank, a Russian-language project of the FrameNet type, is intended as a hybrid of a dictionary of constructions with an annotated corpus. Another

WordNet related projects, such as Open Multilingual WordNet, can even relate synsets in various languages to each other

Russian-language project called Lexicograph mainly studies the relations between different meanings of a word and how each meaning predicts the way the word is used (and covers only verbs in its current version). As for the Typological database "The Vocabulary of Pain" it was developed to study a specific thematic class of words across languages, and its design, though remarkably flexible (Kostyrkin et al. 2012), does not suit the features of words from our semantic field (as what is studied in our case is not always explicitly present in the text).

In general it should be noted that in most of these projects the semantic field of human character is rather poorly covered (if intended to be included at all) (for example, such synonymic row as "greedy, covetous, avaricious" is not present in the current version of FrameNet). More importantly, the above mentioned projects seem more focused on the meaning of a word as a whole undergoing various processes like meaning shifts, and none intend to go deeper into the meaning itself. This research aims (beside a major goal of building a more adequate semantic representation of character words by identifying their behaviour patterns as well as their prototypical ("best") examples) to reach such a level of detail where distinctions between near-synonyms will be visible.

All of this called for developing a DB specifically designed for the purposes of the present project.

"Human Characters" Database: major problems

The present paper discusses a general outlay of "Human Characters" database (in its pilot version)². It is intended as a resource providing a format for comprehensive analysis of character words usage in different languages³. The DB is supposed

The work is being done within the framework of a seminar "Human characters through the prism of language" at the Department of Theoretical and Applied Linguistics, Philological Faculty, Moscow State University. At the moment it encompasses examples from the main subcorpus of the Russian National Corpus on several Russian adjectives and nouns from Greediness and Candidness semantic groups. The initial analysis of contexts is mostly done by students who fill in the DB in accordance with the provided guidelines After that everything is reviewed by the developers (i.e. the authors of the present paper), which means that every context is analysed at least by three persons (all native Russian speakers). Those issues which cause disagreement are then discussed at the seminar. If anything still remains unclear at this stage, such issues may be tested through specially designed experiments with native speakers (as in (Lukashevich 2004)).

Some time ago contexts from the British National Corpus displaying the use of several English adjectives and adverbs from Candidness group were taken for preliminary analysis. Some results of this analysis were presented in (Kobozeva, Lukashevich 2012). However, it became clear that BNC does not contain enough contexts to make reliable conclusions about fine distinctions between near-synonyms in a row. Besides, extracting all the relevant contexts from BNC presents certain technical problems, which means that it will be necessary to use material from other (bigger) corpora of English. Another crucial point will be to ensure that the results are reviewed by native speakers. (It should be noted here that the examples in the paper are given not only in Russian, but also in English on purpose. This is done to avoid unnecessary translation (in cases where a similar English example was easy to give) and to show that at least some features are shared by words from this group across languages.)

to list all the contexts in which the studied word is used in a given language corpus. Each (relevant) context is described with great detail from various angles: not only the exact meaning of the word in question, but most varied aspects of its usage and features of the context are identified (like the grammatical construction in which it is used, whether it conveys any evaluative meaning, whether there is a reference to a fiction character, etc). By generalising from contextual details in every aspect one can identify the major features of a typical situation associated with the named character trait and then formulate its behaviour pattern (i.e. specify which actions in what conditions are usually performed by a person so that this person may be assigned the named character trait).

Thus, for a comprehensive study of a character word various aspects of its usage should be taken into account. It should be remarked that not all these aspects and consequently not all details of the DB structure will be discussed here, but only those which significantly influence the analysis procedures.

1. Identifying relevant contexts

First of all, an important task is to identify the contexts which are relevant for this analysis. The thing is that it is typical for many adjectives and adverbs belonging to this group to have more than one meaning, or at least more than one usage. Character adjectives may be used with nouns (both animate and inanimate), as well as in predicative constructions. For example, *candid* can be found in the following contexts:

- (1) a) Maybe I delivered my opinion more bluntly than I should have, but I had always been candid with Ted.
 - b) He is either a **candid** friend or an honest enemy who disdains to tell lies.
 - c) ...her femininity was exquisitely candid.
 - d) ...his candid eyes never left Pitt's.
 - e) We take a **candid** look at the choices now open to you.
 - f) We are doing a **candid** camera in here today. (the British National Corpus (BNC))

Of the above only (1a) and (1b) describe a person and their actions directly and can be used to find explicit information related to why the person was assigned the characteristics in question. Although other examples may also be useful on further stages of study, especially when fine distinctions between near–synonyms are elucidated, such contexts are not included into the "core" analysis.

Character nouns can also display certain ambiguity, though this is not typical of the "core" members of the semantic field (derived from character adjectives—zhadina 'cheapskate' from zhadny 'greedy' in Russian, meanie from mean in English). But there are less central cases of metonymic or metaphoric nominations derived from names belonging to other semantic classes. For instance, lisa 'fox' in Russian has two meanings: 1) 'an animal'; 2) 'a cunning, honey-mouthed person' (Ozhegov 1990). For the purposes of this research only such contexts are considered relevant where the word in question either characterises a person in accordance with the

person's behaviour, or metonymically signifies some aspects of behaviour typical for a person with a particular character trait.

As for character adverbs, they are also used in more than one way, and not all of their uses need to be taken into account for the purposes of this research (Kobozeva, Lukashevich 2012).

This means that it is important for a researcher to be able to sort corpora contexts in accordance with their relevance for analysis. (It should be noted here that at the moment all the contexts which are extracted from a corpus in accordance with the search conditions are included into the DB irrespective of their relevance for several reasons. On the one hand, it is highly problematic to sort contexts automatically, so that only relevant ones are imported into the DB. On the other hand, the contexts which are not relevant for the "core" analysis may prove necessary on later stages of study, as it has already been mentioned above. Besides, the information they provide may be used for other purposes (e.g., to analyse regular polysemy models).)

Defining grammatical construction is helpful, but it is not enough to separate all the relevant contexts. For example, adjectives in predicative constructions can relate both to animate and inanimate nouns, and only sentences with animate nouns would be normally relevant here. On the other hand, (2) shows an example where personification takes place and (2) should therefore be taken into account:

(2) The bathroom mirror was **candid**, almost disapproving, whereas her bedroom mirror took and returned a more indulgent view, softening lines and contours. (BNC)

To resolve this, a special field "Relevant example" was introduced, so that when marked it signals that the context is relevant for analysis. (Nevertheless grammatical construction still needs to be identified, as these distinctions may prove significant for the choice of contextual meaning.)

2. Context length

It should be noted that whenever it is possible, it is important to have at hand not only sentences with the words in question, but longer contexts (as long as can be extracted from the corpus). The reason is that quite often the information necessary to specify the meaning in which the word is used is contained elsewhere in the text (mostly within the range of several sentences before or after the sentence with the word (e.g. as in 3a below)). (Here it should be noted that this is relevant only for those corpora which allow extracting more than one sentence with the word used, like the Russian National Corpus (RNC), for example. Some resources, such as the BNC, do not allow this, which often makes examples useless for the semantic analysis we perform.)

As quite lengthy pieces of text are required sometimes to understand the situation represented in the context (i.e. to be able to identify the conditions and the action of the person characterised by the word in question), a special field "Pattern Instance" (i.e. a specific realisation of a particular behaviour pattern linked with a particular

character trait) was introduced in the DB to make analysis and further discussion easier: short descriptions, or synopses, of the situations from the contexts should be recorded there. The general idea is to get an additional intermediate level of generalisation which would allow working with the example without having to read the whole text once again every time. These short descriptions will contain example information in a form which will make it easier to see in what way the given situation instantiates behaviour pattern. Such paraphrases should keep all the relevant features of the example and yet be short and concise. They should include only essential details of the situation, such that if these are left out, the example becomes unclear, or it is no longer enough to assign the named character trait to a person. (3a) and (3b) below show an English translation of a sample context from RNC illustrating the use of *korystolyubivy* 'avaricious' and a resulting Pattern Instance:

- (3) (a) ... the Russian revolution brought ... the petty bourgeoisie to the front row. Yes indeed, the very crass, avaricious petty bourgeoisie, which back at the beginning of our century was ridiculed by all liberal Russian writers, from Chekhov to Gorky. ... It was exactly the petty bourgeois—selfish, apolitical and devoid of ideology—who had waited out in the backwater for the Civil war storms to calm down, to then crawl out safe and sound and serve the Soviet power; it did not matter what kind of power to serve—just power, in order to grab a piece of the government pie as big as possible⁴. (the Russian National Corpus—Anatoly Gladilin, A long race day (1976–1981)⁵)
 - (b) Pattern Instance: *lower middle class waited out the revolution and the civil war and began to serve the current power* (no matter which one, here—the Soviet power) *to get as much money, property, etc. as possible*

3. Defining features of a typical situation

Another important task is to find the defining features of a typical situation in which the character trait in question is revealed. Initially the plan was to identify the Action, Motivation and briefly describe the overall Conditions of the situation (for example, just to indicate that it concerned "paying somebody for the work done"). However, it was decided at a certain point that it would be preferable to identify roles or participants of the situation in every relevant context as this information may give

^{4 ...} русская революция выдвинула на авансцену... мещанство. Да, да, то самое дремучее, корыстолюбивое мещанство, которое еще в начале нашего века осмеивали все прогрессивные русские писатели, начиная с Чехова и кончая Горьким. ... Именно мещанин, эгоистичный, аполитичный и безыдейный, переждал в тихой заводи, когда успокоятся бури Гражданской войны, и целым и невредимым вылез наружу, чтобы служить советской власти, не важно какой, важно, что власти, чтобы отхватить себе кусок правительственного пирога побольше. (НКРЯ — Анатолий Гладилин. Большой беговой день (1976–1981))

⁵ Authors' translation

clues to the differences between near-synonyms⁶. Thus, such fields as Subject, Object-Theme and 2nd Participant were introduced: Subject defines a participant of the situation which is characterised by the word in question; Object-Theme names such entity the attitude to which serves as the basis for distinguishing the character trait (or, to be more precise, the cluster of character traits); and 2nd Participant describes the participant whose interests are affected by the Subject's behaviour. For example, in (3) "lower middle class" is the Subject, "material wealth" is the Object, and 2nd Participant is missing. The Action in this example is "serve the current power", Motivation is "to obtain as much material wealth as possible", Conditions may be defined as "when there is a choice between obtaining material wealth and following one's political principles".

4. Representing the meaning

The third task constituting the crucial part of analysis is to find differences and similarities between the exact meanings in which the words in question are used in contexts.

Initially a sort of a two-level tree structure was used to make such generalisations. The main point was to get non-overlapping subsets of contexts and to specify what makes the cases of human behaviour similar within each subset. This helped to identify the main features of actions typical for various instances of behaviour associated with the studied character trait. As it was merely an intermediate step, it was not of much consequence that the resulting groups of contexts were not always of "equal weight", e.g. not always on the same level of generalisation. For example, for Russian *otkrovenny* ('frank,candid') the groups were as follows:

- 1) the person says smth about him/herself which is not obvious and this may lead to negative consequences to the person:
 - tells smth negative about him/herself;
 - speaks of his/her real feelings;
 - says what (s)he is thinking;
- 2) the person does not hide anything;
- 3) the person says what (s)he is thinking although it violates etiquette rules;
- 4) the person tells the listener smth negative about the others which is not obvious; etc.

Thus it was clear that for *otkrovenny* a second level of generalisation could be identified as several meanings shared a common part.

However, one level of generalisation could also be enough, as was the case for Russian *iskrenny* ('sincere') where the groups were as follows:

- 1) the person says what (s)he is really thinking;
- 2) what the person is doing, corresponds to what (s)he is saying and thinking;

⁶ Here not the semantic-syntactic roles of the sentence with the character word are meant, but roles of the participants of the behaviour pattern.

- the person says what (s)he is really thinking and this leads to negative consequences for others;
- 4) the person says what (s)he is really thinking and this leads to negative consequences for this person;
- 5) the person demonstrates feelings, which (s)he is really experiencing;
- 6) the person says smth about his/her real feelings and thoughts which is unpleasant for the listener; etc.

It is easy to note that certain parts of the meanings roughly formulated above are repeated in various combinations. The way they are combined does not allow to fit them into a clear tree structure. However, the resulting picture is of much help in identifying differences between these two near-synonyms and making higher level generalisations for behaviour patterns.

Taking all of this into account a two-level structure was introduced in the DB consisting of Meaning General and Meaning Subtype. The fields were supposed to divide the contexts into non-overlapping subsets, such that each context within a subset depicts one type of situation sharing some features and distinctly different from situations in other subsets. Meaning Subtype is supposed to specify Meaning General, in a way that Meaning General is subdivided into several Meaning Subtypes.

However, an attempt to analyse in a consistent way large numbers of examples with several words belonging to the same group (Greediness) showed that this would not work as intended. Firstly, it proved quite difficult to get non-overlapping groups (no matter how small they were). Secondly, there were cases when some elements of sense appeared sporadically and their presence did not seem to be linked to any particular Meaning General or Meaning Subtype. (For example, such was the pleasure a person felt when in direct physical contact with money he possessed—a sense identified in several contexts for *skupoy* 'stingy' and *zhadny* 'greedy' in Russian.)

Thus it soon became clear that though still being helpful, this two-level structure is not enough for capturing the distinctions identified.

The above mentioned difficulties with this group of words were only to be expected due to such feature of their semantics as Wittgensteinian family resemblance (a situation when objects in a set share common features in pairs, but no object has all the features at once (Wittgenstein 1953)). Family resemblance effects have already been mentioned as the reason why it is difficult to identify the boundaries between rows of synonyms and between volumes of near-synonyms meanings in a row (Lukashevich 2004, Kobozeva, Lukashevich 2011). The above mentioned difficulties show that similar effects are seen not only on the level of boundaries between the meanings of words, but also between various usages of one word.

Understandable and foreseeable as these difficulties are, it was necessary for the research purposes to find some way around them as this interim step is a key part of the analysis. Obviously this is yet another example of non-discreteness in language which has to be accounted for using discrete tools (in the absence of any other ones which would better suit the nature of the material), a problem discussed in (Kibrik 2013).

To overcome them we use the following technique: we introduce a list of component elements of meaning and in each context choose the ones relevant to this

particular instance. These elements are such "chunks" of meaning which have been identified on the previous stages of analysis as parts which are used as whole units. For example, it is clear in the roughly named meaning groups for otkrovennny and iskrenny that a part of meaning 'what X is saying may lead to negative consequences to X' may be added to other "blocks" like that and this way form a meaning in which the character word is used. When analysing examples with words which belong to the Candidness group in English (candid, frank, sincere, open) it was obvious that this "block" is also often present in their meanings, as well as other "blocks" of meanings which can be identified for othrovennny and iskrenny (Kobozeva, Lukashevich 2011). Therefore it seems possible to make a list of such parts of meaning for words instantiating the Candidness frame (in the FrameNet sense) and use it to describe various usages of words belonging to this group (possibly not only in English and Russian). Thus, obtained through a procedure similar to the one used by A. Wierzbicka when identifying semantic primitives (Wierzbicka 1972), such components of meaning represent the level on which it is possible to compare the meanings of near-synonyms in a row in one language as well as the meanings of translation equivalents across different languages.

In the course of analysing Russian adjectives and nouns from the Greediness group the following list was identified:

- seek not to spend resources
- seek not to spend money on others
- · seek to increase their material wealth
- · seek to obtain material wealth
- attach primary importance to money
- · attach primary importance to material values
- violate moral (=ethical) rules with regard to others
- · violate legal norms
- seek to be the sole user of a resource
- commit an action undesirable for its object/experiencer
- etc.

At the moment the list includes 27 elements. There will most likely be additions when greediness in other languages comes under analysis, however, the "core" list presumably will not change greatly. It is worth mentioning that such components of meaning should be expressed using a semantic metalanguage (in a sense that they should contain units from a final list, such units should be used in a consistent manner, and their meanings may not be identical to the meanings of these words when used in a real language)⁷.

From the example above it is clear that such elements will most likely have different "weight" in the list: some will be more central for a particular group (like the ones related to the attitude to material values for the Greediness group), whereas others, more "peripheral" ones, will include concepts from other spheres (e.g. 'violate moral rules with regard to others' is relevant not only for *korystolyubivy* 'avaricious' but also for *verolomny* 'deceitful', etc.).

⁷ In this project metalanguage is being built incrementally, one character group after another.

Whether such lists can be universal requires further research, but they will at least provide a more flexible tool for capturing similarities and differences between contexts within one language, and at the same time give a ground for comparing the usage of translation equivalents on the other.

To illustrate what has been said above we show how semantic information for (3a) and (3b) may be recorded in the DB:

(level 2) Meaning Subtype = the Subject attaches great importance to material values and seeks to obtain material wealth using his/her proximity to power

Components of meaning: 'attach great importance to material values' + 'seek to obtain material wealth' + 'use their proximity to power' + 'as much as possible'

(level 1) Meaning General = attach great importance to material values and seek to obtain material wealth violating, in so doing, certain rules

(level 0) Behaviour Pattern (tentative):

Korystolyubivy =

= when obtaining material wealth is possible such people prefer material values over other values and seek to obtain or increase their material wealth though it may lead to negative consequences for them and/or other people.

It is typical for such a person:

- to choose a job offer with the highest payment (when it is not necessary) though it may force them and their family to live in extreme conditions;
- to serve no matter which power so that they can use their position to get money, property, etc.;
- to betray a friend for money;
- for a doctor to refuse to go and see the patient for the second time if (s)he does not get paid for the visit;
- for an official to take bribes.

The DB also contains fields describing other important aspects of use (such as the presence of an evaluative component) or important features of the context (e.g., the presence of synonyms in conjunction with the word, etc.), which will not be discussed here. To give an idea of what it looks like, Fig. 1 shows a print-screen from the DB with an example from the Corpus of Contemporary American English. It is important that the suggested DB should allow to use the various recorded analysis results for the purposes other than the main research goals (e.g. use the information on evaluative component for sentiment analysis).



References

- 1. British National Corpus, available at http://www.natcorp.ox.ac.uk/ (as of May 2012)
- 2. Corpus of Contemporary American English, available at http://corpus.byu.edu/coca/ (as of 20.04.2014)
- 3. FrameBank, available at http://framebank.ru/ (as of 20.04.2014)
- 4. FrameNet Project (II), available at https://framenet.icsi.berkeley.edu/fndrupal/
- 5. *Kibrik A. A.* (2013) Non-discreteness in language and focal structure [Nediskretnost' v jazyke i fokalnaja struktura], Cognitive Modelling: Proceedings of the 1st International Forum on Cognitive Modelling [Kognitivnoe modelirovanie: Trudy Pervogo Mezhdunarodnogo foruma po kognitivnomu modelirovaniju], Italy, Milano-Marittima, part 1, pp. 113–116.
- Kobozeva I. M., Lukashevich N. Ju. (2012), Human characters through the prism
 of adverbs, Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialog 2012" [Komp'iuternaia Lingvistika
 i Intellektual'nye Tekhnologii: Po materialam ezhegodnoi Mezhdunarodnoii
 Konferentsii "Dialog 2012"], Bekasovo, pp. 277–287.

- Kostyrkin A. V., Panina A. S., Reznikova T. I., Bonch-Osmolovskaya A. A. (2012), Constructing a lexico-typological database (for a study of pain predicates), Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialog 2012" [Komp'iuternaia Lingvistika i Intellektual'nye Tekhnologii: Po materialam ezhegodnoi Mezhdunarodnoii Konferentsii "Dialog 2012"], Bekasovo, pp. 288–295.
- 8. Lexicograph Project (2010), available at http://lexicograph.ruslang.ru/(in Russian)
- 9. *Lukashevich N. Ju.* (2002), Predicates describing character traits and behaviour patterns [Kharakterologicheskie predikaty i shablony povedeniia], Journal of MSU, Philology series [Vestnik MGU, ser. Filologiia], Moscow, 2002. № 5, pp. 131–141.
- 10. *Lukashevich N. Ju.* (2004), Cognitive semantic analysis of predicates denoting human character traits [Kognitivno-semanticheskii analiz predikatov, oboznachaiushchikh cherty kharaktera cheloveka], PhD thesis, Moscow, 2004.
- 11. *Lukashevich N. Ju., Kobozeva I. M.* (2011), Character nominations in ontological perspective, Computational Linguistics and Intelligent Technologies: Proceedings of the International Conference "Dialog 2011" [Komp'iuternaia Lingvistika i Intellektual'nye Tekhnologii: Po materialam ezhegodnoi Mezhdunarodnoii Konferentsii "Dialog 2011"], Bekasovo, pp. 468–477.
- 12. *Martem'ianov Ju. S.* (1999), Metalanguages of sentence and text description [Metaiazyki opisaniia predlozheniia I teksta], Text processing and cognitive technologies [Obrabotka tekst i kognitivnye tekhnologii], Moscow, 1999. № 3, pp. 124–129.
- 13. *Martem'ianov Ju. S., and Dorofeev G. V.* (1969), Logical inferences and eliciting relations between sentences in the text [Logicheskii vyvod i vyiavlenie sviazei mezhdu predlozheniiami v tekste], Machine translation and applied linguistics [Mashinnyi perevod i prikladnaia lingvistika], Moscow, 1969, issue 12, pp. 36–60.
- 14. Open Multilingual WordNet (1.0), available at http://compling.hss.ntu.edu.sg/omw/cgi-bin/wn-grid.cgi
- 15. *Ozhegov S. I.* (1990), Dictionary of the Russian language [Slovar' russkogo jazyka], Russkij jazyk, Moscow.
- 16. Russian National Corpus, available at http://www.ruscorpora.ru/ (as of 01.2014)
- 17. Typological database "The Vocabulary of Pain" available at http://orientling.ru/bolit/ (as of 20.04.2014),
- 18. Wierzbicka A. (1972), Semantic Primitives, Athenäum-Verlag.
- 19. Wittgenstein L. (1953), Philosophical Investigations, Blackwell Publishing. 2001.
- 20. WordNet Lexical Database (version 3.1), available at http://wordnetweb.princeton. edu/perl/webwn (as of 20.04.2014)