FRAMES REVISITED: AUTOMATIC EXTRACTION OF SEMANTIC PATTERNS FROM A NATURAL TEXT[1]

Kotov A. A. (kotov_aa@nrcki.ru),
Zaidelman L. Y. (zaydelman_ly@nrcki.ru),
Arinkin N. A. (arinkin_na@nrcki.ru)
Kurchatov Institute; Russian State University for the Humanities;
Moscow, Russia

Zinina A. A. (zinina_aa@nrcki.ru)
Kurchatov Institute, Moscow, Russia

Filatov A. A. (alexander.filatov@hp.com)
HP Inc., Moscow, Russia

Our project aims to design a syntactic parser, which constructs a semantic representation in a frame format: a clause is represented as a table of valencies, filled in with semantic markers. This representation is compared to a list of scripts—used to disambiguate and classify the semantic representation as well as to select an appropriate reaction for a companion robot F-2.

Key words: semantic frames, syntax parser, semantic analysis, sentiment analysis

ВОЗВРАЩЕНИЕ К ФРЕЙМАМ: АВТОМАТИЧЕСКОЕ ВЫДЕЛЕНИЕ СЕМАНТИЧЕСКИХ ПАТТЕРНОВ ИЗ ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЫКЕ

Котов А. А. (kotov_aa@nrcki.ru),
Зайдельман Л. Я. (zaydelman_ly@nrcki.ru),
Аринкин Н. А. (arinkin_na@nrcki.ru)
Курчатовский институт; Российский государственный гуманитарный университет; Москва, Россия

Зинина А. А. (zinina_aa@nrcki.ru)
Курчатовский институт, Москва, Россия

Филатов А. А. (alexander.filatov@hp.com)
HP Inc., Москва, Россия

1. Introduction

The notion of *frame* [Minsky, 1988] has been a key concept for semantics for several decades, until it was recently replaced by a network approach where information is stored in a trained classifier and is not readily available for linguistic analysis. We developed an automatic system for the analysis of natural language text, designed to invoke emotional reactions (gestures, facial expression, speech replies) from the meaning of incoming texts. The system can animate a companion robot F-2 or can store the extracted meaning with a suggested emotional reaction to a database. In this work we describe the processing of the incoming text up to its meaning which is stored to the database. We return to the classical form of semantic representation—*frame*, and describe semantics as a set of semantic features (markers) divided up into a set of valencies (*agens, patient,* etc.). To construct the semantic representation we develop a parser, implementing morphological, syntactic and semantic processing of an incoming text as suggested by theoretical linguistic models, e.g. [Melčuk, 1999]. The parser is written in C#, the grammar of Russian is described in syntXML language and the dictionary is stored in an SQL database. On each step of processing, the parser may save the intermediate results of analysis to an SQL database or transfer them to the next software component—Fig. 1. In this work we present a general architecture of the parser and the methods to analyze the semantic representations, extracted during a text analysis.
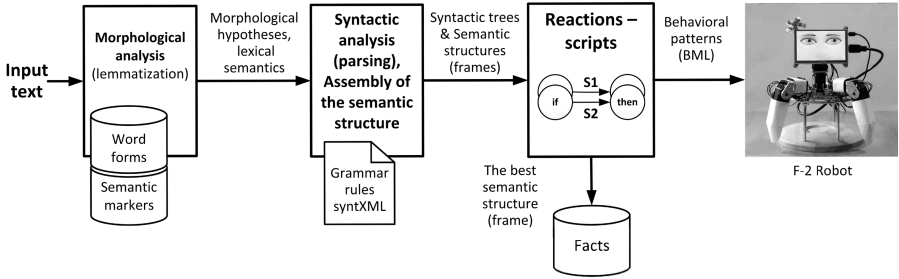
Fig. 1. The architecture of the text analysis

2.   The architecture of the parser

Scripts (the models of emotional reactions) are designed to detect emotional (input) patterns in incoming texts [Kotov, in press]. The representations, to be searched by scripts, are quite shallow: '*X* is a tasty food' or 'I was hit by someone' are sufficient patterns to suggest an emotional reaction to the robot. To detect the patterns, parser should store the main taxonomic markers for natural words—suggest that *John* is 'a person' and *strawberry* is 'a food', as well as the emotional markers—*idiot* is 'an inadequate person'. The parser should also allocate the markers to specific valencies in order to distinguish the antagonist and protagonist roles in an emotional event. Emotions (unlike formal logical notation) are sensitive to the syntactic structure of a sentence: as suggested in a classic work by Blakar, *The police took in the demonstrators* sounds more violent and emotional than *The demonstrators were taken in by the police* [Blakar, 1979]. This implies that the changes of diathesis should be revealed in the required semantic representation in a form of semantic/syntactic roles—valencies. Therefore, the task of the parser is to construct a semantic representation, where semantic markers are applied to a set of valencies. This exactly correspond to the classic representation of *frame*. Table 1 shows the semantic representation of an utterance *Linguists have noticed psychologists at the conference* (*Lingvisty otmetili psihologov na konferencii*). Such semantic representations can be quite shallow, sufficient only to detect the input patterns for the activation of scripts. In the architecture of the parser, scripts play the major role—they select the best tree with the precision, sufficient to extract the input patterns. The distinction between trees, equally weighted by the emotions, is ignored.

Table 1. Semantic representation (frame) of the sentence
Linguists have noticed psychologists at the conference

Valency	*p*	*ag*	*cont*	*loc*
Semantic markers	think pay-attention	somebody profession	somebody profession	abstract-entity container abstract-container

2.1. Morphological processing: lemmatization and dictionary

For a given wordform, the lemmatizer should extract (a) a required set of semantic markers and (b) a set of grammemes to construct a syntactic tree and further allocate the valency for the specific word. The morphological dictionary is based on the Open-Corpora project [Granovsky, Bocharov, Bichineva, 2010] and stores all wordforms and grammemes for 98,000 lemmas. To each incoming segment the lemmatizer assigns a number of morphological hypotheses. In case no hypothesis is assigned, a segment is consecutively processed by (a) a user dictionary, where one can store a noun with an inflectional class and (b) a guesser, which assigns 5 morphological hypotheses, based on a trained Keras neural network (for standard words this offers 96,7% recall). Words from a user dictionary and guesser do not get any semantic representation and are considered as semantically void. 30,000 words in the main dictionary are manually annotated by semantic markers (from 1 to 18 markers per word). We rely on semantic markers from two main groups: taxonomy markers and emotional markers.

(a) **Taxonomy markers** are used to classify the word and the whole predication. We use markers like 'somebody', '[named after] profession', 'move', 'move body', here we rely on the annotation suggested in the Russian semantic dictionary [Shvedova, 1998]. Unlike in traditional ontologies, a word may keep semantic markers from different classes: *bank* has the markers for 'organization', 'building' and 'abstract container'. This polysemy allows us to simulate *situational effects* [Yeh, Barsalou, 2006], where a word meaning is shifted depending on the situation or the emotion to be invoked by semantics—top-down emotional processing [Clore, Ortony, 2000]. In our case, different scripts may detect different markers in the lexical semantics, addressing different word meanings.

(b) **Emotional markers**, on one side, are markers, explicitly expressing a position in an emotional situation and an emotional evaluation like words *fool*, *great* or *useless*. These words are annotated by specific script markers: 'agent/antagonist in a situation of inadequate actions' (INADEQ script), 'protagonist in a situation of superiority' (SUPER), 'antagonist/excessive for the actions' (FREEDOM) or 'protagonist/someone, incorrectly ignored' (UNNEED). On the other side, we annotate ambiguous markers, that may contribute to emotions—*emotionally sensitive markers*, like 'intensity'—relevant in the emotional phrases like *Why do you push me?* [Glovinskaya, 2004].

Emotional markers have attributes, indicating their position in the focal/background semantics of the corresponding word. For example, the word *fool* has the markers 'inadequate' in its focal zone, and 'somebody' ('human') in its background zone, while the word *man* has the marker 'somebody' in its focal zone.

Words may have different meanings (homonymic and polysemic), that is represented by the number of subsets of semantic markers—one subset for each meaning. Different input patterns of scripts may address different markers in the semantics of a word. A script will select the meaning that fits best its input pattern.

2.2. Syntactic processing: syntXML

The task of the syntactic processor is to construct a syntactic representation where each actant gets its valency. The parser implements a left-to-right processing and combines rule-based approach to construct syntactic links with the trained classifier approach to evaluate the syntactic structures. Syntactic rules are designed on syntXML language [Kotov, Zinina, Filatov, 2015]. A rule is defined as a possible reduction, where the right-hand side (one or many segments *a, b, … n*) can be reduced to the left-hand head h (1). Head h can also be a member of the right-hand side and subordinate other segments (2).

$$h \rightarrow <a, b, … n> \tag{1}$$

$$h \rightarrow <a, b, h, … n> \text{ or } <a, b, h^{\text{head}}, … n> \tag{2}$$

Right-hand side of a rule may contain one or many segments *a, b, … n*, segments except *n* can be optional. This allows us to describe immediate constituent rules, where h is not a member of the right-hand side rule section, as well as dependency rules—both with any number of subordinate segments, so the grammar is not limited to binary relations. A rule can (a) verify if a segment matches a specific lexeme, (b) check the presence of a particular grammeme within a segment, (c) check grammatical agreement between rule segments, (d) set or rewrite grammemes for the rule head h, (e) copy grammemes/markers to the rule head h. This combination is represented in an XML form—as a syntXML rule. The grammar for Russian uses 550 rules and structurally corresponds to the System of Syntactic Groups by Gladkiy [Gladkiy, 1985]. In particular—conjunctions are defined as immediate constituents with a virtual head, as the features of the whole construction cannot be fully represented by any segment, whereas other grammar structures are represented as dependency trees.

Each incoming segment is added to a stack and the parser tries to reduce the stack head with a grammar rule, upon the reduction the subordinate segments are excluded from the stack and are linked to the rule head. Morphological and syntactic ambiguities are accounted: a separate stack instance is constructed for each morphological hypothesis and for each ambiguity in the application of syntactic rules. At the same time, the maximum number of stacks m is limited, currently by $m = 512$. Each rule application is evaluated, taking into account the corpus data (basing on a converted SynTagRus corpus): co-occurrence of (a) particular words within this rule, (b) word2vec vectors of the words within the rule, (c) sets of grammemes of the words within the rule. In other words, a syntactic link is more probable, if it exists in the corpus between: (a) the same words, (b) similar words, where similarity is calculated by word2vec distance, or (c) words with the same sets of grammemes. An aggregated score is calculated for each stack, after that m best stacks are preserved for further analysis. Stacks with lower scores (stacks with less probable syntactic links and stacks with few syntactic links) are discarded. In case several trees are generated by the end of the sentence (the possible number is from 0 to m), the best tree is determined and selected depending on its comparison with scripts, as explained in 2.3.

The semantic structure for a tree is constructed in the following way. For each predicate (finite verb, predicative, etc.) a valency p is assigned. Once a rule binds

a predicate with an actant (usually, a noun phrase), it may assign a valency to the noun phrase. In our project we rely on the list of valencies, suggested by [Fillmore, 1968], and use a list of 22 valencies: *ag* (agent), *pat* (patient), *instr* (instrument) etc. We have marked subcategorization frames for 13,000 Russian verbs, in part, relying on data from FrameBank project [Lyashevskaya, 2010]. A verb gets a number of syntactic markers, where each marker allows the verb to trigger a specific rule to assign an actant to a particular valency. For example, the subcategorization frame of the verb *zvenet'* ('to ring'—*звенеть*) has 8 such valencies, as shown in Table 2.

Table 2. Subcategorization frame of the verb *zvenet'* ('to ring'—*звенеть*)

Actant: case and preposition	NP, nominative, no prep.	NP, instrumental, no prep.	NP, accusative, prep. *v* (in)	NP, genitive, prep. *ot* (from)	NP, genitive, prep. *u* (at)	NP, accusative, prep. *o* (about)	Adverb or adverbial NP	Adverb or adverbial NP
Valency	*ag* (agent)	*instr* (instrument)	*targ* (target)	*caus* (cause)	*pos* (possessor)	*ca* (counterparty)	*t* (time)	*loc* (location)

Thus, the syntactic structure of the sentence *U menya zvenit v ushah* ('It rings in my ears'—*у меня звенит в ушах*) will assign the predicate *zvenit* ('rings') to **p** valency, NP *u menya* (~'at me') to **pos** valency (this triggers 1-y-Sgen-pos marker of the predicate) and NP *v ushah* ('in ears') to **loc** valency (this triggers 1-loc marker).

Following the syntactic analysis, the semantic representation is constructed based on the lexical semantics of each word and the semantic valencies assigned to the words in the syntactic structure. Semantics of a single clause is represented by a *semantic predication*—a two-level tree with predicate **p** at the top and a number of actants at the bottom—where each node (predicate or actant) is a set of semantic markers. Such a semantic tree can be represented as a table—*frame*—taking into account that **p** is the head of the whole clause. The semantic frame of the sentence *Lingvisty otmetili psihologov na konferencii* ('Linguists have noticed psychologists at the conference') is shown in Table 1, while *U menya zvenit v ushah* ('It rings in my ears')—in Table 3.

Table 3. Semantic representation (frame) of the sentence
U menya zvenit v ushah ('It rings in my ears')

Valency	*p*	*pos*	*loc*
Markers	to-sound	object somebody other-person principal (speaker)	object container container-small part part-of-human

For a compound (multi-clause) sentence several frames are constructed: one frame for each predicate—with co-reference or attributive links between the valencies of the frames. Further, the semantics of each tree is compared to the list of scripts (their input patterns) to determine the best tree and to suggest an emotional reaction for the F2 robot.
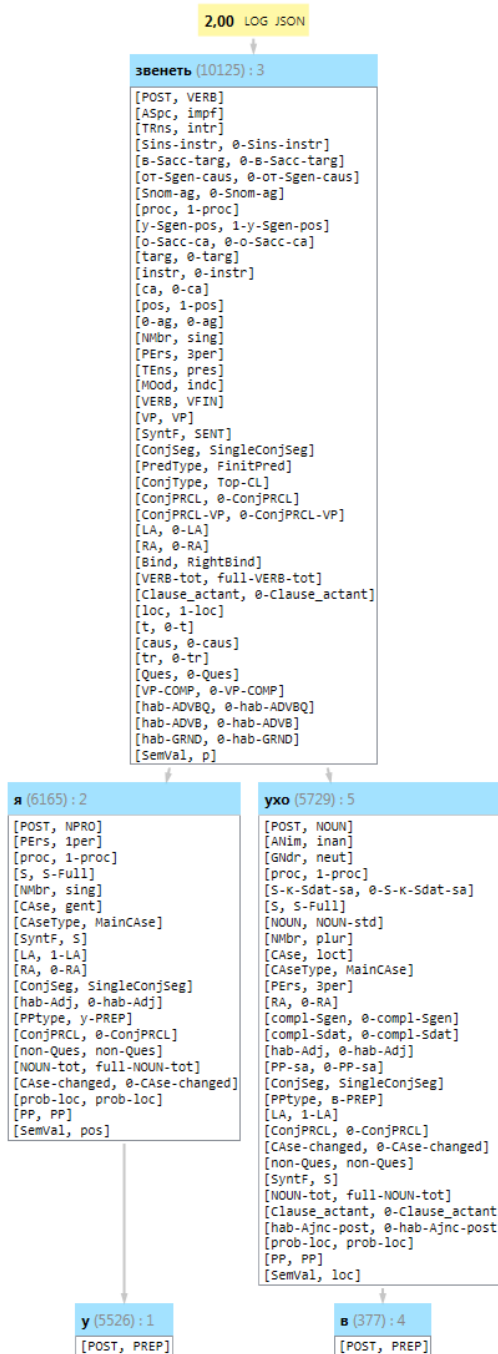
Fig. 2. Syntactic structure of the sentence *U menya zvenit v ushah* ('It rings in my ears'—*у меня звенит в ушах*)

2.3. Evaluation with scripts

To classify and evaluate the sentences, as well as to simulate the emotional dynamics for the F-2 robot, the parser relies on the set of *scripts*. Scripts are defined as *if-then* productions, a script is activated once an incoming semantic predication matches its' *if*-condition—*input pattern*. We consider an input pattern as an implementation of *frame*—a structure, similar to the semantic representation of a clause. For each incoming semantic predication, the parser calculates the distances with the input patterns of all the existing scripts. A tree with the highest similarity to the scripts is selected. We rely on the list of emotional scripts, represented in [Kotov, in press]. The list includes 13 scripts for negative situations: DANGER, APPROPR ("Appropriation"), SUBJV ("Subjectivity"—e. g. 'all he thinks about is himself!') etc., and 21 scripts for positive situations: CONTROL, CARE, COMFORT, ATTENTION (e. g. 'they all adore you!'), APPROVAL (e. g. 'you did it like a real man!') etc. By comparison of each incoming meaning (frame) to the script's input pattern the parser tries to classify a situation as a negative event—'this is dangerous', 'he is inadequate', or as a positive event—'they all adore me!'

Table 4. Distances with scripts and the possible responses for the utterance *Lingvisty otmetili psihologov na konferencii* ('Linguists have noticed psychologists at the conference')

Similarity	Script	Possible response
1.14	ATTENTION—gratitude	*Thank you for your support!*
1.14	ATTENTION	*It seems, everyone notices me!*
1.09	PLAN	*Did they mean that? They plan something evil!*
1.09	SUBJV	*They only think about this!*
1.01	UNUSUAL	*It's a whole new and unusual world!*
1.01	COMFORT	*It's nice to be here!*

As shown in Table 4, ATTENTION script is responsible for the "experience" of emotion, while "ATTENTION—gratitude" is responsible for a communicative reaction—the expression of gratitude, where the internal "expression" of emotion is not definitively required. In both scripts the robot associates itself with the **pat** valency ('a psychologist')—the most emotionally relevant position for the scripts. If any marker suggests a better reference to the robot (*Robot has noticed psychologists, Linguists have noticed the robot*)—the activation of scripts changes accordingly. Both attention scripts are sensitive to 'pay-attention' marker in **p** valency and 'somebody' markers in **ag** and **pat** valencies. Further scripts PLAN and SUBJV are responsible for negative reactions—they can be used to simulate a depressive mood or irony [Kotov, 2005]. Scripts UNUSUAL and COMFORT are sensitive to the situations where an agent acts inside or is entering some "magic world" or a "comfort place". They are sensitive to the actions, taking part in 'containers': UNUSUAL treats 'the conference' as a 'big and unusual container', like *walking in a wild forest*, while COMFORT considers the conference as a tiny and cozy container. As each semantic representation (frame) is associated with a script input pattern, each pair *<script input pattern, semantics>* serves as an illustration of *frame* and its *actual fillers*, like 'an idiot' = *John* or 'a cozy place' = *conference*.

Adverbs and adjectives may add additional markers to the predicate structures, changing their evaluation. For an utterance *Linguists always notice psychologists at the conferences*—subjv ("Subjectivity") becomes the leading script with the similarity 1.53—robot judges the 'linguists' as "limited" persons, who 'can do nothing else but noticing the psychologists'. This is an example of *emotionally sensitive marker*, as described in 2.1(b). Although the most active script is considered as the most relevant classifier of the situation and the best reaction for the robot, other scripts may be preferred for specific reactions: *They plan something evil!* reaction (plan) may be invoked to simulate robot's depressive state or irony [Kotov, 2009]. This ambiguity in the perception of text semantics may be used to study the architecture of consciousness—where one script (and the corresponding representation) constitutes the "accurate" understanding, while other scripts may form "ironical" or "hypothetical" understanding of the text [Kotov, 2017]. Different scripts may even suggest the selection of different syntactic trees, this option is not actually used, although it can be useful for computational simulation of humor, where the second semantics (tree) is used for the humorous semantic shift.

3. Facts database

Parser operates in a daily test mode: it collects through RSS and processes about 6,000 sentences per day from 25 information sources—15 media sites and 10 most popular LiveJournal blogs. The texts are analyzed by syntactic and semantic modules at a speed of 100 sentences per minute, the extracted semantic representations are stored in a database (PostgreSQL). As the parser is designed for the extraction of a specific set of scenarios (input patterns), the results of analysis cannot be directly compared to the results of other parsers: the analysis is not sensitive to the changes in the syntactic trees, not relevant to the input patterns. As we suggest, one of the possible extensions of the parser is the conversion of the extracted semantic representations to the input patterns of scenarios: up to now the parser distinguishes only the emotionally relevant representations. However, the extracted facts (as in Tables 1 and 3) can be converted to input patterns to recognize 'possible' or 'regular' situations. These situations do not offer specific communicative reactions to the robot (and thus their development is a side task for the project) but may extend the parser accuracy in a large number of the recognized situations.

The extracted shallow semantic representations can be easily stored and processed in a database. They offer diverse types of requests for information retrieval and classification.

Lexical analysis. For a given lexeme (head of a noun phrase or head of a verb phrase) it is possible to retrieve its participation in different valencies with other surrounding actants. As shown in Table 5, the analysis of a word distribution across valencies reveals the semantic features and polysemy of the word—where a single notion with certain polysemy may be treated as an active agent, patient, cause, location, time, source/target, or a trajectory. This approach may be used to extend the semantics of a particular word: e. g. add markers for a possible actor, time, 2d or 3d-location, etc.

Table 5. Participation of a lexeme *snow* in different valencies

Valency	Examples
ag (agent)	**Snow** falls. **Snow** melts under black feet.
pat (patient)	John has been raking **snow** with bare hands. Their children cannot eat **snow**.
cont (content)	He simply didn't notice the **snow**. It resembled a fire-breathing **snow**.
instr (instrument)	Traces would be drifted with the **snow**. His hands were frozen but he immediately rubbed them with **snow**.
caus (cause)	Everything is white because of the **snow**. The wreckage could happen due to the **snow**.
foc (focus)	We were lucky with the **snow**. Their ingression with the **snow** to the water was the reason of the crimson color.
it (interpretation)	He was surprised: it was a sweet **snow**! The construction has melted and turned into a simple **snow**.
loc (location)	Do they sleep right on the **snow**? People saw crows on the **snow**.
t (time)	After the **snow** in his dreams he knew, everything would be fine in the morning.
src (source)	An arrogant exclamation came out from the **snow**.
targ (target)	They threw rifles to the **snow**. He fell from the stairs and dug his head into the **snow**.
tr (trajectory)	He managed to get through the spring **snow**. He dragged the case in the **snow**.

Analysis of a semantic/syntactic valency. For a given verb (head of a verb phrase) and its valency it is possible to study typical fillers—like *what do people drink?* (patient for a verb *drink*—see Table 6) or *who is usually a patient of violence?* (patient of any violence verb). This aggregation may serve as a basis for question answering—for questions to a specific participant of a situation, as well as to extend lexical semantics: e. g. add markers of a possible 'drink' to the fillers of the frame.

Table 6. Fillers of the *pat* valency for diverse 'drink' predicates

pat of 'drink' predicates	Number of cases
tea	144
water	78
wine	72
coffee	64
beer	51
what	49
vodka	40
blood	31
something	19
champagne	17

---

[2]   Wrong analysis of a clause *drink [is a] sin*.

pat of 'drink' predicates	Number of cases
milk	14
glass	13
bottle	11
drink	10
cognac	10
it	10
sin^2	10
cocktail	9
liquid	9

Search and aggregation of semantic frames. The annotation of the valencies with semantic markers makes it possible to search for all the facts (semantic predications) corresponding to a given pattern. For example, in the cases of extremism, where 'a person causes harm to someone', named after his/her 'nationality' or 'profession'—as in the following examples:

(1) *His girlfriend has beaten the driver with an unidentified object.*

(2) *Following the investigation records, the businessman has shot the director of the plant.*

(3) *Husband has also started to beat the policemen.*

The aggregation of specific examples, meeting the requested pattern may be further used to extend the accuracy of the input patterns. In this sense, the parser represented here is *frame*-centric: frames (input patterns of scripts) are used to select a tree and each word meaning, further, frames can be used to extend lexical semantics, new frames can be designed basing on the clustering of the extracted meanings, and the existing frames may be refined by the new examples falling within the same script.

References

1. *Blakar, R. M.* (1979), Language as a means of social power. In: Pragmalinguistics, J. Mey (ed.). The Hague-Paris, Mouton, pp. 131–169.
2. *Clore, G. L., & Ortony, A.* (2000), Cognition in Emotion: Always, Sometimes, or Never? In R. D. Lane & L. Nadel (Eds.), Cognitive Neuroscience of Emotion: Oxford Univ. Press, pp. 24–61.
3. *Fillmore, C. J.* (1968), The Case for Case. In Universals in linguistic theory. New York: Holt, Rinehart & Winston, pp. 1–68.
4. *Gladkiy, A. V.* (1985), Syntactic Structures of Natural Language in Computer-Aided Communication Systems [Sintaksicheskie struktury estestvennogo yazyka v avtomatizirovannyh sistemah obshcheniya], Nauka, Moscow.
5. *Golovinskaya, M. Ya.* (2004), Hidden Hyperbole as a Manifestation and Justification of Verbal Aggression. The Hidden Meanings: Word. Text. Culture [Skrytaya giperbola kak proyavlenie i opravdanie rechevoj agressii Sokrovennye smysly: Slovo. Tekst. Kul'tura], Yazyki slavyanskoj kul'tury, Moscow, pp. 69–76.

6. *Granovsky, D. V., Bocharov, V. V., Bichineva, S. V.* (2010), The OpenCorpora: Principles and Prospects [Otkrytyj korpus: principy raboty i perspektivy], Computational Linguistics and Development of Semantic Search on the Internet: Proceedings of the Scientific Workshop of the XIII All-Russian Joint Conference "Internet and modern society" [Komp'yuternaya lingvistika i razvitie semanticheskogo poiska v Internete: Trudy nauchnogo seminara XIII Vserossijskoj ob"edinennoj konferencii «Internet i sovremennoe obshchestvo»], St. Petersburg, p. 94.

7. *Kotov, A.* (2017), A computational model of consciousness for artificial emotional agents. Psychology in Russia: State of the Art, 10 (3), pp. 57–73.

8. *Kotov, A.* (2005), Application of Psychological Characteristics to D-Script Model for Emotional Speech Processing J. Tao, T. Tan, and R. W. Picard (Eds.) ACII 2005, LNCS 3784. Berlin, Heidelberg: Springer-Verlag, pp. 294–302.

9. *Kotov, A.* (2009), Accounting for irony and emotional oscillation in computer architectures Proceedings of International Conference on Affective Computing and Intelligent Interaction ACII 2009. Amsterdam: IEEE, pp. 506–511.

10. *Kotov, A.* (in print), Mechanisms of speech influence. M.: Kurchatov Institute.

11. *Kotov, A., Zinina, A., & Filatov, A.* (2015), Semantic Parser for Sentiment Analysis and the Emotional Computer Agents Proceedings of the AINL-ISMW FRUCT 2015 (pp. 167–170).

12. *Lyashevskaya O.* (2010), Bank of Russian constructions and valencies // LREC 2010. Malta, Valletta, May 19–21.

13. *Melčuk, I. A.* (1999), The Theory of Linguistic Models "Meaning ⇔ Text" [Opyt teorii lingvisticheskih modelej «SMYSL ⇔ TEKST»], Jazyki russkoj kultury, Moscow.

14. *Minsky, M. L.* (1988), The Society of Mind. New-York, London: Touchstone Book.

15. *Shvedova, N. Yu.* (1998) The Russian Semantic Dictionary [Russkij semanticheskij slovar'. Tolkovyj slovar', sistematizirovannyj po klassam slov i znachenij], Azbukovnik, Moscow.

16. *Yeh, W., & Barsalou, L. W.* (2006), The situated nature of concepts. American Journal of Psychology, 119(3), pp. 349–384.