

REVIEW

of dissertation for obtaining the educational and scientific degree *doctor* (Doctor of Philosophy, PhD)

Author of the dissertation: Hristina Kukova

Department of Computational Linguistics

Institute of Bulgarian Language "Prof. Lubomir Andreychin

Title of the dissertation: "Formal description of perceptual and mental predicates in Bulgarian (removal of semantic ambiguity)"

Reviewer: Prof. Yovka Tisheva, PhD, Sofia University "St. Kliment Ohridski"

1. Data for the author and for the doctoral dissertation

Hristina Kukova graduated from the bachelor's program "Russian Philology" at the Faculty of Slavic Philology at Sofia University "St. Kliment Ohridski" in 2001, and in 2002 – from the bachelor's program "Bulgarian Philology" at the same faculty. In the period 2003 - 2004 she studied in the master's program "Computational Linguistics. Internet technologies in the humanities. She expanded her training in the field of computational linguistics within the project "Integration of new practices and knowledge in the teaching of computational linguistics" (2012-2013). In the same period she worked on the Bulgarian National Corps. In 2010, Hr. Kukova was appointed a researcher in the Department of Computational Linguistics at Institute of Bulgarian language, and later continued her work as an assistant. She is currently a philologist in the same section and is engaged in the development and testing of language resources.

From March 1, 2004 to March 1, 2008 Hristina Kukova was enrolled as a doctoral student in the professional field 2.1. Philology, doctoral program "General and Comparative Linguistics (Mathematical Linguistics)" in the Department of Computational Linguistics of the Institute of Bulgarian Language "Prof. L. Andreychin". Her training is in regular form with supervisor Prof. Dr. Radka Vlahova. As of September 1, 2008, she was expelled with the right to defense. During his doctoral studies Hr. Kukova increases her scientific and professional training by participating in the courses organized by the Training Center of BAS, Curriculum for PhD students, as well as in specialized trainings on probabilistic models and formal semantics. By decision of the Scientific Council of IBE of December 2, 2021, the exams taken by the doctoral student during her studies are recognized in the procedure for the defense of the dissertation.

Hristina Kukova has taken part in a national project of the Institute of Bulgarian Language - "What is written remains! Write correctly!" She is part of the team of two ongoing research projects: "Ontology of situation situations - linguistic modeling. Comparative Study for Bulgarian and Russian" and "Enriching the Wordnet Semantic Network with Conceptual Frames".

The text presented under the procedure has a total volume of 379 pages, of which 280 pages are the main text. It contains 38 tables with summaries and 4 figures. Two applications are also included. The first is a table with synonymous sets (ID №, literals and definition) from Wordnet,

assigned a frame from Framenet and type according to the typology proposed in Chapter 4, and the second - a comparison of vocabulary meanings of verbs for perception in the Dictionary of Bulgarian and Wordnet. The list of references includes 147 titles.

2. Relevance of the problem developed in the dissertation

The dissertation of Hristina Kukova "Formal description of perceptual and mental predicates in Bulgarian (removal of semantic ambiguity)" is developed in a topical and important field of computational linguistics. Theoretically, the topic is a crossroads of research issues addressed in a number of cognitive sciences and linguistics, as it reflects the relationship between semantic and syntactic structures, between the mechanisms of generation and explication of lexical and syntactic meanings. The work is a comprehensive and in-depth study of the semantic features of one class of predicates in the Bulgarian language, which in certain contexts are realized with meanings relating to another lexical-semantic class. Ambiguity, also referred to as polysemy, homonymy, vagueness, syncretism or conversion, has been studied in the light of the processes that give rise to it, of specific classes of words in which this phenomenon is noted, e.g. in nouns, or at a certain linguistic level, e.g. morphological ambiguity. The work of Hr. Kukova is a comprehensive study of the relationship between the verbs for perception in Bulgarian and the semantic class of mental predicates.

Apart from the linguistic plan, the topicality of the issue is connected with the ongoing attempts in computational linguistics to develop models and algorithms for the automatic removal of semantic ambiguity. The automatic removal of semantic ambiguity is a significant problem in computer processing in natural language. Unlike the removal of grammatical ambiguity, which has been overcome to one degree or another for most European languages, resolving semantic ambiguity is a much more complex and still unresolved issue. The automatic removal of this type of ambiguity involves the selection of the appropriate meaning for a word from a set of predetermined possible meanings. In the Bulgarian language the predicates of the considered class are characterized by many lexical meanings, as well as by a great structural diversity in the syntactic realization. As the author notes, for the purposes of automatic removal of semantic ambiguity, all possible combinations should be described, i.e. syntactic constructions in which the considered predicates are realized with their basic meaning perception or as mental predicates.

The object of the study is clearly and precisely defined. The author sets three scientific goals: enlarging the previous descriptions and typology of the predicates for perception; comparative analysis of the argumentative and complementary characteristics of predicates with an emphasis on those representatives of the verbs of perception that could be found in both groups (perception and mentality) in order to support the automatic removal of semantic ambiguity; typological representation of the frames and frame elements of perceptual verbs. They are realistic, achievable and correspond to the topic of the dissertation. The nine scientific tasks derive from the main goal and accurately represent the stages that the study goes through. I believe that the first of the three hypotheses formulates reflects a common postulate about the relationship between the lexical semantics of verbs (not only those for perception) and their syntactic realization, proven in a number of previous studies.

3. Knowledge of the state of the problem and the scientific literature on the subject

The theoretical framework of the research was developed after the study of 147 bibliographic sources in Bulgarian, Russian and English. Hristina Kukova demonstrates a good knowledge concerning the general theoretical models for interpretation of the phenomenon of ambiguity, as well as of the lexical ambiguity (polysemy and homonymy). She skillfully interprets, summarizes and creatively applies in the analytical parts of the dissertation criteria for determining the phenomenon and classifications of the types of ambiguity proposed by J. Lakeff, J. Katz and J. Fodor, D. Cruz, J. Pustejowski et al. The study of N. Mihailova-Stalyanova is commented on the Bulgarian language, in which the theory of component analysis is applied. Hr. Kukova notes that she will consider "three types of ambiguity, namely ambiguity, homonymy (although less common in verbs) and metaphorical uses. Perceptual predicates demonstrate high productivity in terms of semantic enrichment through linguistic metaphor" (p. 42).

The doctoral student knows the theoretical foundations and practical experience related to the automatic removal of semantic ambiguity. At the heart of the process is determining the meaning with which the word is used in a given context. From a dictionary or resource representing the possible meanings attributed to the token, algorithms must be used to determine the meaning that has been updated in the particular case. In the second chapter of the dissertation Hr. Kukova reviews the different approaches to removing semantic ambiguity, noting that they differ in the type and amount of information being processed. "In earlier and fundamental scientific research in this direction, approaches use one or two types of information. Newer and upgrading approaches most often combine different types of information and statistics." (p. 48) The author comments on the results achieved through three groups of approaches: approaches using syntagmatic and statistical information (usually working on corpus data); approaches using conceptual and statistical information (taking into account the relationship between the topic / field of use of the particular word and the meaning); approaches using paradigmatic, syntagmatic and conceptual information (working with vocabulary definitions of words to detect the semantic similarity of words from definitions with words from the contextual environment). A variant of the latter is the Wordnet approach, which extends the subcategorizing information model.

Hristina Kukova's research is related to a specific group of predicates, which presupposes knowledge of the theoretical principles for linguistic classification of these units. Semantic classification of predicates for different languages. The author refers to the classical research of Z. Wendler, W. Cheif, J. Lions, B. Levin, R. Van Valin and R. LaPola, and in view of the Slavic languages - L. Shcherba, O. Selivyorstova, Y. Apresyan, E. Paducheva. Reviews of specific studies related to semantic classifications (division of semantic classes) or to the aspectual classifications of predicates (actions, processes, states or qualities) have done. The typological system developed within Framenet is mainly used to present the arguments to the analyzed predicates. It takes into account the semantic types of complements (objects, events and proposals according to the classification of S. Schule) and the structural and semantic characteristics of the frame elements with which they are expressed. In the fourth chapter of the dissertation an attempt is made to typologize the complements to the perceptual predicates in the Bulgarian language. Hr. Kukova defines the frame elements, considers the "intersection" between perceptual and mental predicates and determines the structural and semantic features of complements in the realization of different meanings. This is how the correspondences and discrepancies in the structural and semantic models that the perceptual and mental predicates

realize are derived. These conclusions allow specifying the approaches in Chapter Five in the typology of arguments and decisions to remove semantic ambiguity.

The focus of the review of research specifically related to perceptual predicates is not on the thematic classification by type of sense. In connection with the topic of the dissertation, Hristina Kukova commented mainly works that analyze the transformation from the specific sensory to the cognitive and perceptual predicates have epistemic, conceptual, indirect or abstract interpretation, which connects them with verbs for mental activity. In non-epistemic meanings, verbs are called pure perception, without the presence of a cognitive process, as opposed to epistemic meanings. In both cases the predicates are realized with different complements.

4. Research methodology

In the course of her research, Hristina Kukova applied a complex methodology, including general scientific logical methods (analysis, synthesis, summary, classification), empirical methods (source research, observation, evaluation, description and comparison, qualitative and quantitative content analysis). The methodology is adequate to the set scientific tasks. Its application leads to an accurate and consistent presentation of the models in the realization of non-epistemic and epistemic meanings of perceptual predicates and the choice of arguments for each type of meaning.

The methodology of the work includes several steps, leading from the empirical data to the typological generalizations. In the initial stage, language material from Wordnet (synonymous sets containing perceptual predicates) was excerpted, which was compared and expanded through data from the Dictionary of the Bulgarian Language. Systematization of the types of perceptual predicates according to four differential features is proposed: temporal localization; orientation (subject experiential, object stimulus); active or passive perception; actual or potential perception.

To determine the verbs where transition from the class of perceptual to that of mental predicates is possible, the author seeks the intersection of synonymous sets from the studied class with synonymous sets from other semantic classes according to Wordnet characteristics (eg for social relations and activities (verb. social), cognitive (verb.cognition), verbs for creation (verb.creation), etc.). Through analysis of the syntactic frames from Framenet and examples from the Bulgarian National Corpus, conclusions have been made about the syntactic realization of verbs depending on the specific meaning (for perception or for cognitive activity). Rules for the construction of verb phrases in different meanings and subcategorization are formulated. After analysis of the types of arguments, complements and complementators, a semantic and syntactic analysis of the frame elements (nuclear, peripheral and extrathematic) was made. The proposals for automatic removal of semantic ambiguity in view of the considered class of predicates are based on the typology of their complements.

5. The scientific and scientific-applied contributions of the dissertation. Significance of contributions to science and practice

The contributions to the work are clearly and accurately presented. They derive from the analysis of empirical data. The research hypotheses formulated at the beginning of the study were confirmed. The scientific goals of the research have been successfully met.

Theoretically, the knowledge of the semantic features and the typology of the predicates for perception in the Bulgarian language has been expanded. The comparative analysis of the syntactic realization of the verbs for perception, which could also be used with the meaning of mental predicates, proves the connection between the syntactic structure and the semantic characteristics of the support. These dependencies are systematized in order to support the automatic removal of semantic ambiguity in perception verbs. Changes in meaning are observed both in terms of naming other types of perceptions and in terms of naming cognitive activities or social relationships. Hr. Kukova separates prototypical and non-prototypical meanings of the considered verbs. Spontaneous perceptual verbs are most susceptible to semantic derivation, and this process is primarily observed in the verbs for visual and auditory perception. Most cognitive extensions can be attributed to verbs of visual perception. In the analysis of prototype uses, two groups of verbs emerge: for visual, auditory and tactile perceptions - for olfactory and taste perceptions. Complements to verbs from the first three groups can be introduced with *che* or *yes*, while in the other two they occur mainly with *che*-sentences. Verbs of the type *see*, *hear*, *feel* can be used with complements with cognitive content, while this is less typical of the other two types of verbs.

The non-prototypical meanings of the verbs for visual perception refer to cognitive activities, denoted by the mental predicates *understand*, *predict*, *imagine*, *judge*, *judge*, to the social sphere, e.g. *meet*, *visit*, etc., to an evaluation group, e.g. *discover*, *care for* and separate meanings as *witness* and *relate*. As non-prototype meanings of the verbs for auditory perception are defined: *paying attention*, *obeying*, *being informed* and *understanding*. Tactile predicates are characterized by the meanings *take (consume)*, *influence*, *reach*, *deal with*. The semantic extensions of olfactory verbs include the uses in which they realize meanings such as *trace*, *suppose*, *suspect*, and *investigate*. Verbs for taste perception can also realize meanings such as *experience*, *enjoy*, and *anticipate*.

The Hristina Kukova's work shows her achievements in solving a scientifically applied problem through methods corresponding to modern achievements of science and practice. The seven contributions mentioned in the dissertation can be summarized in two groups. A significant scientific-applied contribution is the presented original solution of a current problem related to the derivation of semantic features in the realization of the considered verbs. The dissertation expands the field of syntactic analysis, as it studies dependencies in the realization of polysemous words. Contributions in the field of syntactic descriptions of the Bulgarian language are the systematized models for presenting the semantic structure of the predicates for perception and the syntactic models in the realization of prototypical and non-prototypical meanings. These models are of practical value for computational linguistics, in the development of programs for automatic removal of semantic ambiguity and automatic translation, and for translation practice and for teaching Bulgarian as a native or foreign language.

6. Evaluation of the publications on the dissertation and the abstract

Hristina Kukova presented a list of eight publications on the topic of the dissertation. One of them is in English and the others in Bulgarian. Two of the articles are in publications referenced and indexed in world-famous databases of scientific information. Two articles have been published in *Bulgarian Language* periodical, and three - in collections of papers from scientific forums. The quantitative and qualitative evaluation of the presented publications shows that

they meet the minimum requirements for obtaining the educational and scientific degree "Doctor" (PhD).

The abstract of the dissertation accurately and clearly presents the structure of the work, the observations and analyzes made and their results, as well as the scientific contributions of the dissertation.

7. Conclusion

The text presented by Hristina Kukova meets the formal and substantive requirements for a dissertation. This is a comprehensive in-depth study of a current scientific problem, as a result of which the author reaches important conclusions in the field of syntax and formal description of Bulgarian language, lexical semantics, and computational linguistics.

Assessing the relevance of the topic of the dissertation "Formal description of perceptual and mental predicates in the Bulgarian language (removal of semantic ambiguity)", the results achieved and the conclusions drawn, I propose to the esteemed Scientific Jury to award Hristina Kukova educational and scientific degree "Doctor" (PhD) in in the field of higher education 2. Humanities, professional field 2.1. Philology, scientific specialty "General and Comparative Linguistics".

15.05.2022

Prof. Yovka Tisheva