

## REVIEW

under the competition for the academic position of Associate Professor  
in field 2.1. *Philology*, scientific specialty *General and Comparative Linguistics*  
as promulgated in State Gazette No 102/23.12.2022  
with the following candidate: Chief Assistant Tsvetana Ivanova Dimitrova, PhD,  
Computational Linguistics Section, Prof. L. Andreychin Institute of Bulgarian Language

Reviewer: Prof. Petya Nacheva Ossenova, PhD, Saint Kliment Ohridski University of Sofia

The set of materials submitted by the candidate includes all the necessary documents required under the procedure. Chief Assistant Tsvetana Dimitrova, PhD thus fulfils the minimum national requirements pursuant to Article 2b(2) and (3) of the Development of Academic Staff in the Republic of Bulgaria Act (DASRBA).

Ms. Dimitrova participates in the competition for the academic position of Associate Professor with the following publications: one monograph and a total of 22 articles and studies in two fields. These fields are as follows: *Language resources, computational language modelling and Historical corpora, description of historical language data*. For the first field, the candidate has submitted a total of 14 publications, all but one of which are co-authored; three are studies and six have been printed in conference proceedings abroad; six publications are indexed in Scopus and two in Web of Science. The second field is presented by a total of 8 publications, of which four are co-authored; two are studies; the majority are journal publications; two are indexed in Scopus. There are 125 citations for publications between 2008 and 2020, many of them in proceedings and publications abroad.

As the publications demonstrate, Ms. Dimitrova works in the field of language technologies and resources as well as in the field of historical linguistics. It is no coincidence that the candidate is a member of international organizations such as the Commission on Computer Supported Processing of Mediaeval Slavonic Manuscripts and Early Printed Books with the International Committee of Slavists; the Slavic Linguistic Society; and the Global WordNet Association. Ms. . Dimitrova is involved in a number of national and international projects. Their topics include the creation and work on resources for the Bulgarian language, including Old Bulgarian (word network, national corpus, image corpus, etc.); comparative studies of phenomena in Bulgarian and other Slavic languages; applied aspects of language technologies (machine translation, automatic name recognition, etc.). I must also mention the fact that in 2011 Ms. Tsvetana Dimitrova received the Professor Marin Drinov Young Scientist Award in

the field of humanities. The candidate's teaching activity includes a course in corpus linguistics with the Master's Programme in *Computational Linguistics. Internet Technologies in the Humanities* of the Faculty of Slavonic Philologies at the Saint Kliment Ohridski University of Sofia.

When evaluating the publications submitted for the competition, I will first briefly focus on the articles and studies, and then on the monograph as the main work. In the first area - *Language Resources, Computer Modelling of Language* - Ms. Dimitrova has submitted mainly co-authored publications. Since the field of language technology predominantly involves teamwork, I believe that the candidate has managed to fit into a particular team and to contribute towards the creation and enrichment of a number of language resources for our language. Ms. Dimitrova has made the following contributions: Description of the WordNet system (publications 3, 4, 5); the study of a number of linguistic phenomena in WordNet such as derivational relations (publication 2), the classification of adjectives (publications 6, 7, 8), the addition of (morpho)semantic relations to verbs and nouns in WordNet (publications 9, 10, 11, 12); the description of the specifics of the Bulgarian-English corpus (publication 13); the description of the specifics of names for persons, geographical objects and organizations in Bulgarian (publication 14); the comparative examination of predicative constructions in Bulgarian and Russian from a corpus point of view (publication 15).

In the second field - *Historical corpora, Description of Historical Linguistic Data* - Ms. Dimitrova has made the following contributions: a thorough description of the behaviour of clitics in historical texts through corpus analysis, including a comparison with modern Bulgarian (publications 16, 17, 19, 20); a description of the development of the future tense in diachronic terms (publication 18); a corpus study of the conjunction *although* (even though, etc.) (publication 21); problems in the segmentation and annotation of old texts (publications 22, 23).

The monograph submitted for the competition is entitled *Bulgarian Clitics: History and Present*. Sofi, Prof. Lyubomir Andreychin Institute for the Bulgarian language with the Bulgarian Academy of Sciences, 2023, 256 pages, ISBN 978-619-908-845-6.

The topic of clitics is extremely interesting because, on the one hand, it is much discussed in Bulgarian and foreign linguistics, but on the other hand, it is not an easy topic as an object of linguistic modelling. Ms. Dimitrova focuses on the modelling of clitics in the process of their diachronic development, which I consider a major contribution of the book.

The monograph consists of a preface, three chapters, a conclusion, and an appendix with quantitative data on the examples given. The preface presents the object of the study together



with the scope of the reviewed linguistic material, the data used, the basic concepts, and an overview of research on clitics in Bulgarian. *Chapter One* describes the nature and features of clitics, and this description is made in relation to two 19th century linguistic laws. These are Wackernagel's Law and Tobler-Mussafia Law. The author's observations are verified through examples from different corpora. *Chapter Two* presents clitics in their diachronic development. Here Ms. Dimitrova works with sources (which she has described in detail) from three periods - Old Bulgarian, Middle Bulgarian, and Early New Bulgarian. Three groups of elements are analysed: discourse clitics, pronominal clitics (together with nuanced semiclitics and quasiclitics) and verbal clitics. *Chapter Three* proposes a model for the variation in the development of clitics according to phase theory within Chomsky's minimalist syntax. Contributive findings are that: the syntactic positions of clitics in the sentence are characterized by relative stability over time; changes in the composition of elements around clitics lead to the dropping of some of them; the composition of pronominal and verbal clitics is expanded; variability is due to changes in some functional elements, such as negation, and to changes in verb forms.

As a member of the jury, I consider that the procedure has complied with all the legal requirements, namely: DASRBA, the Regulations on the Implementation of DASRBA, of the Regulations for the Conditions and Procedure for the Acquisition of Scientific Degrees and for the Occupation of Academic Positions at the Bulgarian Academy of Sciences, of the Regulations for the Conditions and Procedure for the Acquisition of Scientific Degrees and for the Occupation of Academic Positions at the Prof. Lyubomir Andreychin Institute for the Bulgarian Language.

Following review of the materials and scientific works of the candidate, as well as analysis of their significance and contributions, I confidently give my positive assessment and recommend to the Scientific Jury to elect Chief Assistant Tsvetana Dimitrova to the academic position of *Associate Professor* in field 2.1. *Philology*, scientific specialty *General and Comparative Linguistics* for the needs of the Computational Linguistics Section at the Prof. Lyubomir Andreychin Institute for the Bulgarian Language.

Date: ....  
(signature)

Prepared by: .....