



Department of Computational Linguistics
Institute for Bulgarian Language
Bulgarian Academy of Sciences

Proceedings of the Third International Conference

**Computational
Linguistics in
Bulgaria**

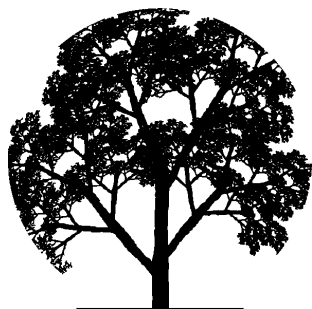


**27-29 May 2018
Sofia, Bulgaria**

The Third International Conference *Computational Linguistics in Bulgaria (CLIB 2018)* is organised with the support of the National Science Fund of the Republic of Bulgaria under the project *Towards a Semantic Network Enriched with a Variety of Relations*, Grant Agreement DN10/3/2016.




CLIB 2018 is organised by:



Department of
Computational Linguistics
Institute for Bulgarian Language
Bulgarian Academy of Sciences

PUBLICATION AND CATALOGUING INFORMATION

Title:	Proceedings of the Third International Conference <i>Computational Linguistics in Bulgaria (CLIB 2018)</i>
ISSN:	2367 5675 (online)
Published and distributed by:	The Institute for Bulgarian Language Bulgarian Academy of Sciences
Editorial address:	Institute for Bulgarian Language Bulgarian Academy of Sciences 52 Shipchenski Prohod Blvd., Bldg. 17 Sofia 1113, Bulgaria +3592/ 872 23 02
Copyright:	Copyright of each paper stays with the respective authors. The works in the Proceedings are licensed under a Creative Commons Attribution 4.0 International Licence (CC BY 4.0).  License details: http://creativecommons.org/licenses/by/4.0 Copyright © 2018

Proceedings of the
Third International Conference
*Computational Linguistics in
Bulgaria*



Sofia, Bulgaria
28-29 May 2018

PROGRAMME COMMITTEE

Chair:

Svetla Koeva – Institute for Bulgarian Language (BAS)

Co-chairs:

Mila Dimitrova-Vulchanova – Norwegian University of Science and Technology

Cvetana Krstev – University of Belgrade

Tania Avgustinova – Saarland University

Iana Atanassova – Centre Tesnière, Université de Franche-Comté, UFR SLHS

Verginica Barbu Mititelu – Research Institute for Artificial Intelligence, Romanian Academy

Mariana Damova – Mozaika, Bulgaria

Ivan Derzhanski – Institute of Mathematics and Informatics (BAS)

Radovan Garabík – L’udovít Štúr Institute of Linguistics, Slovak Academy of Sciences

Stefan Gerdjikov – Sofia University

Filip Ginter – University of Turku

Kjetil Rå Hauge – University of Oslo

Zornitsa Kozareva – Google

Ivan Koychev – Sofia University

Hristo Krushkov – Plovdiv University

Eric Laporte – Université Paris-Est Marne-la-Vallée

Denis Maurel – University of Tours

Stoyan Mihov – Institute of Information and Communication Technologies

Ruslan Mitkov – University of Wolverhampton

Preslav Nakov – Qatar Computing Research Institute

Karel Oliva – Institute of the Czech Language, Academy of Sciences of the Czech Republic

Maciej Ogrodniczuk – Institute of Computer Science, Polish Academy of Sciences

Maciej Piasecki – Wrocław University of Technology

Dragomir Radev – Yale University

Marko Tadić – University of Zagreb

Hristo Tanev – Joint Research Centre of the European Commission

Tinko Tincev – Sofia University

Dan Tufis – Research Institute for Artificial Intelligence, Romanian Academy

Duško Vitas – University of Belgrade

Radka Vlahova – Sofia University

Victoria Yaneva – University of Wolverhampton

ORGANISING COMMITTEE

Svetlozara Leseva – Institute for Bulgarian Language (BAS)

Iana Atanassova – Centre Tesnière, Université de Franche-Comté, UFR SLHS

Rositsa Dekova – Plovdiv University, Faculty of Philology, Department of English Studies

Tsvetana Dimitrova – Institute for Bulgarian Language (BAS)

Dimitar Hristov – Institute for Bulgarian Language (BAS)

Alexander Popov – Institute of Information and Communication Technologies

Borislav Rizov – Institute for Bulgarian Language (BAS)

Katya Saint-Amand – Senior Linguist, Chenope

Valentina Stefanova – Institute for Bulgarian Language (BAS)

Ivelina Stoyanova – Institute for Bulgarian Language (BAS)

Ekaterina Tarpomanova – Sofia University, Faculty of Slavic Studies

Maria Todorova – Institute for Bulgarian Language (BAS)

Martin Yalamov – Institute for Bulgarian Language (BAS)

Victoria Yaneva – University of Wolverhampton

PLENARY TALKS

LINGUISTIC INTELLIGENCE: COMPUTERS VS. HUMANS

Prof. Ruslan Mitkov (University of Wolverhampton)

Computers are ubiquitous – they are and are used everywhere. But how good are computers at understanding and producing natural languages (e.g. English or Bulgarian)? In other words, what is the level of their linguistic intelligence? This presentation will examine the linguistic intelligence of the computers and will look at the challenges ahead...

I shall begin by a brief historical flashback. I shall plot the timeline of the linguistic intelligence of computers against that of humans. Natural Language Processing (NLP) advances in the last 20 years have made it possible for the linguistic intelligence of computers to increase significantly but they are still behind humans...

The presentation will explain why it is so difficult for computers to understand, generate and in general, to process natural language texts – it is a steep road/learning curve, it is long and winding road for both computers and researchers who seek to develop intelligent programs. The talk will also briefly present well-established NLP techniques computers follow when 'learning' to 'speak' our languages, including rule-based and knowledge-based methods initially and machine learning and deep learning methods more recently, the latter being regarded as highly promising. A selection of Natural Language Processing applications will be outlined next. Finally, a preview will be offered of selected slides from my plenary talk at CLIB'2018 (which will be given on the following day).

I am not a clairvoyant, but at some of my plenary talks I have been asked to predict how far will computers go... At the end of my presentation in Sofia I shall share with you what I predict for the future and in general, what my vision is.

WITH A LITTLE HELP FROM NLP: MY LANGUAGE TECHNOLOGY APPLICATIONS WITH IMPACT ON SOCIETY

Prof. Ruslan Mitkov (University of Wolverhampton)

The talk will present three original methodologies developed by the speaker, underpinning implemented Language Technology tools which are already having an impact on the following areas of society: e-learning, translation and interpreting and care for people with language disabilities.

The first part of the presentation will introduce an original methodology and tool for generating multiple-choice tests from electronic textbooks. The application draws on a variety of Natural Language Processing (NLP) techniques which include term extraction, semantic computing and sentence transformation. The presentation will include an evaluation of the tool which demonstrates that generation of multiple-choice tests items with the help of this tool is almost four times faster than manual construction and the quality of the test items is not compromised. This application benefits e-learning users (both teachers and students) and is an example of how NLP can have a positive societal impact, in which the speaker passionately believes.

The talk will go on to outline two other original recent projects which are also related to the application of NLP beyond academia. First, a project, whose objective is to develop next-generation translation memory tools for translators and, in the near future, for interpreters, will be briefly presented. Finally, an original methodology and system will be outlined which helps users with autism to read and better understand texts.

BUILDING CONVERSATIONAL ASSISTANTS USING DEEP LEARNING

Dr Zornitsa Kozareva (Google)

Over the years there has been a paradigm shift in how humans interact with machines. Today's users are no longer satisfied with seeing a list of relevant web pages, instead they want to complete tasks and take actions. This raises the questions: "How do we teach machines to become useful in a human-centered environment?" and "How do we build machines that help us organize our daily schedules, arrange our travel and be aware of our preferences and habits?". In this talk, I will describe these challenges in the context of conversational assistants. Then, I will delve into deep learning algorithms for entity extraction, user intent prediction and question answering. Finally, I will highlight findings on user intent prediction from shopping, movies, restaurant and sport domains.

NEURAL GRAPH LEARNING

Dr Sujith Ravi (Google)

Recent machine learning advances have enabled us to build intelligent systems that understand semantics from speech, natural language text and images. While great progress has been made in many AI fields, building scalable intelligent systems from "scratch" still remains a daunting challenge for many applications. To overcome this, we exploit the power of graph algorithms since they offer a simple elegant way to express different types of relationships observed in data and can concisely encode structure underlying a problem. In this talk I will focus on "How can we combine the flexibility of graphs with the power of machine learning?"

I will describe how we address these challenges and design efficient algorithms by employing graph-based machine learning as a computing mechanism to solve real-world prediction tasks. Our graph-based machine learning framework can operate at large scale and easily handle massive graphs (containing billions of vertices and trillions of edges) and make predictions over billions of output labels while achieving $O(1)$ space complexity per vertex. In particular, we combine graph learning with deep neural networks to power a number of machine intelligence applications, including Smart Reply, image recognition and video summarization to tackle complex language understanding and computer vision problems. I will also introduce some of our latest research and share results on "neural graph learning", a new joint optimization framework for combining graph learning with deep neural network models.

Table of Contents

Ruslan Mitkov <i>With a little help from NLP: My Language Technology applications with impact on society</i>	1
Vito Pirrelli <i>NLP-based Assessment of Reading Efficiency in Early Grade Children</i>	5
Mila Vulchanova and Valentin Vulchanov <i>Figurative language processing: A developmental and NLP Perspective</i>	7
Nikola Taushanov, Ivan Koychev and Preslav Nakov <i>Abstractive Text Summarization with Application to Bulgarian News Articles</i>	15
Maria Gritz <i>Towards Lexical Meaning Formal Representation by virtue of the NL-DL Definition Transformation Method</i>	23
Junya Morita <i>Narrow Productivity, Competition, and Blocking in Word Formation</i>	34
Cvetana Krstev, Ranka Stanković and Duško Vitas <i>Knowledge and Rule-Based Diacritic Restoration in Serbian</i>	41
Anton Zinoviev <i>Perfect Bulgarian Hyphenation, or How not to Stutter at End-of-line</i>	52
Anna Roitberg and Denis Khachko <i>Russian Bridging Anaphora Corpus</i>	62
Ekaterina Tarpomanova <i>Aspectual and Temporal Characteristics of the Past Active Participles in Bulgarian – a Corpus-based Study</i>	69
Olena Siruk and Ivan Derzhanski <i>Unmatched Femininitives in a Corpus of Bulgarian and Ukrainian Parallel Texts</i>	77
Viktoriya Petrova <i>The Bulgarian Summaries Corpus</i>	85
Natalia Loukachevitch and Boris Dobrov <i>Ontologies for Natural Language Processing: Case of Russian</i>	93
Ranka Stanković, Miljana Mladenović, Ivan Obradović, Marko Vitas and Cvetana Krstev <i>Resource-based WordNet Augmentation and Enrichment</i>	104
Svetlozara Leseva, Ivelina Stoyanova and Maria Todorova <i>Classifying Verbs in WordNet by Harnessing Semantic Resources</i>	115

Maria Mitrofan, Verginica Barbu Mititelu and Grigorina Mitrofan <i>A Pilot Study for Enriching the Romanian WordNet with Medical Terms</i>	126
Ivelina Stoyanova <i>Factors and Features Determining the Inheritance of Semantic Primes between Verbs and Nouns within WordNet</i>	135
Borislav Rizov and Tsvetana Dimitrova <i>Online Editor for WordNets</i>	146
Amir Bakarov <i>The Effect of Unobserved Word-Context Co-occurrences on a Vector-Mixture Approach for Compositional Distributional Semantics</i>	153
Rositsa Dekova and Adelina Radeva <i>Introducing Computational Linguistics and NLP to High School Students</i>	162
Ivan Derzhanski and Milena Veneva <i>Linguistic Problems on Number Names</i>	169
Marina Dzhonova, Kjetil Rå Hauge and Yovka Tisheva <i>Parallel Web Display of Transcribed Spoken Bulgarian with its Normalised Version and an Indexed List of Lemmas</i>	177
Georgi Dzhumayov <i>Integrating Crowdsourcing in Language Learning</i>	185
Todor Lazarov <i>Bulgarian–English Parallel Corpus for the Purposes of Creating Statistical Translation Model of the Verb Forms. General Conception, Structure, Resources and Annotation</i>	193
Branislava Šandrih <i>Fingerprints in SMS messages: Automatic Recognition of a Short Message Sender Using Gradient Boosting</i>	203

Aspectual and temporal characteristics of the past active participles in Bulgarian – a corpus-based study

Ekaterina Tarpomanova
University of Sofia St. Kliment Ohridski
tochitsaaa@gmail.com

Abstract

The paper presents a corpus-based study of the past active participles in Bulgarian with respect of their aspectual and temporal characteristics. As this type of participles combine two morphological markers, a special attention is paid on their interaction in different tenses, moods and evidentials. The source of language material used for the study is the Bulgarian National Corpus. The paper is organized in terms of morphological oppositions, aspectual and temporal, analyzing the functions of the participles in compound verbal forms.

1. Introduction

In the modern Bulgarian there are five types of participles: present active, aorist active, imperfect active, past passive, and gerund. Being a verbal form, participles can be marked for tense, aspect and voice, but they also may share some of the categories of the adjective as gender, number and definiteness. However, their hybrid nature does not allow for the complete manifestation of the grammatical categories and especially with regard to the verbal categories participles are only partially marked with the respective grammatical meanings (GSBKE: 374).

The paper aims at studying the functions of the aorist and imperfect active participles by using the empirical data of the Bulgarian National Corpus. The aorist active participle is an old form that is found in all the Slavic languages. It is considered a formally, semantically and functionally stable form both in dialectal and standard varieties of Bulgarian. The imperfect active participle is an innovation in Bulgarian and a number of researchers share the opinion that its formation is connected to the grammaticalisation of the category of evidentiality. The study is organised in terms of morphological oppositions – aspectual and temporal, based on the respective characteristics of the participles. In such situations the speaker has to make a choice between morphologically marked forms according to his/her communicative intentions and the context that can enhance or restrict the usage of a certain form.

2. Research method

The Bulgarian National Corpus (BulNC) is used as a source of empiric language material being the largest electronic resource for Bulgarian (its monolingual part contains over 1,2 billion words). BulNC has been designed mainly for computational linguistic tasks focusing on volume and structure. Although representativeness and balance are not considered key features, the corpus covers the language production since 1945 up to now and the language varieties of different text types. The online search system and the linguistic annotation make it suitable for linguistic research too. For detailed description of BulNC, see Koeva et al. 2012.

Participles in BulNC are annotated as deverbal forms with several grammatical characteristics, for example *npaβen* {V PE T s q}: V = verb, PE = perfective, T = transitive, s = singular, q = imperfect participle. Theoretically the combination of two features – aspect (imperfective or perfective) and type of participle (past aorist or past imperfect) should provide all the grammatical information through the respective morphological markers for a correct annotation, but in fact there are many instances of

Keywords: past active participles, Bulgarian, corpus-based study

incorrect interpretation, especially concerning specific forms such as imperfect participles derived from perfective verbs. For that reason I chose three imperfective verbs representative for the three conjugations in Bulgarian, and their perfective counterparts¹: *пиша, напиша* ‘write’; *правя, направя* ‘do; make’; *казвам, кажа* ‘say’. The verbs are frequent and semantically neutral, so that the results of a search by word form allow for general conclusions about the types of participles under study.

3. Past active participles in Bulgarian: formation, meaning and usage

The aorist past participle is formed by adding the suffix *-l* to the aorist stem of an imperfective or a perfective verb:

пиша – *писал* ‘write, impf.’
напиша – *написал* ‘write, pf.’
правя – *правил* ‘do; make, impf.’
направя – *направил* ‘do; make, pf.’
казвам – *казвал* ‘say, impf.’
кажа – *казал* ‘say, pf.’

The aorist past participle denotes property of an action that is performed before a given interval of reference (GSBKE: 379; Nitsolova 2008: 434). It is used in the resultative tenses of indicative (perfectum, plusquamperfectum, futurum exactum, futurum exactum praeteriti), in the structure of indirect evidentials (renarrative, inferential and dubitative) and admirative, and in the Slavic type of the conditional mood.

The imperfect past participle is formed from the imperfect stem of an imperfective or a perfective verb and the suffix *-l*:

пиша – *пишел* ‘write, impf.’
напиша – *напишел* ‘write, pf.’
правя – *правел* ‘do; make, impf.’
направя – *направел* ‘do; make, pf.’
казвам – *казвал* ‘say, impf.’
кажа – *кажел* ‘say, pf.’

According to Nitsolova, the imperfect active participle denotes an action whose interval is larger than a present or a past interval of reference (Nitsolova 2008: 436). It can be used only in indirect evidential forms (renarrative, inferential or dubitative) and in admirative. Unlike the aorist participle, it cannot function as an adjective.

The 3rd conjugation verbs have only one stem for all the tenses, i.e. the present stem, and for that reason the aorist and the imperfect participles are homonymous.

Past active participles in Bulgarian are organised in a complicated system with two morphological markers: for aspect (imperfective vs. perfective) and for tense (aorist vs. imperfect). Their functioning can be analysed in terms of two oppositions: aspectual and temporal (as they are all active, the opposition by voice is not relevant).

4. Aspectual oppositions

4.1. Indicative

Perfect and pluperfect

писал vs. *написал*, *правил* vs. *направил*, *казвал* vs. *казал*

Participles display the common characteristics of the respective aspect, i.e. participles of imperfective verbs present the event as atelic, more often iterative, non-concrete (general) or processual², while the participles derived from perfective verbs view the event as telic, usually single and/or concrete. The examples of BulNC show that there are several typical contexts of each type of participle.

¹ The prevalent opinion for the aspectual oppositions in Bulgarian is that a basic imperfective verb (*пиша*) and a prefixed perfective verb (*напиша*) do not form an aspectual pair, but for the purpose of this study verbs are selected for their frequency and variety of forms.

² For the concrete aspectual meanings I use the classification of Valentin Stankov (Stankov 1980).

Imperfective: iterativity

The imperfective verbs and participles respectively are typically used for unbounded iterativity and habituality, while the perfective is associated with bounded iterativity. Iterativity is often enhanced by adverbials of the type ‘many times’, and habituality by adverbs and adverbials with the ‘always’.

(1) *Всичко това той го е казвал и преди безброй пъти.*

‘He has said that before, thousands of times.’

(2) *Да, тъкмо това беше правил винаги – носеше се по пързалката на течението.*

‘Yes, he had always done this – he was drifting on the stream.’

Perfective: bounded iterativity

In Bulgarian bounded iterativity is regularly expressed by perfective verbs, usually in a lexical context specifying the number of times the event is repeated. With respect to the system of participles, bounded iterations are connected with the aorist participle. Still, in the structure of the perfect tense this is not a central meaning of this type of participle. A possible explanation is that the bounded iterativity combines better with aorist than with perfect because the event is presented as localized in a past moment, which contradicts to the main meaning of the perfect. Another restricting factor is the extension of the scope of the inferential in the field of the perfect, especially in 2nd and 3rd person.

The language data in BulNC illustrate the clear preference for aorist instead of perfect with aorist participle: 16 instances of aorist vs. 1 instance of perfect for the verb *кажа* ‘say, pf.’ in 1st p. sg. In the examples below the usage of perfect in (4) should be interpreted as emphatic.

(3) *Три пъти казах “добър вечер”.*

‘I said “good evening” three times.’

(4) *Хиляди пъти съм казал, че ненавиждам боя...*

‘I have said thousands of times that I hate fight.’

Imperfective: general factuality

The imperfective participle is used when the event is viewed as a general fact, without any specifications of its properties (Stankov 1980). This is one of the typical meanings of the imperfective aspect, but it is also strongly connected with the perfect tense. A very frequent lexical context in interrogative sentences are adverbials with the meaning ‘ever’.

(5) *Писмото звучеше сякаш го бе писал той.*

‘The letter sounded as if he had written it.’

(6) *Да съм казвал някога, че планът е съвършен?*

‘Have I ever said that the plan was perfect?’

General factuality is often expressed in negative context, and in such cases it can be enhanced by adverbials ‘never’, ‘at all’, etc.

(7) *Аз например никога не съм писал нещо криминално.*

‘As for me, I have never written detective stories.’

(8) *Никога по-рано не съм правил това!*

‘I have never done this before!’

Perfective: concrete factuality

According to Stankov (1980), the concrete factual meaning is the central particular meaning of the perfective that expresses a single complete event stated as a fact in the concrete circumstances of its realization. Among the past tenses it is connected mostly with the aorist, denoting a concrete and a completed event in the past, but it is compatible with the perfect too. As compared to the aorist, the perfect meaning can be more expressive or to put an emphasis on the event. In a sentence with a perfective verb its arguments describe explicitly the situation of the event realization.

(9) *Егон, не си го измислям. – Не съм казал това.*

‘Egon, I’m not making it up. – I didn’t say that.’

(10) *Направил съм това проследяващо устройство.*

‘I made this tracking device.’

Imperfective: process

Processuality is a central meaning of the imperfective aspect. To express a process, participles of imperfective verbs are more often used in pluperfect with a taxis function.

(11) *Разбира се, знаеше всичко това, докато беше писал текста.*

‘Of course, he knew all that while he was writing the text.’

Imperfective instead of perfective

A perfective reading of basic imperfective verbs (non-prefixed and non-suffixed) is inherited by the respective participles.

(12) *Казвай де, какво ти е писал?*

‘Come on, tell me what he wrote you.’

Futurum exactum and futurum exactum praeteriti

Due to their meaning both tenses more often comprise in their structure perfective participles. FE and FEP refer to an event whose result is situated before the completion of another event (FEP is mostly used in conditional sentences). The fact that the second event is completed generally implies the completion of the first event too, that is why those two tenses usually choose participles of perfective verbs.

(13) ... *тридесет минути след приемането на химикала, ще е казал на търтеите всичко, което Елиът иска да знае.*

‘... thirty minutes after consuming the substance, he will have told the drones everything Eliot wanted to know.’

(14) *Станеше ли то, за няколко месеца щеше да е направил кариера и то каква!*

‘If this happened, for a few months he would have made a career, and a great one!’

The combination *ще е* + participle of imperfective verb has usually a presumptive reading. In fact, all the examples of such combination found in BulNC are presumptives (130 results):

(15) *Някога, на младини, тя ще е била стройна и хубава.*

‘When she was young, she must have been slender and beautiful.’

(16) *Когато най-сетне се приготви, беше невероятен – така ще е блестял Харун ал Рашид на първата си сватба.*

‘When he finally got ready, he looked amazing – Harun al-Rashid must have shined like this at his first wedding.’

(17) *От това може да се съди, че в Букурещ тя ще е преболедувала доста сериозно.*

‘One may conclude that in Bucharest she must have been very sick.’

A few exceptions are found, for biaspectual verbs with perfective interpretation:

(18) *След около два часа вашата нервна система ще е асимилирала вече Зеко.*

‘In a couple of hours your nervous system will have assimilated Zeko.’

(19) *На практика, населението ще е гласувало за умерено, традиционно или поне реформистко правителство, а ще се установи режим на твърдата левица...*

‘In actual fact, the people will have voted for a moderate, traditionalist or at least reformist government, but a regime of the hard left will establish itself.’

However, out of the corpus examples are found in which FE form with an imperfective participle is used to express a general fact in the future situated before a future event. Therefore, despite of the corpus data, the usage of imperfective participle in the structure of FE and FEP is possible, although limited in terms of frequency.

(20) *Просто ще е правил секс, а стеснителността му ще си остане...*

‘He will just have had sex, but his shyness will remain the same.’

4.2. Conditional

писал vs. написал, правил vs. направил, казвал vs. казал

Aspectual opposition between past active participles is present in the Slavic type of the conditional mood formed by the auxiliary *бих* and the aorist participle of an imperfective or perfective verb. As the conditional forms are unambiguous, some statistical data may be obtained by a searching by word forms (2 and 3 p. sg. of the selected verbs).

imperfective	би писал	31	би правил	78	би казвал	12
perfective	би написал	59	би направил	141 2	би казал	1290

Table 1: Instances of the imperfective and perfective aorist participles in conditional.

The results presented in Table 1 show a clear predominance of the perfective verbs in conditional, except for the verbs *пиша* / *напиша* 'write' with a ratio of only 1:2 between imperfective and perfective. The conclusion is that conditionals combine better with telic events, while atelic are peripheral. This observation is apparent for the aspectual pair *казвам* / *кажа* 'say' where the imperfective verb is suffixed and cannot be used with a perfective meaning, unlike *пиша* 'write' and *правя* 'do; make'.

Taking into account the usage of the selected verbs, the most frequent concrete meaning of the imperfective participles in conditional forms is general factual (21), while iterative (22), habitual (23) and processual (24) meanings are occasional. A very frequent context for the imperfective participles is a *what*-question – 41 instances of the 78 occurrences of the form *би правил* 'would do'.

(21) *Съвсем други скокове би правил този тигър на свобода.*

'This tiger would make quite different jumps if he was free.'

(22) *Ако не чакаха някаква облага, и говорещият истината би лъгал колкото лъжеца, и лъжецът би казвал истината, колкото нелъжещия.*

'If they did not expect some benefit, the truth teller would lie as much as the liar and the liar would tell the truth as often as the truth teller.'

(23) *Какво би правил обикновено? – Нищо особено.*

'What would he usually do? – Nothing special.'

(24) *Днес някой спяха от Аинтаб би казвал: ...*

'Today, some spahi from Aintab would say: ...'

Perfective conditional forms refer to a concrete event.

(25) *Така би направил един обикновен гражданин.*

'That's what a common citizen would do.'

4.3. Evidentials

Evidential present and imperfect

пишел vs. *напишел*, *правел* vs. *направел*, *казвал* vs. *кажел*

The evidential present and imperfect formed with imperfective participle display the characteristics of the respective tenses of indicative. The most frequent aspectual meanings associated with these tenses are the following: processual (26), iterative (27), habitual (28) and general factual (29). The examples below illustrate the usage of the participles in renarrative.

(26) *Но Зайо Байо не **правел** нищо особено.*

'But the Rabbit wasn't doing anything special.'

(27) *Кажете им, че ей сега тръгвам – **казвал** той на пратениците, а това "ей сега" нямаше край.*

'Tell them that I'm leaving right away – he used to say every time to the messengers, and this right away was endless.'

(28) *Ахав, великият миротворец на Вискос, често **казвал**: ...*

'Ahav, the great peacemaker of Viskos, used to say: ...'

(29) *Не **правела** така.*

'She never does that, she said.'

The imperfect participle of perfective verbs can be used in dependent clauses only, or in imperative and optative clauses, which corresponds to the usage of the perfective verbs in indicative. The dependent clauses are more often introduced by the conjunction *да* 'to', other subordinating conjunctions (*за да* 'in order to'), relative pronouns and adverbs.

(30) *Който **кажел** една нова истина, вдигали му паметник.*

'Whoever told a new truth, they raised him a monument.'

(31) *... **треста** пъти да **кажел** „Отче наш“ и **треста** пъти „Аве Мария“.*

'He had to say three hundred times the Lord's Prayer and three hundred times Ave Maria.'

(32) *Бащата запази Анри при себе си, за да го **откъснел** от влиянието на майката и да го **направел** добър католик.*

'The father kept Henry for himself, so as to bar him from his mother's influence and to raise him as a good catholic.'

Evidential perfect and pluperfect

писал vs. написал, правил vs. направил, казвал vs. казал

The evidential perfect and pluperfect expressed by a single form are formed with the past active participle of the auxiliary *съм* 'be', i.e. *бил*, and aorist participle of the lexical verb. The usage of the participles in the renarrative tenses is identical with their functioning in indicative and they are found in a similar lexical context.

The imperfective participles refer to repetitive events (33), general facts (34), in many cases in negative context.

(33) *Колко пъти й бил казвал на тази патка, че краката му са вечно студени и ако не бъдат добре затоплени, той изобщо не може да заспи!*

'How many times he told this idiot that his feet are always cold and if they don't get heated up well, he can't fall asleep at all.'

(34) *От разменените приказки разбрах, че не се е мил от Пролетния празник. Никой не му бил казвал да го направи след смъртта на майка му.*

'I understood from what he said that he hadn't washed since the spring holiday. Nobody told him to do that after his mother's death.'

The perfective participles denote a concrete fact (35), a few examples are found with the particular meaning of bounded iterativity (36). In the majority of cases the combination of the evidential auxiliary *бил* + aorist participle with iterative meaning is a dubitative aorist.

(35) *Престояло цяла седмица в храма, защото никой не им бил казал къде си отседнал.*

'It remained a whole week in the temple, because nobody had told them where you had put up.'

(36) *100 пъти му бил казал...*

'He told him 100 times.'

5. Temporal opposition

The temporal opposition holds between the aorist and the imperfect participles, which is only possible within the evidentiality system, where the two types of participles are used to form the temporal structure of the category. Tenses are organized by pairs expressed by a single form: present and imperfect; perfect and pluperfect; future and *futurum praeteriti*; *futurum exactum* and *futurum exactum praeteriti*; aorist. Thus the temporal opposition imperfect vs. aorist is expressed by the imperfect and aorist participles, respectively. Due to the two participial paradigms the evidential temporal system can express all types of events and their relations as the indicative tenses.

5.1. Renarrative/inferential imperfect vs. aorist

пишел, напишел vs. писал, написал

правел, направел vs. правил, направил

казвал, казал vs. казвал, казал

The imperfect denotes an event that is simultaneous to a past moment, while the aorist refers to a completed event in the past. The temporal relations are illustrated with two text excerpts in renarrative (37) and inferential (38), which are the evidentials that may be used in longer texts.

(37) *Щом свършил първият танц, Петер се наредил с дамата си горе на площадката до Краля на танца и щом онзи рипнел три стъпки над земята, Петер скачал четири. Направел ли онзи чудни, изящни стъпки, Петер започвал да усуква и върти краката си така, че хората, които го гледали, се захласвали от удоволствие и възторг.*

'When the first dance finished, Peter lined up himself and his lady on the stage next to the King of the Dance and when he jumped three feet from the floor, Peter jumped four. If he did those wondrous, elegant steps, Peter started to fling and twist his feet in such a way that people who looked at him were struck with delight and amazement.'

(38) *"Който и да е бил, трябва първо бавно и безшумно да е убил дежурния, след това – Ту Май, като е запушил устата на младия хан с ръка, докато го е събарял надолу." Бен се обърна. "Да, и е трябвало вратата да остане затворена, докато го е правел, иначе е щял да бъде видян от мъжете около масата." Затвори очи, видял всичко ясно. "Офицерът се е*

оттеглял, когато се е обърнал с лице към Брок, извадил е оръжието си, без да даде на Брок време да стане от стола.”

“Whoever he was, first he must have killed the guard slowly and quietly, and then Tu Mai by gagging the young khan with his hand while wrestling him down.” Ben turned around. “Yes, and the door must have remained closed while he was doing it, otherwise the men around the table would have seen him.” He closed his eyes and saw everything clearly. “The officer must have been withdrawing when he turned to face Brock and drew his weapon without giving Brock time to get up from his chair.”

Example (31) describes a competition in dancing between two characters, Peter and the King of the Dance. The story begins by two single completed actions expressed by perfective aorist participles denoting the renarrative aorist tense: *свършил* ‘finished’, *се наредил* ‘lined up’. The following sentences comprise repetitive events expressed by imperfective imperfect participles in the main clauses that refer to imperfect tense (*скачал* ‘jumped’, *започвал* ‘started’), and in the dependent time clauses two specific verbal forms occur – *рухнал* ‘jumped’ and *направел* ‘did’, which are imperfect participles derived from perfective verbs and correspond to a peculiar meaning of the perfective aspect when combined with imperfect tense to denote repetitive events through a single example (Maslov 1959: 232). The excerpt ends with two continuous actions (*гледали* ‘looked at’, *се захласвали* ‘were struck’) expressed by imperfective imperfect participles.

Similarly, in (32) imperfect and aorist participles are used in inferential forms to express temporal relations in a murder scene inferred by a character in the novel. In that excerpt the typical contrast between aorist and imperfect can be seen, the aorist referring to single and completed events in the past (*е запушил* ‘gagged’, *се е обърнал* ‘turned’, *извадил е* ‘drew’), and the imperfect denoting continuous and incompleted acts that serve as a background for the completed ones (*е събарял* ‘wrestling’, *е правел* ‘was doing’, *се е оттеглял* ‘withdrawing’).

6. Aspect, tense and adjectives

The aorist participles may have adjectival usage and in these cases the perfective stem is preferred. Nevertheless in particular contexts both perfective and imperfective participles may be used as adjectives inheriting the aspectual and the temporal characteristics of the respective participle.

(39) *Четящият впоследствие ще почувства душата на писалия.*

‘The reader will afterwards feel the soul of the writer.’

(40) „Часът на зеления прилив“ очевидно е някакво предварително определено време между *написалия* документа и онзи, който трябва да го прочете.

“The hour of the green flow” is obviously some time period between the writer of the document and the one who has to read it.’

7. Distribution

The general distribution of the past active participles of the verbs *правя* ‘do; make impf.’, *направя* ‘do; make pf.’, and *пиша* ‘write impf.’, *напиша* ‘write pf.’ without specification of the compound verb form is shown in Table 2. The verbs are chosen to illustrate the forms distribution with respect to their frequency and the possibility to compare all four participles.

	aorist		imperfect	
imperfective	<i>правил</i>	6 023	<i>правел</i>	1 121
perfective	<i>направил</i>	24 847	<i>направел</i>	43
imperfective	<i>писал</i>	3 163	<i>пишел</i>	385
perfective	<i>написал</i>	4 529	<i>напишел</i>	3

Table 2: Distribution of the past active participles

The number of occurrences may be analyzed in several viewpoints. Aorist participles are more frequent than imperfect participles as they may be used in perfect tenses of indicative, in conditional and in evidential tenses. The usage of the imperfect participles is limited to few tenses of the indirect evidentiality. With respect to the compatibility of the grammatical features aorist participles derive more often from perfective verbs and imperfect participles – from imperfective verbs. Taking into

account those two trends, the highest frequency of the perfective aorist participle is not unexpected, as well as the smallest number of occurrences of the perfective imperfect participle.

8. Conclusion

Past active participles in Bulgarian form a complex system combining aspectual and temporal characteristics. Their usage in different tenses, moods and evidentials depends on the compatibility of the respective grammatical meanings. Corpus-based studies outline the general tendencies of their usage, the specific contexts that require a given type of participle and the restrictions due to incompatible aspectual and temporal meaning. In general, participles in compound temporal forms cover all the central particular meanings of verb aspects in Bulgarian, thus creating a possibility to express aspectual opposition within perfect tenses, conditional mood and evidentiality. Matching the general trends in aspectual functions, perfective participles have homogenous meaning and usage, and, on the contrary, imperfective ones display much more diversity in their functions and none of their particular meanings can be pointed out as predominant. In terms of frequency, aorist participles prevail considerably above imperfect, the latter being restricted within the evidential system.

Acknowledgement

This research was supported by the project *The Balkan languages as an emanation of the ethnical and cultural community of the Balkans (verb typology)*, financed by the Scientific Research Fund at the Ministry of Education and Science, contract ДН 20/9/11.12.2017.

References

- GSBKE: *Gramatika na savremenniya balgarski knizhoven ezik. Tom II. Morfologiya*. Sofia: Bulgarian Academy of Sciences, 1983.
- Koeva, S., Stoyanova, I., Leseva, S., Dimitrova, T., Dekova, R., and Tarpomanova, E. (2012). The Bulgarian National Corpus: Theory and Practice in Corpus Design. *Journal of Language Modelling*, 0(1): 65–110.
- Maslov, Y. S. (1959). Glagolnii vid v sovremennom bolgarskom literaturnom iazyke (znachenie i upotreblenie). – *Voprosy grammatiki bolgarskogo literaturnogo iazyka*. Moscow, pp. 157 – 312.
- Nitsolova, R. (2008). *Balgarska gramatika. Morfologiya*. Sofia: University Press „Sv. Kliment Ohridski“.
- Stankov, V. (1980). *Glagolniiyat vid v balgarskiya knizhoven ezik*. Sofia: Nauka i izkustvo.