

A New Version for Bulgarian MTE Morphosyntactic Specifications for Some Verbal Forms¹

Ludmila Dimitrova¹, Peter Rashkov²

¹Institute of Mathematics and Informatics
Bulgarian Academy of Sciences, Sofia, Bulgaria

ludmila@cc.bas.bg

²Jacobs University, Bremen, Germany

Abstract

In this paper we propose a new version for MULTEXT-East morphosyntactic specifications for Bulgarian participles.

Keywords: *Bulgarian, impersonal verbal forms, grammatical attribute, morphosyntactic descriptors*

Introduction

The morphosyntactic descriptors for Bulgarian in the framework of the MULTEXT-East (MTE for short) project were developed 12 years ago. Some of them are not strictly adequate to the particular morphosyntactic properties of the respective parts-of-speech – Tense, Number, Gender, Voice, Definiteness – especially in the system of impersonal verbal forms (participles) (MTE 2004).

The system of participles in modern Bulgarian contains a multitude of forms with different origins. Its main elements are the two past participles (perfect and imperfect) and the past passive participial which are characteristic of the spoken language for a long time, as well as the present active participial, which has been re-introduced in the literary language in the nineteenth century. The indeclinable gerund has its own special place, because its distribution does not cover all Bulgarian dialects, yet it is widely used in the written language.

According to the Bulgarian grammarians, Bulgarian participles do not possess the grammatical attribute Voice because the relation between the subject and the verb action conveyed by a participle is attributive (*плачещо дете, подранила зима*), not predicative (*детето плаче, зимата подрани*). In fact, the Bulgarian language (as well as all other Slavic languages) have no special forms for passive voice at all and express it periphrastically.

Yet the Bulgarian classification distinguishes between active participles (деятелни причастия)

- present active participial (Bulg. *сегашно деятелно причастие*):

четящ, ходещ, разказващ,

- perfect active participle (in Bulg. *минало свършено деятелно причастие, formed from the aorist stem of the verb + suffix -I*):

живя-х → живя-л, ходи-х → ходи-л, писа-х → писа-л;

imperfect active participle (Bulg. *минало несвършено деятелно причастие, formed from the imperfect stem of the verb + suffix -I*):

живее-х → живее-л, ходе-х → ходе-л, пише-х → пише-л;

and passive participles (*страдателни причастия*)

- past passive participial (Bulg. *минало страдателно причастие, formed from the imperfect stem of the verb + suffix -н- or -т-*):

писа-х → писа-н, игра-х → игра-н, разказва-х → разказва-н,

ши-х → ши-т, би-х → би-т, чу-х → чу-т, обу-х → обу-т;

¹ The study and preparation of these results have received funding from the EC's Seventh Framework Programme [FP7/2007-2013] under Grant Agreement 211938 MONDILEX

- present passive participial (Bulg. *сегашно страдателно причастие*): *любим, необходим, неделим*, which existed in its own right in Old Bulgarian, but survives only in a few cases and is considered an adjective.

The participles are differentiated according to tense and aspect:

The grammatical category tense is not fully developed either, and the distinction between the so-called present (*сегашни причастия*) and past participles (*минали причастия*), according to the Bulgarian classification is rather due to their historical origins, namely the present or past verb stem, from which the participle is formed, than to any kind of temporal relations.

The grammatical category aspect has also a limited place in the participle system – only imperfective verbs can form the present active participial, while both perfective and imperfective verbs form past participles.

The grammatical categories gender, number, and definiteness apply to participles in a similar way as to adjectives. The categories gender and number are natural to all participles, and in this way the participles resemble adjectives. Definiteness applies to the active participial, past passive and perfect active participles. Hence we may speak of the Bulgarian participle as a hybrid part-of-speech, where each grammatical attribute is only partially represented. This hybridism has attracted the attention of the first modern Bulgarian grammarians (such as Neofit Rilski¹) and made them classify the participles as a separate POS.

Present active participial

The MSD of this type of participle (Bulg. *сегашно деятелно причастие*) contains VForm=participle(p), Tense=present(p), Voice=active(a), for example,

въртящото	въртя	Vmpp-sna-y
дължаща	дължа	Vmpp-sfa-n
изгряващата	изгрявам	Vmpp-sfa-y
изгряващи	изгрявам	Vmpp-p-a-n

The present active participial is a new phenomenon in modern Bulgarian. It disappeared from the spoken language in the historical development of the language during 15-16 century, but was frequently used in Old Bulgarian, for example:

- **Хода же при мори галиленецѣ видѣ Симона и Андрейѣ, брата того Симона, въметаѣшта мрежѣ въ море.** (Mar.) /As he went along the Sea of Galilee, he saw Simon and Andrew, Simon's brother, casting a net into the sea./ [Mark 1:16]

The reintroduction of this verbal form started in the 19 century under Russian and Church Slavic influence. Today this participial has its own specific structure (stem and suffix) in contrast to the same category in other Slavic languages. It represents a property/attribute of a person or an object resulting from an action of the person or the object regardless of any temporal orientation. This participial expresses the property/attribute independently and regardless of the tense of the main verb in the sentence. It corresponds semantically to the English *ing*-form (English present participle), when it modifies a noun or serves as a noun with active sense. For example:

- *Преподавателят посреща (посреща is verb in present tense) идващите ученици.* /The lecturer meets the *coming* students/

- *Преподавателят посрещна (посрещна is verb in aorist tense) идващите ученици.* /The lecturer met the *coming* students/

- *Преподавателят ще посрещне (ще посрещне is verb in future tense) идващите ученици.* /The lecturer will meet the *incoming* students/.

It also functions as an attribute and behaves as a noun or an adjective in the sentence:

- *В стаята имаше много **правостоящи**.* /There were many *standing* [people] in the room/.

- *Пази се от **падащи** предмети!* /Beware of *falling* objects!/.

Furthermore, it can serve as a marker for a subordinate clause, for instance:

- *По улиците се мяркаха тук-там минавачи, **неспокойно поглеждащи** към сивото небе.* /On the streets one caught here and there a glimpse of passers-by, *restlessly glancing* toward the grey sky./

¹ Neofit Rilski published *Bolgarska Grammatika*, the first systematic Bulgarian grammar, in 1835.

This participial cannot possess the Tense attribute because it expresses the property/attribute independently and regardless of the tense of the main verb in the sentence (Bulgarian Grammar 1990). The Voice attribute is also implicit from the context. Hence we propose a new attribute VForm=active_participial(a) in MSD, which replaces the old descriptor in the following manner:

		MSD old version		MSD new proposal
въртящото	въртя	Vmpp-sna-y	->	Vma--sn--y
дължаща	дължа	Vmpp-sfa-n	->	Vma--sf--n
изгрыващата	изгрывам	Vmpp-sfa-y	->	Vma--sf--y
изгрыващи	изгрывам	Vmpp-p-a-n	->	Vma--p---n

Perfect active participle

The MSD of the perfect active participle (Bulg. *Минало свършено деятелно причастие*) contains VForm=participle(p), Tense=aorist(a), Voice=active(a), for example,

отиш ля	отида	Vmpa-sfa-n
отиш ли	отида	Vmpa-p-a-n
отиш ъл	отида	Vmpa-sma-n
отр екли	отрека	Vmpa-p-a-n

From a historical perspective, this participle is quite old, and it is frequently used in Old Bulgarian in the formation of compound tenses, for example,

- **И по чѣто не въдасть моего сѣребра пѣнасъникомъ и азъ пришедъ съ лихвожъ нстазалъ є вимъ?** (Mar.) /Why then didn't you put my money in the bank, so that when I returned I *could have collected* it with interest?/ [Luke 19:23]

It is consequently stable across all Bulgarian dialects in form, semantics and functionality. Its forms are derived from the aorist stem of the verbs + suffix *-л, -ла, -ло, -ли: писа-х > писа-л, живя-х > живя-л, пи-х > пи-л*, etc. This participle has significantly expanded its usage and in modern Bulgarian it expresses a property/attribute resulting from an action that has been completed before the action of the main verb in the sentence, but regardless of the moment of speaking. For example:

- *Падналите ни другари* /our fallen comrades/
- *Гледам дошлия войник.* /I am looking at the soldier who came./
- *Гледах дошлия войник.* /I was looking at the soldier who had come./
- *Ще гледам дошлия войник.* /I will look at the soldier who will come./

The perfect active participle shares all morphosyntactic characteristics of the noun and adjective – number, gender and definiteness. For example:

- *Закъснелите ученици стояха настрана.* /The students who were late stood aside./
- *Пътниците слизаха бързо от пристигналия влак.* /The passengers quickly got off the arrived train (the train that had arrived)./

- *Дошлите се бяха насъбрали пред училището.* /The ones who had come had gathered in front of the school./

Furthermore, it can serve as a marker for a subordinate clause (Bulgarian Grammar 1990), for instance:

- *Година беше се минало откак Василчо, яхнал белия си кон, се изгуби от очите на Божура.* (Y. Yovkov, 1976) /A year passed since Vasilčo, who had straddled his white horse, had disappeared from Božura's sight./

This participle plays a major role in the formation of compound verb tenses and moods (Bulgarian Grammar 1990), for example:

- Indicative: *ходил съм* (perfect), *бях ходил* (plusquamperfect), *ще съм ходил* (future perfect), *щях да съм ходил* (future perfect in the past)

- Re-narrative (inferential): *ходил съм* (aorist) /I [allegedly] went/, *ходил съм бил* (perfect, plusquamperfect) /I [allegedly] had gone/, *щях да ходя* (future and future in the past) /I [allegedly] will go, I [allegedly] would go/, *щях да съм ходил* (future perfect and future perfect in the past) /I [allegedly] will have gone, I [allegedly] would have gone/

- Conditional: *бих ходил* /I would go/.

Taking all this into account, we see that its MSD is not quite adequate since the attribute Tense=аorist(a) merely reflects the historic origin of the perfect active participle, derived from the aorist stem of the verbs (*живя-, лежа-, ходи-, гледа-*). We see that this participle is used to form many more verb tenses and moods, hence its MSD should not contain the morphosyntactic attribute Tense. This participle has no morphosyntactic attribute Voice, because it is an impersonal verbal form; the grammatical voice is implicit from the context. Hence, we only propose the following simplification of the MSD:

		MSD old version		MSD new proposal
буталото	бутам	Vmpa-sna-y	->	Vmp--sn--y
бухналата	бухна	Vmpa-sfa-y	->	Vmp--sf--y
возилото	возя	Vmpa-sna-y	->	Vmp--sn--y
отишла	отида	Vmpa-sfa-n	->	Vmp--sf--n
отишли	отида	Vmpa-p-a-n	->	Vmp--p---n
отишъл	отида	Vmpa-sma-n	->	Vmp--sm--n
отрекли	отрека	Vmpa-p-a-n	->	Vmp--p---n

Imperfect active participle

The MSD of the imperfect active participle (Bulgarian *минало несвършено деятелно причастие*) contains VForm=participle(p), Tense=imperfect(i), Voice=active(a), for instance,

благодаряла	благодаря	Vmpi-sfa-n
вадел	вадя	Vmpi-sma-n
грешала	греша	Vmpi-sfa-n
деряла	дера	Vmpi-sfa-n
правел	правя	Vmpi-sma-n
правела	правя	Vmpi-sfa-n

This participle is a relatively new POS in Bulgarian morphology, as it did not exist in Old Bulgarian and has developed in parallel to the re-narrative mood of the Bulgarian verb. Hence, its MSD is not quite adequate – the Tense=imperfect(i) attribute reflects merely the historic origin of the imperfect participle, as its forms are derived from the imperfect stem of the verb (*живее-, лежа-/леже-, ходе-, гледа-*). **The main role of this participle is to express renarrative, or non-witnessed modality, hence it cannot be used as an attribute. It possesses only the grammatical attributes number and gender, and is exclusively used to form some compound verb tenses in renarrative (or inferential) mood.**

It is formed from the imperfect stem + suffix -л: *пише-х > пише-л, живее-х > живее-л, тие-х > тие-л*, etc. Sometimes the form of this participle coincides with the form of the first past participle because the imperfect and aorist stems of the verbs coincide. This is the case for some verbs from second conjugation (*вървял, търпял*, etc.) and all verbs from third conjugation (*гледал, стрелял*, etc.).

Some example of the usage are for forming present and imperfect tense in renarrative mood, e.g. *ходел съм* /I [allegedly] have been going/, as well as in the modal form *нямало* used to form negative forms of future, future in the past and future perfect in renarrative mood: *нямало да ходя, нямало да съм донесъл*.

Due to this highly specialized semantic function we propose a new language-specific VForm=renarrative(r) in MSD, which replaces the old descriptor in the following manner:

		MSD old version		MSD new proposal
благодаряла	благодаря	Vmpi-sfa-n	->	Vmr--sf
вадел	вадя	Vmpi-sma-n	->	Vmr--sm
грешала	греша	Vmpi-sfa-n	->	Vmr--sf
деряла	дера	Vmpi-sfa-n	->	Vmr--sf
правел	правя	Vmpi-sma-n	->	Vmr--sm
правела	правя	Vmpi-sfa-n	->	Vmr--sf
отидели	отида	Vmpi-p-a-n	->	Vmr--p

Past passive participial

The MSD of the past passive participial (Bulgarian *минало страдателно причастие*) contains VForm=participle(p), Tense=past(s), Voice=passive(p), for instance,

отрязани	отрежа	Vmps-p-p-n
отрязаните	отрежа	Vmps-p-p-y
предаден	предам	Vmps-smp-n
прочетен	прочета	Vmps-smp-n

This participial is also an old form, which has not changed much semantically or formally over the centuries, for example it is used in Old Bulgarian as an attribute:

- **Придѣте къ мѣнѣ въси трѹждаѣштѣи сѧ и оврѣменени.** (Mar.) /Come to me, all you who are weary and *burdened*./ [Matthew 11:28]

It is formed from the aorist verb stem + suffix *-н/-т*: *казах* > *казан*, *пренесох* > *пренесен*, *ших* > *шит*, *носих* > *носен*, *измих* > *измит*, etc. The main function of this participial is to convey the result of an action as an attribute of a person or object (Bulgarian Grammar 1990). In this way it is very close to the adjective and the noun, with which it shares the morphosyntactic attributes Number, Gender and Definiteness. It corresponds to the English past participle when it modifies or serves as a noun with passive sense. For example:

- *Вечер [...] грееха огнените сияния на **запалилите** села.* (Y. Yovkov, 1976) /In the evenings the fiery radiance of the *burnt* villages was shining./

- ***Поканените** чакат в приемната зала.* /The *invited* ones wait in the parlour./

Furthermore, it can serve as a marker for a subordinate clause, for instance:

*И те носели една книга, **намерена** в Света гора, [...]* (Y. Yovkov, 1976) /And they supposedly carried a book *found* in Mount Athos, [...]/

Additionally it is used to form one type of periphrastic passive voice in Bulgarian, similar to the English formation.

*Стоките **ще бъдат прегледани** от комисията.* /The merchandise will be inspected by the commission./, but compare the reflexive-passive form: *Стоките **ще се преглеждат** от комисията.*

This participial cannot possess the Tense or Voice attribute because it is an impersonal verb form and both the temporal characteristic and the relationship between the action (or state) expressed by the verb stem and the participant (object) are presented attributively, not predicatively. Hence we propose a new attribute VForm=passive_participial(v) in MSD, which replaces the old descriptor in the following manner:

		MSD old version		MSD new proposal
отрязани	отрежа	Vmps-p-p-n	->	Vmv--p---n
отрязаните	отрежа	Vmps-p-p-y	->	Vmv--p---y
предаден	предам	Vmps-smp-n	->	Vmv--sm--n
прочетена	прочета	Vmps-smp-n	->	Vmv--sf--n

Conclusion

We hope that these changes, which bring the morphosyntactic description in line with the grammatical characteristics of the Bulgarian participles, will be more useful for automated processing of Bulgarian texts.

References

MTE, 2004: MULTEXT-East Morphosyntactic Specifications – version 3, edition 10th May 2004.

Bulgarian Grammar, 1993: Главна редакция Д. Тилков, Ст. Стоянов, К. Попов. Граматика на съвременния български книжовен език. Том 2 / МОРФОЛОГИЯ. Издателство на БАН. София. (In Bulgarian).

Y. Yovkov, 1976: Йордан Йовков. Старопланински легенди. В Събрани съчинения в шест тома. Том втори, Издателство "Български писател", София, под общата редакция на Симеон Султанов. (In Bulgarian).

Mar.: Codex Marianus (10 century manuscript), 7-bit ASCII code at <http://www.slav.helsinki.fi/ccmh/marianus.html>

Appendix

1. Common table for Verb:

Verb (V)
14 Positions

```

**** **** **** **** **** **** **** -----
PoS  Type VFrm  Tens Pers Numb  Gend Voic  Neg  Def  Cltc Case Anim Clt2
**** **** **** **** **** **** **** -----

= ===== = EN  RO  SL  CS  BG  ET  HU
P ATT          VAL          C  x  x  x  x  x  x  x
= ===== =

1 Type          main          m  x  x  x  x  x  x  x
          auxiliary      a  x  x  x  x  x  x  x
          modal          o  x  x  x  x          x
          copula         c          x  x  x
          l.s. base       b  x

-----
2 VForm          indicative      i  x  x  x  x  x  x  x
          subjunctive     s          x          x  x  x
          imperative      m          x  x  x  x  x  x
          conditional      c  x          x  x          x  x
          infinitive       n  x  x  x  x          x  x
          participle       p  x  x  x  x  x  x
          gerund           g          x          x  x
          supine           u          x          x
          transgressive    t          x
          l.s. quotative    q          x
          l.s. renarrative  r          x
          l.s. active_participial  a          x
          l.s. passive_participial  v          x

-----
3 Tense          present          p  x  x  x  x  x  x  x
          imperfect       i          x          x  x
          future          f          x  x
          past           s  x  x  x  x          x  x
          l.s. pluperfect  l          x
          l.s. aorist      a          x

-----
4 Person          first          1  x  x  x  x  x  x  x
          second          2  x  x  x  x  x  x  x
          third          3  x  x  x  x  x  x  x

-----
5 Number          singular        s  x  x  x  x  x  x  x
          plural          p  x  x  x  x  x  x  x
          l.s. dual         d          x
          collective       l

-----
6 Gender          masculine       m          x  x  x  x
          feminine       f          x  x  x  x
          neuter          n          x  x  x  x

*****
7 Voice          active          a          x  x  x  x
          passive         p          x  x  x
-----

```

2. Application to Bulgarian Verb

Verb (V)

P	ATT	VAL	Example	C
1	Type	main auxiliary	govorya sym	m a
2	VForm	indicative imperative participle l.s. renarrative l.s. active_participial l.s. passive_participial gerund	govorya govorete govoril govorel govorest govoreno govorejki	i m p r a v g
3	Tense	present imperfect l.s. aorist	govorya govoreh govorih	p i a
4	Person	first second third	govorya govorish govori	1 2 3
5	Number	singular plural	govorya govoryat	s p
6	Gender	masculine feminine neuter	govoril govorila govorilo	m f n

7	Voice	active	govorya	a
8	Negative			-
9	Definiteness	no yes short_art full_art	govoril govorilite govoriliya govoriliyat	n y s f
10	Clitic			-
11	Case			-
12	Animate			-
13	Clitic_s			-

Combinations

PoS	Type	VFrM	Tens	Pers	Numb	Gend	Voic	Neg	Def	Cl1	Case	An	Cl2	Example
V	[ma]	i	[pai]	[123]	[sp]	-	a	-	-	-	-	-	-	1.
V	[ma]	m	-	2	[sp]	-	a	-	-	-	-	-	-	2.
V	m	p	-	-	s	[mfn]	-	-	n	-	-	-	-	3.
V	m	a	-	-	s	m	-	-	[sf]	-	-	-	-	4.
V	m	p	-	-	s	[fn]	-	-	[ny]	-	-	-	-	5.
V	m	r	-	-	s	[mfn]	-	-	-	-	-	-	-	6.
V	m	a	-	-	p	-	-	-	[ny]	-	-	-	-	7.
V	m	v	-	-	s	[mfn]	-	-	n	-	-	-	-	8.
V	m	v	-	-	s	m	-	-	[sf]	-	-	-	-	9.
V	m	v	-	-	s	[fn]	-	-	y	-	-	-	-	10.
V	m	v	-	-	p	-	-	-	[ny]	-	-	-	-	11.
V	a	[pr]	-	-	s	[mf]	-	-	n	-	-	-	-	12.
V	a	p	-	-	s	m	-	-	[sf]	-	-	-	-	13.
V	a	p	-	-	s	[fn]	-	-	y	-	-	-	-	14.
V	a	[pr]	-	-	p	-	-	-	[n-]	-	-	-	-	15.
V	a	p	[a]	-	p	-	-	-	y	-	-	-	-	16.
V	[ma]	g	-	-	-	-	-	-	-	-	-	-	-	17.
V	m	r	-	-	p	-	-	-	-	-	-	-	-	18.
V	m	a	-	-	s	[mfn]	-	-	n	-	-	-	-	19.

Examples:

1. govorya, sym
2. govori, bydete
3. govovil, govovila, govovilo - like noun/adjective
4. govovestiya, govovestiyat - like noun/adjective
5. govovila, govoviloto - like noun/adjective
6. govovrel, govovrela, govovrelo -renarrative, only in verbal use
7. govovresti, govovrestite - like noun/adjective
8. govovren, govovrena, govovreno - like noun/adjective
9. govovreniya, govovreniyat - like noun/adjective
10. govovrenata, govovrenoto - like noun/adjective
11. govovreni, govovrenite - like noun/adjective
12. bil, bila - like noun/adjective
13. biliya, biliyat - like noun/adjective
14. bilata, biloto - like noun/adjective
15. bili -renarrative and participle form coincide - see table
16. bilite - like noun/adjective
17. govovrejki, bidejki
18. govovreli -renarrative, only in verbal use
19. govovrest, govovresta, govovresto - like noun/adjective