On the Diachrony of the Clitic Cluster in Bulgarian

Tsvetana Dimitrova
Institute for Bulgarian Language, Sofia

Abstract The article traces back the formation of the clitic cluster in Bulgarian starting from the Old Church Slavonic through Middle Bulgarian up to the Early Modern Bulgarian and beyond. It offers a hypothetical two-layer structure of the cluster – with the main layer consisting of a (pronominal) core and a (verbal) periphery, and a secondary layer hosting (‘quasi-clitical’) elements that exhibit, both diachronically and synchronically, a behaviour that is not strictly consistent with that of the clitical elements. The language material from three corpora shows that there was no change in the positions of the elements in the core, while the observed changes in the periphery are mainly due to changes in the set of the elements (as a result of the restructuring of the pronoun system and changes in the auxiliary system, as well as the loss of some early clitics, such as the discourse markers).


Summary 1 Introduction. – 2 Clitics in Bulgarian. – 3 Clustering: The History. – 3.1 Discourse Clitics. – 3.2 Pronominal Clitics. – 3.3 Verbal Clitics and Clitic-Like Elements.
Introduction

Slavic clitics have been the focus of attention in a number of works in the recent decades (Franks, King 2000; Franks 2017; Zimmerling 2013; Migdalski 2016; among many others), with accounts also on the diachrony of cliticization in (one or more) Slavic languages.\(^1\)

The diachrony of the Bulgarian clitics has drawn particular interest because Bulgarian is part of the Balkan Sprachbund and as such it exhibits verb-adjacent cliticization pattern, but also because it displays various patterns in its history.\(^2\) Although it has been claimed that the position of clitics in Old Church Slavonic (OCS)\(^3\) texts mostly echoes their placement in the Greek originals, there are numerous exceptions reflecting patterns in the language of the scribes (especially in constructions without parallels in Greek; see Slawski 1946, among others). In her seminal study on the clitics in the history of Bulgarian, Pancheva (2005) argues that verb-adjacent clitics found in Old Bulgarian (and OCS) temporarily switched to second-position (up to and beyond the Middle Bulgarian period), only to become verb-adjacent again (in modern Bulgarian). Migdalski (2016) claims that in the majority of cases pronominal clitics in OCS were verb-adjacent while the second position (2P) was obligatory only for the so-called ‘operator’ clitics (discourse markers and ‘ethical datives’); he relates the verb-adjacency to the presence of tense morphology in a language.

The discussion in the present article is based on data excerpted from manuscripts that presumably reflect the diachrony of the Bulgarian language and are closer to the vernacular than to the literary language. The main aim of the article is to sketch the evolvement of the (sentential) clitic cluster in the history of Bulgarian and to check

---


\(^3\) Here I will use the term Old Church Slavonic (OCS), which reflects the status of the language as used by the Slavic orthodox community (cf. Bujukliev et al. 1993). However, the linguistic features of the texts in the monuments are assumed to reflect those of Old Bulgarian.
whether there has been a significant change in the clustering pattern. The observations on this specific set of data confirm the old hypothesis that the clitics in Bulgarian are (and were) verb-adjacent (and also target 2P). The necessary ramification here is that any conclusions based on heterogeneous diachronic data highly depend on the sources and their analysis. The notion ‘second position’, for example, depends on syntactic constituency within the language and on the interpretation of the data that are taken into account. With diachronic texts, the segmentation is also a matter of interpretation: for the OCS data here, the texts were already segmented within the corpus used for reference; with the other data (Middle Bulgarian and Early Middle Bulgarian), however, segmentation was done aside. The difficulties with the specifics of the data remain: with the OCS data I have tried to consider variant readings if accessible, with interesting results; this should be done over all the data if possible (for example, one should consider variant readings within the texts of the damaskin collections). In order for such a task to be accomplished, however, one needs a really big and smartly annotated corpus, which is unfortunately missing at this time.

2 Clitics in Bulgarian

To define ‘what a clitic is’ is not a trivial task because, as Franks (2017, 146) puts it, “the idiosyncrasies of clitics all go in the same direction, i.e., a negative one”. It is easier to say what clitics are not rather than what they are, because they are – prosodically, morphologically, lexically, and syntactically – ‘deficient’. Firstly, clitics are prosodically deficient as they do not have an independent stress, but form a prosodic unit with another, prosodically ‘strong’, word. However, in many languages clitics may host the stress when preceding or following other elements that remain prosodically ‘weak’ (or may receive secondary stress). In present-day standard Bulgarian, clitics may bear stress after negation as in (1) (there is also a hypothesis for a secondary stress here; Krăsteva (2020, 119), on the basis of experimental data, claims that in interrogative sentences with negation, even when the clitics receive stress, they do not have “an independent intonational peak”).

4 A more detailed review cannot be supplied here but ca. 43% of the occurrences of the datives ми ‘I.dat’ and ти ‘you.dat’ in Codex Marianus (CM) have variant readings in other New Testament (NT) texts as witnessed in Codex Zographensis (CZ), Codex Assemani (CA) and Codex Sabbææ (CSb), and/or in the respective Greek editions.
Secondly, clitics are a ‘bag’ of grammatical features (Sadock 1991) and do not have lexico-conceptual features, i.e. they are also semantically deficient (Franks 2017, 154). Historically, the short dative forms of the first- and second-person singular pronouns ми ‘I.dat’ and ти ‘you.dat’ were already clitics in Old Bulgarian/OCS (as were the parallel pronouns in Greek, see Wackernagel 2009; Večerka 1989; 1993); other short pronominal forms, however, became clitics as a result of early attrition: 3pSg masculine accusative е-го (ѥ-го)5 ‘he.acc’ > го ‘he.acc’, with further reconsideration of г- as a third-person marker resulting in a generalised third-person plural with the ending -и (ги ‘they.acc’); еѫ/e > ѭ > я ‘she.acc’ (3pSg feminine accusative); емоу > моу ‘he.dat’ (3pSg masculine and neuter dative), еи > ей > й ‘she.dat’ (3pSg feminine dative), etc.; see Byjukliev et al. 1991, 234-6; Mirčev 1963, 165-6).

Thirdly, clitics are syntactically dependent although different elements may exhibit different behaviour. Clitics can be categorised into ‘simple’ (non-paradigmatic) and ‘special’ (paradigmatic) (Zwicky 1977; 1985): in present-day Bulgarian, the latter are present-tense auxiliaries/copulas, short forms of personal and possessive pronouns, which have corresponding ‘full’ (prosodically and syntactically ‘strong’) forms.

Clitics are known for their ability to clusterise – in some languages the clitic cluster has a fixed position in the clause or in the phrase (in standard Bulgarian these are the ‘special’ clitics and the ‘simple’ interrogative ли). The so-called Wackernagel clitics in Indo-European languages were all second-position (2P) clitics. At issue here would be the definition of the second position: whether the clitic is positioned after the first word in the sentence, as originally formulated by Wackernagel (2009), or after the first syntactic constituent (including a phrasal one) (Halpern 1995), or after the first constituent that immediately follows the so-called rhythmic-syntactic barrier of Zaliznjak (2008). Zaliznjak introduced this barrier as an ad-

---

5 Orthographic variants throughout the text are generalised if possible (e.g. ero = erw; мн = мі etc.), except for certain examples. The same is true for jotated variants (i.e. ero is given instead of ⶶро etc.).
ditional (and obligatory) condition to make the Wackernagel effect work for the Old Russian clitics. Additionally, the restrictions known as the Tobler-Mussafia’s law (Tobler 1875; Mussafia 1888), originally formulated for the Romance languages, do not allow ‘special’ clitics to be first in the sentence.

All these constraints are partially valid for present-day standard Bulgarian: clitics are mostly verb-adjacent, and are found either after the verb in the first position in the clause, as in (2a), or before the verb if there is another constituent in first position, as in (2b). Although (2c) is ungrammatical with clitics following both the subject and the verb and coming third in the linear order of the clause, clitics can still be found further up in the clause following more than one constituent, as in (2d). Clitics may also land third after a preceding verb that, however, is analysed as first after the rhythmic-syntactic barrier (with other sentential element to its left), as in (2e).

(2) a.

Даде му я Иван книгата

give he.dat she.acc Ivan book.def

b.

Иван му я даде книгата

Ivan he.dat she.acc gave book.def

c.

*Иван даде му я книгата

*Ivan gave he.dat she.acc book.def

d.

В събота изненадващо Иван му я даде книгата

In Saturday surprisingly Ivan he.dat she.acc gave book.def

e.

Да, даде му я книгата Иван

Yes, gave he.dat she.acc book.def Ivan

According to Zimmerling (2013, 89), there are different language systems regarding the position of the clitics, and Bulgarian belongs to the class of languages with so-called ‘extended WP+ system’, in which the clustering clitics are expected to be verb-adjacent or rather the verb in sentences with 2P clitics is expected to be clitic-adjacent, i.e.
the positional constraint here is both on the clitic(s) and on the verb.\(^6\) In addition, the clustering clitics are in more or less fixed position relative to the clause boundary – generally second after an element in complementiser position (conjunctions, wh-words, other relativisers, etc.) or following the first constituent that comes immediately after the so-called rhythmic-syntactic barrier of Zaliznjak (2008).

The following considerations are applicable for the discussion that follows. In the article, I assume that second-position of the clitics is the position following: (a) a specific set of conjunctions (да ‘to’, че ‘that’, ўко ‘as’, аште ‘if’, ами ‘but’); (b) wh-words; (c) other relativisers (дето, де, щото ‘that, which’);\(^7\) (d) a constituent in focus or topic; or (e) verbal element in first position. Sentential constituents located outside the boundary of the clause such as vocatives and some extra-sentential particles do not count as first position. Differently from other Slavic languages, clitics in Bulgarian cannot be found first in the sentence, and in spite of some exceptions in historical texts, Bulgarian clitics in general comply with the Tobler-Mussafia’s law if and only if the latter applies to the clause and not to the level of the intonational phrase (i.e. relative to the rhythmic-syntactic barrier of Zaliznjak (2008). In what follows, we will see that these considerations are also valid for the history of Bulgarian in spite of some counter-examples.

3 Clustering: The History

The hypothesis probed in the article is that the order of the elements within the cluster is a follow-up on the diachronic development of elements targeting the second position, from left to right. Thus, the oldest clitics (2P) would be on the far left (or first) of the clustering elements, while the newest ones would land on the (far) right. Beside the special clitics and the discourse clitics, elements that are found around the cluster and are considered (semi-)part of it by some authors (see more in Zaliznjak 2008), are: the interrogative clitic ли, which is found on the far left but also, given that it does not impose any syntactic restrictions on the preceding element, it may show up after each of the other elements in the cluster; the old adverbial particles и ‘and’, тоу ‘then’; and the adverbs пакъ, пакꙑ ‘again’.

---

\(^6\) As one of the reviewers rightly put it "a decisive factor could be the position of the verb, and in that scenario the second position requirement is irrelevant", although there are still some exceptions, though obscure or stylistically marked.

\(^7\) These form prosodic words (linked by _) with the clitics in some texts - in fact, in these positions, the clitics are part of the first prosodic complex: DLv (17th c.; Mladenova, Velčeva 2013): дето_м_е саде войла оном%H_и влък_ шото_е сега АНТИХРИСТЬ. И_да_съ поклонять влък, че_м_е работата кафо на_влък.
Not all clustering elements in present-day Bulgarian show similar behaviour. Monosyllabic auxiliaries, other than the present-tense ones, such as the conditional би ‘would’, the past-tense бе ‘was/were’, the future ще ‘will/shall’, may also appear first in the clause. Both in present-day standard Bulgarian and in the historical data observed in the article, these auxiliaries may be found either to the left or to the right of the pronominal elements (the short forms of dative and accusative personal pronouns, the short forms of possessive personal pronouns, which have the same form as the dative personal ones, and the short reflexive pronouns – accusative се and dative си). Within the cluster, the pronominal clitics appear in a fixed order with respect to one another: the dative always precedes the accusative (with reflexives behaving in a parallel way), while the present-tense auxiliaries show a split behaviour: all persons except third person singular show up to the left of the pronominals, while the third-person singular auxiliary (‘е’is.3pSg’) shows up to its right.

In the present article, I propose that the cluster is composed of a primary layer and a secondary one. The primary layer is further split into a core (pronominal) and a periphery (verbal). This is also historically motivated as the auxiliary clitics have joined the clustering elements later (present-tense auxiliary clitics are derived via attrition to the left or to the right of the remaining elements: first-person singular е-съм > съм ‘am.1pSg’; second-person singular е-си > си ‘are.2pSg’; third-person singular е-стъ > е ‘is.3pSg’; third-person plural сѫ-тъ > са ‘are.3pPl’). The secondary layer hosts elements that are clitic-adjacent such as past-tense auxiliaries whose position also varies as in (3a) vs. (3b), as well as the conditional (see Nicolova 2008).

(3) a.

Иван бе/беше/бил му го дал.
Ivan be.aor/impf/renarr he.dat he.acc given

b.

Иван му го бе/беше/бил дал.
Ivan he.dat he.acc be.aor/impf/renarr given

In the pattern in (4), the primary layer is in bold. Some dialects have retained only present-tense AUX2, i.e. the verbal clitic is to the right (Antonova-Vasileva et al. 2016, map 144).

(4) (qu) (aux0) AUX1 DAT ACC AUX2 (aux *)

The ordering in (4) illustrates the overall pattern in present-day standard Bulgarian and if we are to judge by the data discussed be-
low, it has undergone relatively few changes, which have rather affected the inventory of the elements that may appear in the respective slots rather than the pattern.

The earliest 2P clitics in Old Bulgarian/OCS are the discourse markers же and бо (the latter, however, was lost early). The pronominal clitics, which presumably joined the cluster at a later point and which build the core of the primary layer also in the history of Bulgarian, can be classified into several subclasses depending on their placement within the cluster: (1) the dative first-person singular ми ‘I.DAT’ and second-person singular ти ‘you.DAT’; (2) the short forms of anaphoric pronouns (which would be later reanalysed as third-person pronouns): и ‘he.ACC’ (masculine singular; neuter and feminine dual), ж ‘she.ACC’ (feminine singular), е ‘it.ACC’ (neuter singular), и ‘they.ACC’ (plural, masculine and feminine), б/б ‘they.DU’ (neuter plural; masculine dual); the accusative short forms of the first- and second-person singular personal pronouns м ‘I.ACC’ and т ‘you.ACC’ (and, possibly, plurals and duals), defined as semi-clitics (Vaillant 1948, 262; Večerka 1989, 42) as they may occur in ‘strong’ positions – after a preposition and (rarely) at the beginning of a sentence; and reflexives с (accusative) and си (dative), also found after prepositions; (3) in later texts, the inventory expands with third-person singular datives моу ‘he.DAT’ (masculine and neuter) and й ‘she.DAT’ (feminine) derived from the earlier dative short forms емоу and еи that also tend to appear more or less adjacent to the position of the cluster in earlier texts.

Both historically and in present-day Bulgarian, verbal elements (‘connectors’; Zimmerling 2013) among the clustering elements are located in the periphery of the primary layer and display variation in placement with respect to other clustering elements in that they can occupy either aux1 or aux2 in (5) below. These are: the monosyllabic auxiliaries of бѫти (optative бѫ/бѫ ‘would’; aorist бѫ ‘was/were. aor’; imperfect б ‘was/were.impf’, which, however, may be found also in first position); the future auxiliary ще/щѫ ‘will/shall’ (with variants ке, хте, че, etc.) appearing not earlier than Middle Bulgarian; present-tense auxiliaries and their successors: есмъ (съм, смъ) ‘am.1pSg’, ес (с) ‘2pSg’, 3pSg есть (е) ‘is.3Sg’, есме (сме) ‘are.1pPl’, есть (сте) ‘are.2pPl’, сѫтъ (сѫ, са) ‘are.3pPl’.

The interrogative clitic ли (to the far left in both (4) and (5)) and adverbial particles are part of the secondary layer. These are: ти ‘and’, и ‘and’, пак(ъ) ‘again’, той ‘then’, etc. The adverbs may show up on both sides of the primary layer, as indicated by the two positions of ADV in (5).

The pattern in (5) gives the relative order of the elements that tend to clusterise as witnessed in the earliest historical texts. With some changes, this order correlates with the one that has been observed by Zaliznjak (2008, 82) in the Novgorod charters. Note that not all elements that appear in bold in (5) are necessarily clitics; as will be seen below, some of them are semi-clitics, while others are clitic-like.
According to the so-called ‘historical principle’, the order of the clitics in the cluster mirrors the chronology in which they have become ‘weak’, a pre-condition for their later reanalysis as clitics. This principle predicts that potential members of the clitic cluster would join the clustering elements only to their right. According to Zimmerling (2013), this principle cannot explain the order but can only account for the change in the status of the elements, which is a mechanism for expanding the cluster. Observations on the data that will be discussed, however, show that this is true for a limited set of elements with relatively fixed slots, either in the clause or in the phrase – the discourse markers and the pronominal clitics, which, as mentioned above, constitute the core of the clitic cluster (while verbal elements fluctuate).

The discussion is based on data that has been excerpted from texts from three diachronic corpora. For Old Church Slavonic (OCS), these are the annotated texts of two monuments – Codex Marianus (CM) and Codex Suprasliensis (CSpr) within the TOROT corpus (Haug, Eckhoff 2011a; 2011b), which are used for the numbers; variant readings here are consulted according to three monuments – Codex Zographensis (CZ), Codex Assemani (CA), and Codex Sabbæ (CSb) in the TITUS database. For Middle Bulgarian, the texts used are the Legend of Troy (LTr) (14th c.; Miklošič 1871) and the Vlacho-Bulgarian Charters (VB) (15th-16th c.; Bogdan 1902) that are available in the Diachronic Corpus of the Bulgarian Language (Totomanova 2015). For Early Modern Bulgarian (EMB), the texts in two damaskins (17th c.) are used: the Damaskin of Troyan (DTr; Ivanova 1967) and the Damaskin of Lovech (DLv; Mladenova, Velčeva 2013). The texts were chosen because of their accessibility and availability for replica and the hypothetical closeness to the vernacular.

The reflexive pronoun сѧ ‘self. acc’ exhibits ‘mixed’ behaviour of a marker that reflexivises the verb and a semi-clitic but is also found among the clustering elements as we will see later. There are a couple of examples with other accusatives, therefore the reflexive is placed within the pattern here, albeit tentatively.

Expectably, some orders that are present in our data are not found in present-day standard Bulgarian: for example, pronominal clitics could occupy the first sentential position that is never the case in present-day standard Bulgarian (though this phenomenon is observed in some dialects, see Iliev 2018, Tiševă 2008, among other). Also expectably, clitics were found in the absolute second prosodic position as well; in some cases, they could split a noun phrase much like what happens in present-day Serbian. This is no longer the case in standard Bulgarian.

Within the TOROT corpus: http://syntacticus.org/; the TITUS database: https://titus.fkidg1.uni-frankfurt.de/framee.htm?/index.htm; and Diachronic Corpus of the Bulgarian Language: https://histdict.uni-sofia.bg/textcorpus/list

Texts in CSpr, VB, DTr and DLv have different linguistic properties, probably following different redactions (and translations) (in the case of CSpr, DTr, DLv; the collection of VB in the Bogdan’s 1902 publication contains texts written over two centu-
3.1 Discourse Clitics

The discourse clitics же (a focus marker or an emphatic particle) and бо (a complementiser or a marker of subordination) are widely attested in OCS (with parallels in Greek; cf. Thayer 1889). These clitics usually appear in the second clausal and/or phrasal position: they are the earliest 2P clitics and both were already used as suffixes (as in же ‘who/which’, никъто же ‘nobody’, неже ‘therefore’, ибо ‘because’, etc.). Migdalski (2016) calls them “operator clitics” since they additionally ‘operate’ on the information structure of the sentence. If found together (rarely in the observed texts), the order is же > бо, as in (6a); pronominal clitics follow them (6b-c).

(6) a. 

\[ \text{ако же бо вълшеб къ дученикомъ} \quad \text{CSpr, 251r, 19-20} \]

\[ \text{if because entered} \quad \text{to disciples} \]

\[ \text{Оти мън яр есълве пръс този мадта} \]

Due to space limitation, glosses are given if there is no appropriate translation and with elements that are in focus in the article, mainly clitics (e.g., is.aux.3pSg: present-tense auxiliary, 3pSg; be.aux.cond: conditional auxiliary; etc.).

b. 

\[ \text{оунѣе бо ти есть} \quad \text{CM, Mt. 5,29} \]

\[ \text{better because you.dat is.cor.3pSg} \]

\[ \text{съмѣрѣй яр сой} \]

d. 

\[ \text{за погъблѣть единъ оудъ твоихъ} \quad \text{CM, Lk. 14,14} \]

\[ \text{to die one members.GEN your.GEN} \]

\[ \text{яна апѣлтѣй енъ товн мелън сой*} \]

* Three Greek NT editions were consulted. Wherever possible, I supply the corresponding Greek text with variants according to other editions if they attest for variant readings with respect to the clitics. If there is no reference to an edition, the text is given according to Nestle 1904; for a discussion on editions vs. manuscripts, see Toufexis 2010.

c. 

\[ \text{въдась бо because ти you.dat са} \quad \text{CM, Lk. 14,14} \]

\[ \text{give} \]

\[ \text{аньтаподобѣсть яр you.dat сой} \quad \text{Nestle 1904} \]

\[ \text{аньтаподобѣсть дѣ сой} \quad \text{Tischendorf 8th ed.} \]

ries ago, which reflect different language varieties and influences). The raw numbers given in the article show the overall trends without aiming at a thorough analysis as the data is quite heterogeneous.
The particle же is used for emphatic marking of the preceding constituent and has the semantics of a weak adversative conjunction, most often translating Greek δέ ‘but, and’ – as in (7a) – and more rarely μέν ‘while’, but also (in various contexts) τέ ‘and, also’, οὖν ‘therefore’, δή ‘but, then’, γάρ ‘because’, γέ ‘in fact, only’, καὶ ‘and’. In some cases, there is no parallel element in Greek, as in (7b); see also the variation in (7c) where οὖν is translated as ‘же оубо’, only ‘оубо’ and only ‘же’ in the three manuscripts cited.

(7) a. 

онъ же рече имъ врагъ [...]
he EMRH told they.dat enemy
Онъ же [р]ече [ймъ] врагъ CA
о δέ ἐφη αὐτοῖς, Ἐχθρός
rabby.ephr VSAT isim owner
slaves EMRH told he.dat
Онъ же ръша [...] OI δέ δοῦλοι λέγουσιν αὐτῷ

b. 

[...] приходитъ же неприѣзнь [...] CM, CZ, Mt. 13,19
comes EMRH evil-one
[...] приходитъ неприѣзнь [...] [...] ερχεται ὁ πονηρός [...] b. 

C. 

ерда же оубо придетъ гъвъ винограда [...]
when therefore comes owner vineyard.gen 
ерда же придетъ гъвъ винограда [...]
CZ
'и эрда оубо приде гъ виноградоу [...]
CSb
отав оун ἐλθη ὁ κύριος τοῦ ἀμπελῶνος [...]

Both же and бо may split a constituent – see (8) where же separates зѣло ‘much’ and зовѫштѹ ‘who is calling (him)’ (ἐπικράζοντος).

(8)

Зѣло же зовѫштѹ ἐγό много ἔ γλαγολѫштѹ CSp, 21r, 28-29
much EMPH calling I.dat he.acc much and speaking
Επικράζοντος δέ μου αὐτῷ καὶ λέγοντος πολλάκις:
because. The use of οὐμόν as an emphatic particle is a very early development, mostly in interrogative sentences – οὐμόν is predominantly found after a wh-word (Cejtlin 1994, 721-2) and translates different elements in Greek: ἀρα μὲν ‘then indeed’, οὖν ‘therefore’, τοίνυν ‘therefore’ (each has specific placement: οὖν is predominantly in 2P, and μὲν can be found after the definite article, i.e. phrase-internally). There are adjacent placements of οὐμόν and бо, with variant readings as in (9).

(9)

ѣко бо οὐμόν σὑβίαράντη σπλѣвелы. CM, Mt. 13,40
as therefore collect weeds
ѣко οὐμόν σὑβίαράντη σα’ CA
ѣко οὐμόν σὑβίαράντη σπλѣвелы CZ
ὡσπερ οὖν σὑλλέγεται τά ζίζάνια

The constituent preceding же and бо can be a verbal element, a noun or a pronoun, an adjective, an adverbial, a wh-word such as къто ‘who’, чъто ‘which/that’, по чъто ‘because of which’, etc., which can be preceded by a conjunction, negation particle, and/or followed by μι ‘I.DAT’, τι ‘YOU.DAT’, τά ‘YOU.ACC’, CA ‘SELF.ACC’, etc.

In Middle Bulgarian texts observed here, both discourse markers are rare. In LTr, же has only 9 occurrences, and бо is found 7 times. The raw numbers given by Dimitrova and Bojadziev (2014) show that many later non-canonical sources exhibit higher numbers for οὐμόν but not for бо. In VB, бо is missing and же occurs 47 times, mainly in fixed expressions such as великъ же и малъ ‘big but also small’ (26 occurrences) and еще же (12 occurrences). же is also found after pronominals such as the dative ви ‘you.2pPl’ in: варе шо ви же речет жѣпан Кр(ѣ)стѣ и жѣпан Ханѫш ‘something that župan Krastyo and župan Hanush told you’.
3.2 Pronominal Clitics, Semi-Clitic and Clitic-Like Elements

The short forms of personal and anaphoric pronouns can be grouped into three subclasses according to their cliticisation status in OCS. The raw numbers in Table 1 support the assumption that they were all on the path to becoming clitics (in some of the accounts cited above, they are already analysed as such) but there are differences as well.

The first subclass includes the datives МИ ‘I.dat’ and ТИ ‘YOU.dat’, which are the first pronominals to exhibit the behaviour of clitics: according to Wackernagel (2009), the Greek datives МОИ ‘I.dat’ and СОИ ‘YOU.dat’ were clitics, hence, their corresponding pronouns in OCS МИ and ТИ might have been clitics, as well (Večerka 1989; Wackernagel 2009, 496, fn. 8 by D. Langslow). In the data observed here, they tend to appear in 2P (ca. 73%) preceded by a verb (ca. 67% of all 2Ps; and 72% of all occurrences) or after a nominal element, an adverb, a conjunction such as ДА ‘to’, АШТЕ ‘if’, ΚО ‘as’, and a wh-word (ca. 25% of all 2Ps; and ca. 21% of all occurrences). Although Migdałski (2016) claims that these were mainly ethical datives functioning as ‘operator clitics’, a study on these specific cases by Krapova and Dimitrova (2015) shows that many of these datives can be analysed either as affected participants, or as inalienable possessors, or as both (i.e. affected participants that are also inalienable possessors, see Minčeva 1964).

The second subclass of pronominal clitics comprises the so-called semi-clitics and can be further divided into two groups: a) the first- and second-person accusatives МА ‘I.acc’ and ТА ‘YOU.acc’, which appear after a preposition and in first position; and b) the third-person accusatives (И ‘he.acc’, Ж ‘she.acc’, А ‘they.acc’, Е ‘it.acc’, Η ‘they.acc’), which often translate non-clitic Greek pronouns. Elements of both groups show a strong tendency to appear in second position, as well as post-verbally. There is a slight precedence of the second group to appear post-verbally: ca. 78% (МА, ТА) vs. ca. 93% (И, Ж, А, Е, Η). When in second position, however, there is a clear difference between the two groups with respect to their pre-verbal placement: ca. 6% (МА, ТА) vs. ca. 25% (И, Ж, А, Е, Η).

The third subclass of pronominals – involving pronouns that will be defined as ‘clitic-like’ in the present discussion – includes the short dative bi-syllabic ЕМОУ ‘he.dat’, ЕИ ‘she.dat’, ИМЪ ‘they.dat’, etc.,

12 In the so-called archaic damaskins, however, both are still present, in the earlier positions.
which have ambiguous behaviour with respect to discourse markers and the first two pronominal subclasses, and whose successors in later stages of Bulgarian (мой ‘he.dat’, й ‘she.dat’, etc.) are clitics. They are also found predominantly in second position (ca. 75%) and, (even) more often, post-verbally (ca. 92%). However, only 7% of the clitic-like pronouns are both 2P and pre-verbal – this percentage is comparable to the one seen in the preceding paragraph with respect to the first- and second-person accusatives мѧ ‘I.acc’ and тѧ ‘you.acc’ (the first subclass of the semi-clitics).

These results may indicate, albeit it is a tentative proposal, that 2P cliticisation was the dominant pattern, at least during the period witnessed by these OCS texts, and that this pattern was not in competition with pre-verbal placement but could co-exist with the latter in case the word order would allow for it. The 2P pattern and the post-verbal one were not in competition either and perhaps can be seen as alternative mechanisms motivated also by word order and/or information-structure principles and/or the syntax of the original Greek text.

Table 1 Pronominal clitics, semi-clitics and clitic-like pronouns in OCS (according to CM and CSpr)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ми, ти</td>
<td>75</td>
<td>204</td>
<td>26</td>
<td>12</td>
<td>97</td>
<td>4</td>
</tr>
<tr>
<td>емоу, еи, имъ</td>
<td>77</td>
<td>1053</td>
<td>17</td>
<td>23</td>
<td>362</td>
<td>7</td>
</tr>
<tr>
<td>мѧ, тѧ, нѣ, вѣ</td>
<td>83</td>
<td>241</td>
<td>15</td>
<td>22</td>
<td>194</td>
<td>4</td>
</tr>
<tr>
<td>и, жѧ, е, ѣ</td>
<td>49</td>
<td>788</td>
<td>11</td>
<td>12</td>
<td>278</td>
<td>3</td>
</tr>
</tbody>
</table>

i Clustering clitics are calculated as part of the cluster (as well as interrogative ли, monosyllabic adverbs и, ти, тоу), i.e. another clitic, which is part of the cluster, can precede or follow (e.g. приведѣте ми.dat і.acс сьмо ‘bring him to me here’ will be calculated both for the dative ми and for the accusative і as 2P, post-V; рьци оубо намъ чѣто ти.dat сьмо ‘therefore, tell us what you think’ will be calculated for the dative ти as 2P, pre-V).

ii The constituent in 1P can be preceded by a coordinating monosyllabic conjunction such as и ‘and’, а ‘but’, нь ‘but’, etc., and the negation не. Elements after infinitives were calculated accordingly (повелѣ неупатъ принести врѣтиште. й висадити і.acс вѣ не. ‘order the proconsul to bring the back and to put it inside (in-it)’ was 2P, post-V for the accusative і) but periphrastic verb constructions are not included.

iii One or more constituents (noun phrases, prepositional phrases, adverbial phrases, including non-monosyllabic adverbs such as пакъ, пакы) can be placed between the clitic/clitic cluster and the verb.

iv Numbers do not involve periphrastic constructions, first-position elements or prepositional phrases.

If clusterised, the dative ми ‘I.dat’ and ти ‘you.dat’ tend to precede the accusative semi-clitics (and the reflexives), as in (10a-c). This is not true for the clitic-like datives, though, as in (10d).
The placement of the pronominal elements often echoes the respective placement of the parallel elements in the Greek text but with variant readings (especially if there were no corresponding constructions in Greek – Sławski 1946 – such as the reflexive constructions and periphrastic verb constructions). The verbal elements are discussed further in the article; here I will only mention that clitics and semi-clitics demonstrate similar behaviour: both may precede the auxiliary and, by transitivity, also the participle (the active l-participle as well as the passive n/m/t-participles). Both may be found in 2P, in the order X > CL > aux > Participle, where X can be a wh-word, as in (11a), a subjunction such as да ‘to’, as in (11b), or some other functional element in first position.
b. [...] съвѣшташа да и бѫхи оубили CM, Mar. 3,6
 [...] consulted to he.ACC are.aux.3Sg killed
 [...] ἐβουλεύσαντο ἵνα ἀποκτείνωσιν αὐτόν.

Semi-clitic reflexive pronoun сѧ ‘self.ACC’ exhibits variation in its behaviour just like the other semi-clitic pronominals: it may appear after the (reflexive) verb (more often) but also among the clustering elements (more rarely) in a slot of another pronominal and, more specifically, in the slot of the accusative pronominals.

We have already distinguished between ‘true’ clitics (Ми ‘I.DAT’ and ти ‘you.DAT’) and clitic-like pronouns (емоу ‘he.DAT’, etc.) above, but their different status with respect to the cliticisation behaviour also manifests with respect to сѧ ‘self.ACC’: Ми and ти always precedes it, as in (10c), while the clitic-like pronouns typically follow it, as in (10d) above.

If the semi-clitic сѧ ‘self.ACC’, however, is found beside the elements in the core in (5), as indicated by (10c), it may also appear in 2P beside other elements as in (12) where it aspires for a slot within the cluster. In (21) below, we will see another example where сѧ is placed in front of the auxiliary but also after the main verb in compound tenses in a variant reading.

(12)
колико сѧ бы трудили CSp, 49r, 11-12
how-much Reflcl be.aux.cond worked
Πόσα аν еκαμες,
b.

<table>
<thead>
<tr>
<th>Clitic Cluster</th>
<th>Transliteration</th>
<th>Translation</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>седьбо иисъимъшамьже.</td>
<td>seven because they.Gen had she.acc wife.acc</td>
<td>CM, Mk. 12,23</td>
<td></td>
</tr>
<tr>
<td>седьбо</td>
<td>ἑκκαθῆμενμεν.</td>
<td>they acc came to he.at</td>
<td>CZ</td>
</tr>
<tr>
<td>ои γὰρ ἐπὶ ἑσχον αὐτῆν γυναῖκα.</td>
<td>seven because they.Gen had she.acc wife.acc</td>
<td>CM, Mk. 12,23</td>
<td></td>
</tr>
</tbody>
</table>

The inventory of clitics (and clitic-like elements) in OCS texts is rather limited. It is to be expanded in Middle Bulgarian but some forms are already attested in the early texts: for example, contracted forms of clitic-like datives, though very rarely and with variant readings, as in (14).

(14)

<table>
<thead>
<tr>
<th>Clitic Cluster</th>
<th>Transliteration</th>
<th>Translation</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ἱς въ цръкъве съштю придошемъкъ немоу</td>
<td>and in temple being he.dat came to he.at</td>
<td>CM, Mk. 11,27</td>
<td></td>
</tr>
<tr>
<td>ι въ цръкъве ходаштю придошемоу</td>
<td>and in temple being he.dat came to he.at</td>
<td>CM, Mk. 11,27</td>
<td></td>
</tr>
<tr>
<td>και ἐν τῷ ιερῷ περιπατοῦντος αὐτόου ἐρχονται πρὸς αὐτόν</td>
<td>and in temple being he.dat came to he.at</td>
<td>Scrivener 1894</td>
<td></td>
</tr>
</tbody>
</table>

The class of pronominal clitics expands in the Middle Bulgarian texts that I have considered. In LTr, the respective clitics occur in the same positions and order with the addition of third-person accusative ро 'he.acc' and dative мовъ 'he.dat', which behave as 'true' clitics. The earlier forms, however (ерo and емоу), are still found in LTr but in different functions. The accusative clitic рo 'he.acc' is used for direct objects only, while the genitive (genitive-accusative) ерo predominantly expresses the possessive genitive. The dative clitic мовъ 'he.dat' refers to external possessors and is found in 2P, while емоу refers to indirect object (and is found after a verb and after some prepositions such as противъ 'against'). Expectably, ерo and емоу do not cluster.

The clustering elements exhibit the pattern in (5), as illustrated by the examples in (15), with the core (DAT > ACC) in 2P, and the verbal elements placed to its right (in aux2).
In VB, pronominal clitics continue to cluster in the order DAT > ACC (including reflexives). As in (16a), the cluster is placed before the negation particle and can be separated by the verb (although both pronominal clitics and the cluster are already mainly verb-adjacent).
The numbers in table 2 show clitics’ overall preference to second position and pre-verbal placement in VB.

**Table 2** Pronominal clitics in VB

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>миктимо,йи,им</td>
<td>560</td>
<td>43</td>
<td>21</td>
<td>14</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>га/го,ю,их,ме,те,ни,ви</td>
<td>237</td>
<td>18</td>
<td>12</td>
<td>8</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

In VB, a certain number of pronominal clitics are found after a mono-syllabic conjunction such as ‘but’, ‘and’, and after a pause, as well as in second prosodic position where it can split a noun phrase.

(17)

И 

**И** **м** доде слъга моа, да 

and **I.DAT** came servant my to **i.DAT is.aux.3Sg**

купиля що 

bought what **he.DAT am.aux.1Sg** said from all

In the two 17th-century damaskins that are took into consideration, pronominal clitics are also found in 2P but pre-verbal and post-verbal positions are almost equally distributed. Although rarely, the cluster may split a noun phrase, as in (18d). The example in (18e) illustrates the portion of the pattern in (5): DAT ACC AUX2.

(18) a.

за_единь чар_мъсеч 

for_one hour **he.DAT.ReflCL** eradicated power.def and courage.def

b.

и Оста́ви 

and leave **ReflPossCL it.acc** on eyes.DEF and on heart.DEF

c.

á любовь не имает ннцо 

and love **neg have nothing i.DAT ReflCL neg** lend
d. И тоги́ва áрмъ мъ се гъвъ и въ́ DTr and then angel he.dat ReflCL of-God showed up

вгъръ на́д ние́го from-above over he.acc

e. [...] амъ си́ ўзми́ сво́та кръ́па, дето́ DTr

[...] but ReflPossCL take own towel that

мѣ́ а си́ д´ла на д´льь
1.DAT she.acc is.aux.2Sg given on deed

<table>
<thead>
<tr>
<th>Table 3 Pronominal clitics in DTr and DLv</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
</tr>
<tr>
<td>ми, ти, моу, й, им...</td>
</tr>
<tr>
<td>ме, те, нь, вы...</td>
</tr>
<tr>
<td>го, я, гьи, ´...</td>
</tr>
</tbody>
</table>

To sum up, the inventory of the clitical pronouns that are found in second position and pre-verbally increases in the history of Bulgarian, mainly as a result of the changes in the inventory of pronominal elements. The accusatives го ‘he.acc’, я ‘she.acc’, гьи ‘they.acc’ and datives мой ‘he.dat’, й ‘she.dat’ have joined the cluster after the re-analysis of the earlier forms of anaphoric pronouns.

3.3 Verbal Clitics and Clitic-Like Elements

Both in the early and in later texts, the behaviour of most monosyllabic verbal elements such as be-auxiliaries/copulas, future auxiliaries щѫ/ще ‘will/shall’, is ambiguous. They may be found in first position, including immediately after negation, and preceding the discourse clitics – as in (19), but also in 2P, within a cluster and in second prosodic position, splitting a noun phrase, as in (19b) – in parallel to the orders in the Greek text.
Auxiliaries are found either to the left, or to the right of the pronominals among the clustering elements, as exemplified by the periphrastic verb constructions in (20).

The positions of the reflexive and the auxiliary often vary, with the former more often found in 2P, while the latter is adjacent to negation, as in (21).
In Middle Bulgarian, fluctuations in the positions of auxiliaries remain, as exemplified in (22) from LTr.

(22) a.

и поиди против Ектору кралю, да

and came against Hector king to

би сѧ оставилъ гръцкьі воискьі

would left Greek army

b.

 [...] и метн преѓ фарижа егова, 

[...] and threw in-front-of Paris his

dавно сѧ би възвратилъ

long-ago would come back

In VB, monosyllabic auxiliaries are mostly in AUX2, i.e. to the far right, where the latest clustering elements were joining, presumably, as in (23a-b), with 6 examples of the reverse order of which 5 are with смо ‘are.1pPl’ and one – with the conditional auxiliary би ‘would’ and negation particle – see (23c) (and 13 instances are with future auxiliary).

(23) a.

 [...] ере не ми е ўзел вражмаш господства ми

[...] that neg i.dat is.aux.3sG taken enemy master my
The trend continues into the Early Modern Bulgarian with auxiliaries placed on the far right (including the present-tense ones), as in (24a-c), in contrast to the situation in present-day standard Bulgarian (compare also their positions in the dialects according to Antonova-Vasileva et al. 2016, map 144 where the majority of dialects instantiates the order aux > dat / acc). There are again single instances of verbal clitics found before the pronominal ones, as in (24c). The order with respect to the reflexive, as in (24c), is the same as in present-day standard Bulgarian.

(24) a.
Защо́ го́ си́ мъчил, и защо́ го́ си́ испъ́дил д́ Tr why he.acc is.aux.2Sg tortured and why he.acc is.aux.2Sg driven away

b.
акъ́ въ́ съ́ стои́вале пръ́ во́ та_e_ ça nále DLv if_are.aux.3pPl defended rightly then_it.aux_are.aux.3pPl eaten

c.
Затовá_ ça_e ве́ ъ е й_ пока́ніе [...] DLv therefore_Refcl_ is.aux.3pSg already also_repentance [...] 
вe_ ça_ce ве́ ъ е й_ ра́йсъ́ кie [...] therefore_are.aux.2pPl_Refcl already also_heavenly [...]

Negation may separate clustering elements, with the reflexive to the left and the auxiliary to the right to form two prosodic words (whose constituents are linked via ‘‘’), as in (25a) (see also (24e) with the possessive reflexive). Pronominal clitics are in second position while auxiliaries are verb-adjacent, as in (25b-c).

(25) a.

Тъ_еще_се_нѣ̀_добрѣ̀ [тѫка_намѣстил]ь [...]. DLv

b.

За́мѣндогдѣ́_ви_нѣ̀ прѣва́рила сѣ́бѣ́ сымрѣ̀нѣ́ DLv

for until_you.dat before neg_is.aux.3pSg overtook sword deadly

c.

и_догдѣ́_са_радѣ́рь нѣ̀ рѣ́здѣ́лильь. DLv

and_until_Reflcl market neg_is.aux.3pSg broken

**Table 4** gives the order of clustering be- and future-auxiliaries (the latter in Middle Bulgarian and Early Modern Bulgarian) and pronominal clitics, semi-clitics, and clitic-like elements with the periphrastic verb constructions. The numbers attest for variation in the placement of auxiliaries among the clustering elements with preference to the right slot aux2 of the pattern in (5). This is also true for the future auxiliary (the form щѫ was attested very early: Mirčev 1956, 202; Haralampiev 2001, 149). In the texts here, the position of щѫ is also ambivalent (in aux1 and aux2), but it is negation-adjacent and also found in second prosodic position.
Table 4 Pronominals and auxiliaries with periphrastic verb constructions

<table>
<thead>
<tr>
<th></th>
<th>Pron &gt; Aux</th>
<th>Aux &gt; Pron</th>
<th>Non-adjacent</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCS</td>
<td>53</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>MB</td>
<td>143</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>EMB</td>
<td>103</td>
<td>25</td>
<td>4</td>
</tr>
</tbody>
</table>

The clitic cluster in the noun phrase is left outside the scope of the present article but its phrasal position (2P) and order is very similar to that of the clausal cluster, with the earliest clitics to the left and the latest to the right (for comparison with Greek, see Eckhoff 2018). This cluster may not be considered viable in present-day Bulgarian if we assume that the definite article is an affix. The presumed pattern of the noun-phrase clitic cluster might have been DISC > DEM/ART > POSS/GEN/DAT > (AUX). There are very early examples of demonstrative pronouns that are in second position in the noun phrase and immediately follow a nominal element in first phrasal position: these are exactly the demonstratives that may be interpreted as clitic-like articles (Dimitrova-Vulchanova, Vulchanov 2012). However, there are also examples of discourse markers followed by demonstrative pronouns, clusterised in the same phrasal second position, which attests that the reanalysis has not been completed yet (as per Kurz 1963).

4 A Tentative Conclusion

The data that have been discussed in the article backs the assumption that, historically, there might have been no significant changes in the position of the elements in the core of the primary clitic cluster, i.e. the pronominal clitics, with variations happening (and still) among the elements in the verbal periphery. The earliest 2P clitics were lost, somewhat expectably, because they were non-paradigmatic unlike those that have remained. The set of elements filling the respective slots among the clustering elements was expanded as a result of the restructuring within the pronominal and verbal systems. Similar developments might have been at play for the (partial) loss of the noun-phrase cluster. A more thorough analysis of the variant readings across the monuments and constructions (matrix vs. subordinate clause, clause types, absolute constructions) is due in order to highlight the numerous intricacies of cliticisation and clitic patterns in the history of Bulgarian.
Bibliography


Tsvetana Dimitrova
On the Diachrony of the Clitic Cluster in Bulgarian


