

Tania Avgustinova

## Clustering Clitics in Bulgarian Nominal Constituents\*

The present study of Bulgarian nominal morphosyntax focuses on the placement of the short possessive pronouns and the interrogative particle within the nominal syntactic domain, in correlation with the distribution of the definite article. In other words, I am primarily interested in what might be called "proper" nominal clitics, while predicative clitics that are introduced in the NP by attributively used participles or accompany deverbal nouns remain beyond the scope of this article. After presenting some preliminary assumptions, I will show that a treatment of Bulgarian definite article based on HALPERN 1995 is quite promising, especially if augmented to adequately cover the relevant data. Setting thus the appropriate context, I will concentrate on the distribution of Bulgarian possessive enclitics which supplies a strong evidence in favour of admitting the existence of a nominal enclitic cluster. I will show that their placement obeys the same mechanism which governs the distribution of definiteness marking in Bulgarian NPs. Finally, the non-verbal interrogative particle, which on certain conditions may join the nominal enclitic cluster, will be considered with respect to its scope interpretations.

Let me begin with terminological issues concerning the *nominal syntactic domain* in Bulgarian. An important background assumption in my approach is that the term *NP* (noun phrase) can adequately be used for any syntactic constituent which is headed by a nominal category. In particular, the collocation of a preposition and a noun is regarded as an NP marked by the preposition rather than as a prepositional phrase. Consequently, the prepositions are interpreted as phrase markers occurring as the leftmost elements of the respective NPs. Since Bulgarian nouns lack morphological case marking, the use of prepositions as syntactic case markers is very extensive in the language<sup>1</sup>. Nevertheless, the category of *syntactic case* has to be incorporated in the language description because it allows for expressing important generalisations. I assume that non-prepositional NPs exhibit *nominative* syntactic case if they function as subjects, subject predicatives<sup>2</sup> or subject appositions<sup>3</sup>, and *accusative* syntactic case if they

---

\* In: *Proceedings of The 2nd International Conference on Formal Description of Slavic Languages (FDSL-2)*, Potsdam 1997.

<sup>1</sup> Old-Bulgarian non-prepositional cases - with the exception of the subject and the direct object ones - are also expressed by prepositions in contemporary Bulgarian: thus, the old dative and genitive are expressed by the preposition *na* 'to/of', while the old non-prepositional instrumental - by the preposition *s* 'with'.

<sup>2</sup> E.g., Toj e lekar. 'He is a doctor.'

<sup>3</sup> E.g., Asen, našijat star prijatel, pristigna. 'Asen, our old friend, arrived.'

function as direct objects or (accusative) experiencer objects<sup>4</sup>. Even though prepositionally marked NPs are most naturally thought of as exhibiting a generalised *prepositional* syntactic case, there are two important exceptions. Within the clausal syntactic domain, indirect objects as well as some experiencer objects<sup>5</sup> are more adequately interpreted as exhibiting *dative* syntactic case, even though they are always marked by the preposition *na* in contemporary Bulgarian<sup>6</sup>. Also, in the nominal syntactic domain, the *na*-NP modification indicating possession / pertainment, which has diachronically evolved from a genitive attribute<sup>7</sup>, is functionally different from other *na*-NP modifiers in which the same preposition indicates, e.g., location<sup>8</sup>. Therefore, I assume that they occur in *possessive* (rather than prepositional) syntactic case<sup>9</sup>. In terms of syntactic cases, generalisations can also be stated about Bulgarian cliticisation. Thus, for example, in the clitic pronominal system we find only accusative, dative and possessive forms which directly correlate with syntactically accusative, dative

<sup>4</sup> E.g., *Moja prijatelj go trese*. 'My friend shivers (has fever).'

<sup>5</sup> E.g., *Na decata im omrâzna da čakat*. 'The kids are fed up with waiting.'

<sup>6</sup> This special status has different motivations, one of them being the fact that such *na*-NPs can be replicated ("doubled") by a dative clitic, as in (i). NPs where the preposition *na* is not a dative-case marker (but indicates, e.g., location) are not replicable - cf. (ii) vs. (iii), and hence are assumed to be in prepositional syntactic case.

(i) *Na decata im dadoxa bonboni*.  
to children\_def.art them-DAT-CL gave-3pl candies  
'They gave candies to the children.'

(ii) *Na masata postavixa vaza s cvetja*.  
on table\_def.art put-3pl vase with flowers  
'They put a vase with flowers on the table.'

(iii) \**Na masata í ostavixa vaza s cvetja*.  
on table\_def.art her-DAT-CL put-3pl vase with flowers  
(ungrammatical on the intended interpretation)

<sup>7</sup> In the Balkan language environment, it is a general tendency for (historically) genitive and dative attributes to merge in form - ASENOVA 1989.

<sup>8</sup> While a pertainment NP is replicable by a possessive clitic which is homonymous with the dative one, as in (i), a locative *na*-NP modifier is not - cf. (ii) vs. (iii).

(i) *Tam beše samo kolata mu na Asen*.  
there was-3sg only car\_def.art his-POSS-CL of Asen  
'Only Asen's car was there.'

(ii) *Tâlpata na brega postojanno narastva*.  
crowd\_def.art on shore\_def.art steadily grow-3sg  
'The crowd on the shore keeps growing.'

(iii) \**Tâlpata mu na brega postojanno narastva*.  
crowd\_def.art his-POSS-CL on shore\_def.art steadily grow-3sg  
(ungrammatical on the intended interpretation)

<sup>9</sup> A further distinctive property of possessive-case NPs is that they cannot be headed by a pronominal (i). The relevant pronominalisation is to be expressed by the corresponding full possessive pronoun which functions as an adjectival modification (ii).

(i) \**Tam beše samo kolata na nego*.  
there was-3sg only car\_def.art of him

(ii) *Tam beše samo negovata kola*.  
there was-3sg only his-POSS-FULL\_def.art car  
'Only his car was there.'

and possessive NPs distinguished above. Moreover, the phenomenon of clitic replication (cf. AVGUSTINOVA 1997) involves exactly these sets of morphosyntactic objects.

The term *nominal clitic* is intended to refer to those clitic elements whose syntactic domain is the Bulgarian nominal constituent. All prosodically weak prepositions are *nominal proclitics*. They form an accentual unit with the immediately following stress-bearing item and are, thus, prosodically hosted by it. Bulgarian nominal proclitics can be characterised as *simple clitics*. Recall that a simple clitic, in the terminology of ZWICKY 1977, is an element of some basic word class and appears in a position in which the rules of the syntax would (or at least could) put it, i.e. it occupies the normal syntactic position for a non-clitic word of its category. The positioning of clitic prepositions within the NP they mark is just the same as that of the orthotonoc ones (like, e.g., *vmEsto* 'instead of', *osvEn* 'apart from', *posrEdstvom* 'through', *sâglAsno* 'in accordance with', etc.<sup>10</sup>); and the function of a phrase marker requires that a preposition (of whatever prosodic nature) occupies the leftmost ("left-edge") position in the NP. More precisely, a prepositional phrase marker can only be preceded by the negation particle *ne*<sup>11</sup> and by certain adverbials, e.g., *dori* 'even', *predimno* 'mostly', *samo* 'only', etc., which have focusing effect on the entire NP. As to Bulgarian *nominal enclitics*, they are clearly interpretable in Zwicky's terminology as *special clitics*, and will be in the main topic of interest here.

Being a morpheme, Bulgarian definite article exhibits a clitic-like behaviour. Although it qualifies the NP as a whole, its actual form is determined by the word it is attached to. I will take under scrutiny the interpretation of this element in the theory of HALPERN 1995 and, after performing the necessary confrontation and verification with relevant data, reconsider certain insightful aspects of the analysis proposed therein.

With respect to the nature of Bulgarian definite article Halpern makes the following generalisation on which his further analysis relies:

"...the definite article is not a true second position clitic, but rather an inflection on the head of the first constituent of the NP."

This statement, however, can only cover the cases where the first constituent of a definite NP is, e.g., a (bare) noun, an adjectivally declined word (possibly with modifiers of its own), a numeral, or a quantifier allowing an articulated form<sup>12</sup>. Apart from coordination of the above-listed constituents (where it is rather unclear what a head would be), problems arise in particular with the possessive / pertainment *na*-NP modifier when it is preposed with respect to the lexical

---

<sup>10</sup> The stressed vowels are marked by capitalisation.

<sup>11</sup> The negation particle *ne* is a stress-bearing item - not a clitic - whenever it specifies (i.e. scopes over) a non-verbal constituent, and I will not be concerned with it here.

<sup>12</sup> E.g., *vsičkite* *starinni knigi* 'all\_def.art old books', *mnogoto / malkoto / dostatačno* *arhitekturni pametnici* 'many\_def.art / few\_def.art / enough\_def.art architectural monuments'

head<sup>13</sup>. Even though in such cases it would in fact be the "first constituent of the NP", its contingent definite inflection does not indicate definiteness of the whole nominal constituent but rather just the fact that the possessive-case NP modifier is definite. Hence, the definiteness of the whole nominal constituent has to be realised on the next appropriate sub-constituent. - cf. (1).

- (1) a. na profesora kolata  
of professor\_def.art car\_def.art  
'professor's car'
- b. na profesora novata kola  
of professor\_def.art new\_def.art car  
'professor's new car'
- c. na profesora dvete novi koli  
of professor\_def.art two\_def.art new car  
'professor's two new cars'

The actual analysis is based on distinguishing trigger features and marking features. Their distribution is regulated by the Marking Feature Principle (MFP) which is formulated by Halpern as a general mechanism regulating trigger-marking feature interaction:

For a trigger-marking feature pair T and M, where T is specified as a LEFT or RIGHT trigger, and M as a FIRST, HEAD or LAST marking feature:

- a. If T, a LEFT/RIGHT trigger, appears on a mother node, M must appear on the first/last daughter.
- b. If M, a FIRST/HEAD/LAST marking feature, appears on a mother node, it must also appear on the first/head/last daughter.
- c. If M appears on a daughter node, either M or T (or both) must appear on the mother node.

The parameterisation of this principle for the distribution of Bulgarian definite article would involve a LEFT trigger feature DEF (a syntactic property of being **definite**) and a HEAD marking feature DM (a morphological property of having the definite suffix, i.e. **definiteness marking**). The former is associated with the top node of a definite NP while the latter - with the element on which the definiteness is inflectionally realised. In other words, a syntactic property of the entire constituent is treated as being responsible for triggering certain morphological property on one of the sub-constituents. Needless to say, if the analysis presupposes acceptance of the generalisation cited above, it is doomed to fail in (1). Yet, a remark in a footnote on the same page indicates Halpern's

---

<sup>13</sup> As a matter of fact, any preposing of a *na*-NP generally presupposes definiteness of the nominal constituent in which it occurs - cf. the ungrammatical

\*na profesora kola  
of professor\_def.art car

awareness of the potential problems:

"... Two sorts of constituents can [...] be "outside" of the scope of DEF. Specifically, certain quantifiers in Rumanian, which are initial in the NP, don't count as the first constituent, and in Bulgarian and Macedonian it is possible, with appropriate context and intonation, to front a prepositional phrase (including a possessive) to the beginning of an NP; again the placement of the article acts as if this fronted constituent were not present. We can account for these two patterns by assuming that the quantifiers and PPs are outside of the constituent specified DEF..."

The idea to restrict the scope of DEF so that it would not include the prepositionally marked possessive-case pre-modifier of a noun is obviously legitimate, and my proposal is to reformulate the underlying generalisation in the following way: *Bulgarian definite article is an inflection on the head of the leftmost major prepositionally unmarked constituent of the NP*. Provided Bulgarian prepositions are considered phrase markers, and hence, a minor category (as I am assuming), the wording "major prepositionally unmarked constituent" would naturally refer to non-prepositional NPs, APs, AdvPs, as well as (trivially) to the lexical head itself. As to clausal modifiers, they are exempted from this generalisation since they never precede the lexical head, and thus do not occur as the leftmost constituent of an NP. So, in order to make the MFP work also for (1), it would suffice to just specify the notion of *first daughter* accordingly<sup>14</sup>.

With this revision made, let us now consider how the core of Halpern's analysis will work. The distribution of the definite article depends on two conditions relating DEF and DM, with the definite article morpheme being the inflectional realisation of the specification DM. The first condition states that the presence of the specification DEF on a mother node, X, requires the presence of DM on the leftmost major prepositionally unmarked daughter. The intended effect is triggering of definiteness marking (DM), as illustrated in Figure 1.

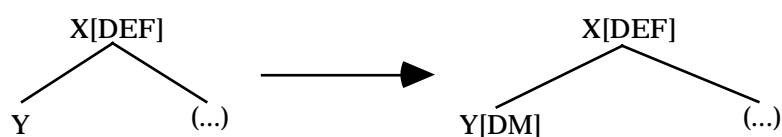


Figure 1

The second condition simply states that the presence of the marking feature DM on a mother node requires the presence of DM on the head daughter. The intended effect is percolation of the definiteness marking as shown on Figure 2.

<sup>14</sup> Articling of conjoined adjectival modifiers requires further adjustments of this principle in order to account for, e.g., plain adjectival coordination vs. NP-coordination via coordinated adjectival modifiers, as well as for certain paratactic phenomena. For lack of space, the topic will not be pursued in this article.

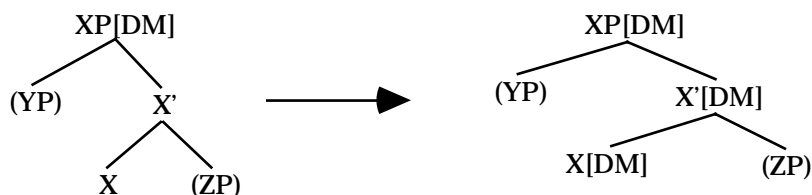


Figure 2

Both conditions force the head of the leftmost major prepositionally unmarked constituent of the NP to be definiteness-marked; inflectional morphology determines the forms of words which are so specified. When demonstratives, which are assumed to be their own definite forms, are initial in a definite NP, the first condition will force them to appear in their definite form, but this will not result in any visible change. However, when they are not initial (e.g., if the demonstrative pronoun is preceded by a quantifier) the definiteness marking passes onto another element, as illustrated in Figure 3.

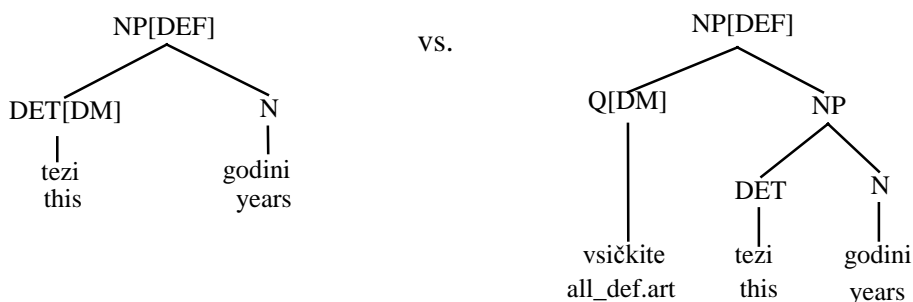


Figure 3

Now, my main claim is that the possessive clitic's placement in the syntactic NP-domain correlates directly with the distribution of the definiteness marking. Usually this is the definite article, as (2) illustrates, but in literary style possessive clitics are also found in NPs where definiteness is expressed by a demonstrative pronoun - cf. (3).

- (2) a. *râkopolite im*  
 manuscripts\_def.art their-POSS-CL  
 'their manuscripts'
- b. *poslednite im râkopolite*  
 last\_def.art their-POSS-CL manuscripts  
 'their last manuscripts'
- c. *dvata im posledni râkopolite*  
 two\_def.art their-POSS-CL last manuscripts  
 'the two last manuscripts of ours'
- d. *mnogoto im nepublikovani râkopolite*  
 many\_def.art their-POSS-CL unpublished manuscripts  
 'the numerous unpublished manuscripts of theirs'

- e. **gotovite im** za pečat râkopisi  
 ready\_def.art their-POSS-CL for print manuscripts  
 'their manuscripts that are ready for print'
- (3) a. **tazi im** slabost<sup>15</sup>  
 this their-POSS-CL weakness  
 'this weakness of theirs'
- b. V **tova si** nastroenie toj beše izključitelno štedâr.  
 in this REFL-POSS-CL mood he was extremely generous  
 'In this mood he was extremely generous.'

The idiosyncrasy of many kinship terms with respect to definiteness is also adequately accommodated if one regards their singular form used with no adjectival pre-modifiers as inherently definite. This inherent definiteness is manifested by the fact that a possessive clitic may immediately follow the kinship term. In reality, any use of short clitic forms of Bulgarian possessive pronouns presupposes syntactic definiteness. If, however, a kinship term is modified by an adjectival pre-modifier, the latter must bear the definiteness marking (cf. Figure 4) in order for the possessive clitic to be used.<sup>16</sup>

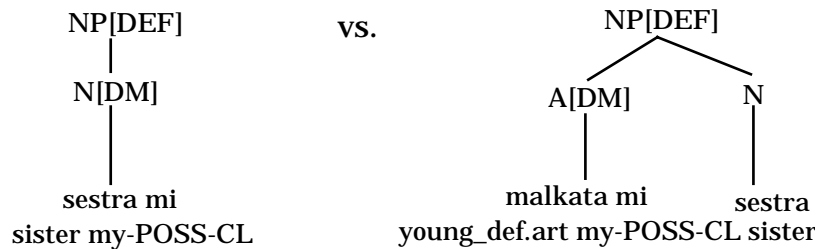


Figure 4

Summing up, a clitic possessive pronoun, which must always follow a morphologically definite word, can occur only in three environments:

- adjacent to the definite article morpheme, i.e. immediately following the articulated word in a definite NP;
- following a demonstrative (a rather special case);

<sup>15</sup> Note that the variants with full (non-clitic) possessive pronouns are completely parallel:

- (i) **tazi tjaxna slabost**  
 this their-FULL weakness
- (ii) **v tova svoe nastroenie**  
 in this REFL-POSS-FULL mood

<sup>16</sup> If the same noun is used either in the plural (i) or not as a kinship term (ii), the possessive clitic has to be licensed by a definite article on the head noun, also if there are no pre-modifiers to it.

- (i) **sestrite mi**  
 sisters\_def.art my-POSS-CL  
 'my sisters', also possible: 'my nurses'
- (ii) **sestrata mi**  
 nurse\_def.art my-POSS-CL  
 'my nurse'

- immediately after a non-articled kinship term in the singular with no attributive (inflected) modifiers preceding it.

Both the definiteness and the "possessed status", indicated by the possessive clitic, have to be interpreted as pertinent to the head noun and, thus, relevant for the entire NP, not, e.g., merely for the inflected modification they would morphologically (in the case of the article) and prosodically (in the case of the possessive enclitic) attach to, whenever the lexical head has the appropriate type of pre-modifiers.

In order to capture the observed regularities, I view the morphonological ensemble inherent definiteness / definite article morpheme + possessive enclitic as a **nominal enclitic cluster** (NEC) whose positioning in the syntactic NP-domain is governed by a special mechanism, the MFP.

It is important further to note that short possessive pronominal forms function as phrasal inflection on NPs by supplying *index* information about the possessor: its person, number, and gender. In order to integrate this observation in the present analysis, I introduce a pair of dedicated features consisting of a trigger feature PI (a property of being inflected for the possessor's index - **possessor inflection**) and a corresponding marking feature PE (a property of enclitically attaching a short possessive pronoun - **possessive enclisis**). The syntactic property of being definite (DEF) is hierarchically superior with respect to the property of being possessor-inflected (PI). Thus, possessor inflection on an NP always presupposes definiteness, but it is not true that all syntactically definite NPs must have possessor inflection. The actual distribution of Bulgarian possessive enclitics in the nominal syntactic domain depends on two conditions relating PI and PE. As expected, these conditions are based on (and derived from) those relating DEF and DM. The first condition states that the presence of a PI specification on a mother node that is marked for DEF requires the presence of PE on the daughter bearing the specification DM. The second condition simply states that the marking feature PE always cooccurs with DM. Such feature interdependence, in essence, models the observed clustering effect.

Let us finally consider one more item which can join the NEC. The non-verbal interrogative particle *li* always (immediately) follows the questioned (sub)constituent it scopes over. If this happens to be the element hosting the NEC, the enclitic *li* joins it as the very last element, following the contingent possessive clitic. A NEC with the shape definite article morpheme + possessive enclitic + *li* is illustrated in (4), and a NEC with the shape possessive enclitic + *li*, hosted by an inherently definite kinship term, in (5). In (4a), as well as in (5), the interrogative particle scopes over the head noun, while in (4b-e) the situation is different - only the modifier that is followed by the interrogative particle is in its scope, namely, the adjective in (4b-c), the numeral in (4d), and the quantifier in (4e). It is obvious that the status of a *simple clitic* can legitimately be claimed for this nominal enclitic.



- (4) a. Râkopisite            **im**            **li** izčeznaxa ?  
 manuscripts\_def.art their-POSS-CL Q disappeared-3pl  
 'Is it their manuscripts that disappeared?'
- b. Poslednite    **im**            **li** râkopisi    izčeznaxa ?  
 last\_def.art    their-POSS-CL Q manuscripts disappeared-3pl  
 'Is it the last ones from their manuscripts that disappeared?'
- c. Gotovite       **im**            **li** za pečat    râkopisi       imate predvid ?  
 ready\_def.art    their-POSS-C    Q for print    manuscripts    have-2pl in-view  
 'Is it the ready-for-print ones of their manuscripts that you mean?'
- d. I dvata         **im**            **li** posledni    râkopisa       izčeznaxa ?  
 and two\_def.art    their-POSS-CL Q last            manuscripts    disappeared-3pl  
 'Did both of their last manuscripts disappear?'
- e. Vsičkite       **im**            **li** nepublikovani    râkopisi       izčeznaxa ?  
 all\_def.art    their-POSS-CL Q unpublished    manuscripts    disappeared-3pl  
 'Did all of their unpublished manuscripts disappeared?'
- (5) Brat            **mu**            **li** šte    dojde ?  
 brother        his-POSS-CL    Q will    come-3sg  
 'Is it his brother who will come?'

Bulgarian interrogative particle *li* has not been investigated as a non-verbal clitic yet<sup>17</sup>, also in the related more or less traditional literature no special attention is paid to this remarkable item of the nominal syntactic domain. So let us consider its possible placements within the NP, and the resulting interpretations. We have already seen how the interrogative particle joins the NEC by virtue of being hosted by the same prosodically strong element. In (4b-e) this results in non-trivial "crossing dependencies". The definite article and the possessive clitic, as already mentioned above, have to be interpreted as related to the head noun, inasmuch as they indicate the definite and the possessed status of the entire NP, while the interrogative particle is related to (and scopes over) a sub-constituent of this NP, i.e. just to the respective pre-modifier of the head noun.

In general, I assume that *li* can have a **wide scope** if it ranges over the NP as a whole, and a **narrow scope** if it ranges over a sub-constituent of this NP. The wide scope is possible only with "right-edge" positioning of this enclitic in the nominal syntactic domain, as in (6a). In the same alignment, a narrow scope of *li* over the contingent prepositional post-modifier is also conceivable. This, however, requires a special intonation pattern with an emphatic stress on the sub-constituent that is immediately followed by *li* and, hence, is to be interpreted as being in its scope - cf. (6b). Note that in both cases *li* is hosted by the rightmost prosodically strong element of the NP. On the other hand, the interrogative

---

<sup>17</sup> For its behaviour as a verbal clitic cf. AVGUSTINOVA 1994.

particle has a clearly and unambiguously narrow scope in all examples in (7).

- (6) a. [Dve izvânredno interesni knigi za moreplavатели li] imaš ?  
[two extremely interesting books about seamen Q] have-2sg  
'Is it two extremely interesting books about seamen what you have?'
- b. Dve izvânredno interesni knigi [za MOREPLAVATELI li] imaš ? <a ne za piloti>  
two extremely interesting books [about SEAMEN Q] have-2sg  
'Is it seamen <not pilots> that you have two extremely interesting books about?'
- (7) a. [dve li] izvânredno interesni knigi za moreplavатели <a ne pet>  
[two Q] extremely interesting books about seamen <and not five>  
(*li* scopes over the cardinal)
- b. dve [izvânredno li] interesni knigi za moreplavатели <t.e. nad srednoto nivo>  
two [extremely Q] interesting books about seamen <i.e., above the average>  
(*li* scopes over the adverb)
- c. dve [izvânredno interesni li] knigi za moreplavатели <a ne prosto aktualni>  
two [extremely interesting Q] books about seamen <and not just up-to-date>  
(*li* scopes over the adjectival phrase...)
- d. dve izvânredno [INTERESNI li] knigi za moreplavатели <a ne modni>  
two extremely [INTERESTING Q] books about seamen <and not fashionable>  
(...or, alternatively, only over the adjective)
- e. dve izvânredno interesni [knigi li] za moreplavатели <a ne albumi>  
two extremely interesting [books Q] about seamen <and not albums>  
(*li* scopes only over the lexical head)

My intention in this article was to show that distinguishing clitics with respect to the syntactic domain they belong to could be crucial for modelling their placement and behaviour. There are two clitics-relevant syntactic domains in Bulgarian: the clause with its verb complex, and the NP. The sets of clitics belonging to these domains are different. I have concentrated on proper nominal clitics, investigating them in the light of clitic-cluster formation.

## References

- ASENOVA, P. (1989), *Balkansko ezikoznanie. Osnovni problemi na Balkanskija ezikov sâjuz*. Sofia: Nauka i izkustvo.
- AVGUSTINOVA, T. (1994), On Bulgarian Verbal Clitics. *Journal of Slavic Linguistics* 2(1): 29-47.
- AVGUSTINOVA, T. (1997), *Word Order and Clitics in Bulgarian*. Doctoral dissertation, Universität des Saarlandes.
- HALPERN, A. (1995), *On the Placement and Morphology of Clitics*. CSLI Lecture Notes. Cambridge University Press.
- ZWICKY, A. (1977), *On Clitics*. Indiana University Linguistics Club, Bloomington.