Chapter 1

The Structure of Bulgarian

1.1. Introduction

Before undertaking the analysis of binding phenomena in Bulgarian, which is the main topic of this study, a brief description of the basic facts of Bulgarian sentence structure and word order will be given. In this chapter, the basic word order patterns and the major categories of the functional domain of the Bulgarian clause will be discussed. I will also present the structure of the Bulgarian DP and finally argue that Bulgarian is head-initial.

Following our analysis for the Bulgarian word order in Van Gelderen and Grozeva (1995), I assume a minimum of two CPs, the most upward hosting the [+wh] feature and the other (or the lowest in the stack) - the focused/topicalized/discourse material. The SVO neutral order in Bulgarian is derived by raising the subject first to SpecTP, then to SpecAspP and the the verb to the T and Asp heads, respectively.¹ These movements are motivated by the strong N-features of the T and Asp heads in Bulgarian. Thus, despite the complexity and intricacy which the word order of the Bulgarian sentence show at first glance, the basic patterns of these alternations can be accounted for in the minimalist framework (Chomsky 1993, 1995).²

¹ TP and AspP stand for Tense Phrase and Aspectual Phrase. Accordingly, these phrases contain the Tense (T) and Aspect (Asp) heads. In the preminimalist research, IP was divided into AgrP and TP (see Kayne 1989, Pollock 1989, Chomsky 1991).
² The investigation of all variations which the Bulgarian word order show needs a more keen and detailed study. I have tried to delineate the framework of this probe as a background of analyzing the binding phenomena in this language.
In Section 1.2., the surface word order patterns are presented. Section 1.3. deals with subject-verb asymmetries. Section 1.4. addresses the impact of structural focus on Yes-No questions. Section 1.5. defines the boundaries of the clitic cluster. In Section 1.6., I present some arguments that Bulgarian is a head-initial language. On the basis of the core empirical data presented in the previous sections, the structure of the Bulgarian DP and the structure of the Bulgarian clause are discussed.

1.2. Surface Word Order Patterns

In Bulgarian, like in Russian and other Slavic languages, the surface orderings of the elements in the sentence are quite free. In a simple sentence like (1), all orderings of the three constituents are possible.

(1)  
   a. Ivan nameri knigata  (SVO)  
       Ivan found  book-the  
       “Ivan found the book.”
   b. Nameri Ivan knigata  (VSO)  
       found  Ivan book-the
   c. Nameri knigata  Ivan  (VOS)  
       found  book-the Ivan
   d. Ivan knigata nameri  (SOV)  
       Ivan book-the found
   e. Knigata nameri Ivan  (OVS)  
       book-the found Ivan
   f. Knigata Ivan nameri  (OSV)  
       book-the Ivan found

Practically all the various orderings in (1a) - (1f) are considered to be natural answers to the questions in (2), unless a particular constituent is brought into the speaker’s attention by means of contrastive stress:

(2)  
   a. Kakvo stana?/Kakvo se sluči?  
      “What happened?”
   b. Kakvo pravi Ivan?  
      “What is Ivan doing?”

6
(1a) has a word order which is related to the neutral intonation, without any given context. In (1b), knigata (book-the) bears contrastive stress and Nameri Ivan (found Ivan) is the topic. In (1c), Ivan represents the focus and Nameri knigata (found book-the) is the topic. In (1d), the focus is either on knigata (book-the) or on nameri (found). In (1e), nameri (found) is in the focus and in (1f), the focus falls on either Ivan or nameri.

The picture becomes even more complex when a structure like (1) allows various intonations. For (1a) we observe the following numerous alternatives.

(3)   a. IVAN nameri knigata
       “IVAN found the book.”/ “It was Ivan who found the book.”
   b. Ivan NAMERI knigata
       “Ivan FOUND the book.”/ “Ivan found the book at last.”
   c. Ivan nameri KNIGATA
       “Ivan found the BOOK.”/ “It was the book that Ivan found.”

The alternatives in (3) are phonologically and interpretively distinct.3 That is, each of them implies something different concerning the point of information which is considered most relevant and important from the speaker’s point of view. Thus emphasis is put on one constituent, backgrounding another. Each position of an accent in (3) identifies a focus (Rochemont 1986, Rochemont and Culicover 1990). The assignment of focus is in part a function of well-formed discourse. The focus of a sentence is predictable in terms of properties of the discourse and context that the sentence occurs in, as well as of the sentence itself. In our case, the relevance of focus and accent to the interpretation of the sentences in (3) is reflected in judgments of well-formedness for question/answer pairs. For example, of the various pronunciations of (1a) given in (3), only (3a) qualifies as a well-formed response to the question in (4):

(4)    Koj nameri knigata?
       “Who found the book?”

---

3 The examples in (3) are also not equal with respect to truth conditions since focus can affect truth conditions. Thanks to J. Nerbonne for pointing this matter to me.
Similarly, questions (5) and (6) are also considered as well-formed discourse question-answer sequences for the sentences in (3b) and (3c), respectively:

(5) Potůrsi li Ivan knigata?
   “Did Ivan look for the book?” / “Has Ivan looked for the book?”

(6) Kakvo nameri Ivan? / Kakvo tūrsi cjal den Ivan?
   “What did Ivan find?” / “What has Ivan been looking for all day long?”
   The pairs will be repeated below in (7) - (9) for clarity of presentation:

(7) a. Q: Koj nameri knigata?
    “Who found the book?”
    A: IVAN nameri knigata
    “IVAN found the book.” / “It was IVAN who found the book.”
   b. A: *Ivan NAMERI knigata
   c. A: *Ivan nameri KNIGATA

(8) a. Q: Potůrsi li Ivan knigata?
    “Did Ivan look for the book?”
    A: Ivan NAMERI knigata
    “Ivan FOUND the book” / “Ivan found the book at last”
   b. A: *IVAN nameri knigata
   c. A: *Ivan nameri KNIGATA

(9) a. Q: Kakvo nameri Ivan?
    “What did Ivan find?” / “What has Ivan found?”
    A: Ivan nameri KNIGATA
    “Ivan found the BOOK” / “It was the book which Ivan found”
   b. A: *IVAN nameri knigata
   c. A: *Ivan NAMERI knigata

Only the answers in (7a), (8a) and (9a) are eligible for the respective questions. The (b) and (c) answers in (7) - (9) cannot be used as appropriate ones in the relevant discourse.

Summing up, we can apply exactly the same line of reasoning for all the other orderings in (1b) - (1f). Again the constituent which bears the
accent will be in focus. The interpretation of the resulting numerous cases can be reduced to the sentences in (7) - (9). The generalization in (10), formulated by Rochemont and Culicover (1990:19), holds throughout all of them:

(10) In a well-formed question/answer sequence, all and only the information provided in the response that is not contained in the question is focused.

As we have already demonstrated, the assignment of accent to the extent that it determines the position of a focus, can affect the ability of a sentence to be embedded naturally in a context.

More striking ambiguities arise in SVO structures which can be also interpreted as OVS, even when no accent on a particular item exists (also observed by Rudin 1986). Compare the following sentences:

(11) Marija vidja Ivan
Marija saw-3SG Ivan
“Marija saw Ivan.”/ “Ivan saw Marija.”

In (11), both NPs Marija and Ivan are semantically possible subjects, [+animate], as well as grammatically possible objects. In these cases, the SVO interpretation is favored, as far as the equally possible OVS interpretation is excluded from the appropriate context.

Similarly, in an NP-V sentence presented in isolation, the NP will be conceived of as the subject:

(12) Marija gleda
Marija watch-PRES-3SG
“Marija is watching.”

However, the same NP can be the object, given the suitable context:

(13) Q: Kakvo pravi Ivan?
    What  do-PRES-3SG Ivan
   "What is Ivan doing?”

There are two possible answers to (13), namely (14a) and (14b):
The ensuing ambiguity is another indication that the different interpretations of a sentence in Bulgarian are highly dependent on discourse factors.

1.3. Subject-Verb Asymmetries

Let us go back to the examples in (1). If Ivan has already been introduced in the discourse, for the neutral statements in (1a) and (1b) we will have:

(15) Nameri knigata
    find-PAST-3SG book-the
    “He found the book.”

Bulgarian shows nominal agreement features on the finite verb. This allows SVO/VSO alternations:

(16) a. Včera Ivan kupi knigata AdvSVO
    yesterday Ivan buy-PAST-3SG book-the
    “Yesterday Ivan bought the book.”

b. Včera kupi knigata Ivan AdvVOS
    yesterday buy-PAST-3SG book-the Ivan
    “Yesterday Ivan bought the book.”

c. Včera kupi Ivan knigata AdvVSO
    yesterday buy-PAST-3SG Ivan book-the
    “Yesterday Ivan bought the book.”

Examples like (16c) are in line with the view that Bulgarian is a scrambling language. More instances of subject inversion follow below:
(17) Telefonirax vednaga
call-PAST-1SG immediately
“I called immediately.”

(18) a. Pristigna v osem
arrive-PAST-3SG at eight
“He/She arrived at eight o’clock.”
b. Trūgnaxme kūso
leave-PAST-1PL late
“We left (very) late.”
c. Igraxa s negovite deca
play-PAST-3PL with his-the children
“They played with his children.”
d. Napisax statijata
write-PAST-1SG article-the
“I wrote the paper.”/ “I have written the paper.”

In Bulgarian, virtually every verb can show up with a pre-verbal or a post-
verbal subject. Compare the sentences with three main verb classes:

(19) a. Ivan pristigna
Ivan arrive-PAST-3SG
“Ivan arrived.”/ “Ivan has arrived.”
b. Pristigna Ivan
arrive-PAST-3SG Ivan
“Ivan arrived.”/ “Ivan has arrived.”
c. Majka ti telefonira
Mother your-POSS CL call-PAST-3SG
“Your mother called/has called.”
d. Telefonira majka ti
call-PAST-3SG mother your-POSS CL
“Your mother called/has called.”
(20)  a. Marija igra
      Marija play-PAST-3SG
      “Marija played/has played.”
  b. Igra                  Marija
      play-PAST-3SG Marija
      “Marija played/has played.”
  c. Majka ti                    plaka
      mother your-POSS CL cry-PAST-3SG
      “Your mother cried/has cried.”
  d. Plaka majka ti
      cry-PAST3SG mother your-POSS CL
      “Your mother cried/has cried.”

(21)  a. Ivan napisa                důlgo pismo
      Ivan write-PAST-3SG long   letter
      “Ivan wrote/has written a long letter.”
  b. Napisa               důlgo pismo Ivan
      write-PAST-3SG long   letter Ivan
      “Ivan wrote/has written a long letter.”
  c. Marija kupi                 interesna   kniga
      Marija buy-PAST-3SG interesting book
      “Marija bought/has bought an interesting book.”
  d. Kupi                interesna   kniga Marija
      buy-PAST-3SG interesting book Marija
      “Marija bought/has bought an interesting book.”

Unaccusatives (19), unergatives (20), and transitives (21) allow both SVO and V(OS) orders. All the instances of postverbal subjects, that is, the (b) and (d) sentences of (19), (20), and (21), are grammatical.

Scrambling is also observed in embedded sentences, as demonstrated in (22) - (24):

(22)  a. Mislja, če Ivan pristigna
      think-PRES-1SG that Ivan arrive-PAST-3SG
      “I think that Ivan arrived/has arrived.”
  b. Mislja, če pristigna Ivan
      think-PRES-1SG that arrive-PAST-3SG Ivan
      “I think that Ivan arrived/has arrived.”
The set of data in (19) - (24) also provides evidence that Bulgarian shows features similar to the “classical” scrambling languages (German, Russian), as well as features of Null-Subject Languages (NSL), characteristic for Romance (Spanish, Catalan, Italian) and some Balkan languages (Albanian, Greek).

1.4. Structural Focus and Yes-No Questions

We have seen in Section 1.2. that accented lexical items identify focus in a Bulgarian sentence. The syntactic form (ordering of constituents) can also affect the sentence in a similar way. Yes-No questions in Bulgarian are formed by means of the question particle *li*, which is also a part of the clitic cluster (see 2.1). *Li* occurs in two environments (Van Gelderen and Grozeva 1995, King 1995, Rudin, King and Izvorski 1998): XP-*li*, as demonstrated below in (25), and V-*li*, in (26):
(25) a. Knigata li nameri Ivan?
   book-the Q find-PAST-3SG Ivan
   “Was it the book that Ivan found?”

b. Ivan li nameri knigata?
   Ivan Q find-PAST-3SG book-the
   “Was it Ivan who found the book?”

c. Na nego li dadoxte knigata?
   to-him Q give-PAST-2PL book-the
   “Was it to him that you gave the book?”

(26) a. Nameri li knigata Ivan?
   find-PAST-3SG Q book-the Ivan
   “Did Ivan find the money?”

b. Dadoxte li mu knigata?
   give-PAST-2PL Q him-CL-DAT book-the
   “Did you give him the book?”

c. Šte dojde li s nas?
   will come-FUT-2SG Q with us
   “Are you coming with us?”/“Are you going to come with us?”

In Van Gelderen and Grozeva (1995) we argue that in Bulgarian the
preposed element most closely to the left of li is focused, like all examples
in (25)4 In V-li constructions in (26), there is no focus involved and the
sentences are assumed to be questions regarding the existence of the clausal
event/state of affairs. In other words, li assigns focus in XP-li sentences and
it remains a “pure” question marker in V-li constructions. Following the
above-mentioned work, I claim further that in (25), the focus features of C
are checked by the NP in SpecCP. On the contrary, in (26) the verbs
incorporate in C, check the features and the result is a regular question. Let
us look next at (27):

4 For a detailed and profound analysis of focus and li see Bošković (1997, 2001a and 2001b).
In (27a) and (27b), Ivan and knigata respectively, move to the specifier of a CP, resulting in topicalization, but no focus effect is observed. It is also possible to have the following sentences:

(28) a. Ivan knigata li nameri?
    Ivan book-the Q find-PAST-3SG
    “As for Ivan, was it the book that he found?”

b. Knigata Ivan li nameri?
    book-the Ivan Q find-PAST-3SG
    “As for the book, was it Ivan who found it?”

In (28a) and (28b), the first NPs Ivan and knigata respectively, are topicalized. The second pair, that is, the NPs knigata and Ivan are questioned and focused since they check the question and focus features. Sentences like (29) and (30) lead us (Van Gelderen and Grozeva 1995) to conclude that there are focused and topicalized elements in SpecCP and the structure presupposes a stacking of CPs (multiple CPs):

(29) Ivan knigata nameri li?
    “As for Ivan and as for the book, did he find it?”

5 With respect to topic, I adopt Yokoyama’s (1986) “four sets of knowledge in contact” (Yokoyama 1986:3), which is further specified by King (1995). In Yokoyama’s classification the information relevant for discourse can be divided in the following four sets:

   (i) A  A’s knowledge set
   B  B’s knowledge set
   Ca  A’s current matter of concern
   Cb  B’s current matter of concern

The information in these sets can overlap, thus producing new sets. The material in the set Ca ∩Cb constitutes the shared matter of current concern. This set is assumed to be the topic of an utterance.
(30) Ivan knigata na Marija dade *li*?
    “As for Ivan and as for the book and as for Marija, did he give it to
    her?”

Because of examples (29) and (30), structure (31) must include at least three
CPs (for an analysis concerned with CP recursion see also Marácz 1989,
Rizzi 1991):

(31)

```
CP
  Spec C'
    Spec C'
      C CP
        Spec C'
          Spec +wh
            C
              li
              +focus
              ...
```

SpecCP nodes are also independently motivated for multiple *wh-*
constructions as in (32):
(32) Koj kakvo na kogo dade včera?
    Who, what, to whom gave yesterday
    “Who gave what to whom yesterday?”

It is also possible to focus each of the wh-elements, as shown in (33). Here li plays a role of a pure focus marker. The translation of (33a) includes “who the hell”, or “who, for god’s sake ...”. Such phrases are also needed after each subsequent wh-word in (33b) and (33c):

(33) a. Koj li kakvo na kogo dade včera?
    “Who, for god’s sake, gave what to whom yesterday?”
  b. Koj kakvo li na kogo dade včera?
  c. Koj kakvo na kogo li dade včera?

If we assume that wh-elements are in SpecCPs, focusing li in (33) would be in the respective C head and therefore the multiple CP structure is plausible. If these claims made in Van Gelderen and Grozeva (1995) are correct, it can be concluded that in Bulgarian only overt movement will result in topicalization and/or question. Thus, the Functional Categories when they appear, have strong features.

1.5. The Clitic Cluster

Bulgarian has pronominal clitics and auxiliary clitics. The set of pronominal clitics in Bulgarian includes personal and reflexive pronouns. They are object clitics and have strong counterparts, which can be contrastively stressed and have the same distributional properties as full noun phrases. Apart from these forms, there is also the negative particle ne which results in Negative Stress Shift, and the question particle li. Bulgarian clitics can never occupy an initial position in the sentence. Compare the sentences in (34) - (36):

---

6 Multiple wh-constructions and focus are also discussed in Bošković (1997), Richards (1997) and Rudin (1986).
(34) a. *Mu kazax
    him-CL-DAT tell-PAST-1SG
    “I told him.” / “I have told him.”
b. Az mu kazax
    I him-CL-DAT tell-PAST-1SG
    “I told him.” / "I have told him."
c. Kazax mu
    tell-PAST-1SG him-CL-DAT
    “I told him.” / “I have told him.”

(35) a. *Se gledam v ogledaloto
    REFL-CL-ACC look-PRES-1SG in mirror-the
    “I am looking at myself in the mirror.”
b. Az se gledam v ogledaloto
    I REFL-CL-ACC look-PRES-1SG in mirror-the
    “I am looking at myself in the mirror.”
c. Gledam se v ogledaloto
    look-PRES-1SG REFL-CL-ACC in mirror-the
    “I am looking at myself in the mirror.”

(36) a. *Sme zakūsneli
    be-PRES-1PL late
    “We are late.”
b. Nie sme zakūsneli
    we be-PRES-1PL late
    “We are late.”
c. Zakūsneli sme
    late be-PRES-1PL
    “We are late.”
All types of clitics are syntactically proclitic to the verb and prosodically enclitic, that is, they need a phonological host to their left. The host can be: one or more stressed words (37), the negative particle ne which does not bear a stress (38), or complementizers, like če (that) and dali (if, whether) (39):

(37) a. Nie ja poznavame
we her-it-CL-ACC know-PRES-1PL
“We know her (it).”/ “We have met her.”
b. Nie otdavna ja poznavame
we long ago her/it-CL-ACC know-PRES-1PL
“We have known her (it) for a long time.”
c. Tja se gleda v ogledaloto
she REFL-CL-ACC look-PRES-3SG in mirror-the
“She is looking at herself in the mirror.”

(38) a. Ne ja poznavam
not her/it-CL-ACC know-PRES-1SG
“I don’t know her (it).”/ “I haven’t met her.”
b. Ne mu go dadox
not him-CL-DAT it-CL-ACC give-PAST-1SG
“I didn’t give it to him.”
c. Ne si kupuvam knigi naposledǔk
not REFL-CL-DAT buy-PRES-1SG books lately
“I don’t buy books for myself lately.”

8 Clitics do not need a phonological host to the left when they are preceded by a pause. Compare:

(i) Nakratko ja razpitaxme
shortly-ADV her-CL-ACC ask to tell-PAST-1PL
“We asked her to tell us briefly what happened.”
(ii) Nakratko, ja razpitaxme
shortly-ADV her-CL-ACC ask to tell-PAST-1PL
“We asked her to tell us BRIEFLY what happened.”

In (ii) the clitic ja does not need a phonological support due to the pause after the dislocated adverb nakratko (shortly).
(39) a. Kazax, če ja poznavam
   say-PAST-1SG that her/it-CL-ACC know-PRES-1SG
   “I said that I had met her before./I said that I knew (had known) it
   (e.g., the book).”

   b. Ne znaja dali go poznavat
      not know-PRES-1SG whether him-CL-ACC know-PRES-3PL
      “I don’t know whether they have met him.”

   c. Mislja, če se poznavam dobre
      think-1SG that REFL-CL-ACC know-PRES-1SG well
      “I think that I know myself well.”

   It is interesting to note that in Bulgarian the negative particle ne
carries inherent stress, which is realized on the following word.9 Thus ja in
(38a), mu in (38b) and si in (38c) are stressed although they represent
pronominal clitics. The interaction of the question/focus particle li with the
negative particle ne and the resulting Negative Stress Shift, can be observed
further in (40):

(40) a. Ne li go pokaza?
    not her-CL-DAT Q him/it-CL-ACC show-PAST-2/3SG
    “Didn’t you/he/she show him/it to her?”

   b. Ne im li go izprati?
      not them-CL-ACC Q it send-PAST-2/3SG
      “Didn’t you/he/she send it to them?”

Since li cliticizes to the first stressed element of the verbal complex, it is
placed after the word which immediately follows ne. Halpern (1995) names
this phenomenon prosodic inversion and elaborates it further. More
recently, it is addressed in Bošković and Franks (2000). In (40) the Dative

9 The negative particle ne carries the stress which is realized on the following noun (i),
adjective (ii) or adverb (iii):

   (i) Iskam ne čaj, a kafe
       want-PRES-1SG not tea but coffee
       “I don’t want tea, I would like some coffee.”

   (ii) Izbirame ne bjala roklja, a červenata
        choose-PRES-1PL not white-the dress but red-the
        “We don’t choose the white dress but the red one.”

   (iii) Vidjaj Ivan ne včera, a onja den
        see-PAST-1SG Ivan not yesterday but that day
        “I didn’t see Ivan yesterday, I saw him the day before.”
clitics i and im are stressed because of the preceding negative particle ne and therefore, li follows them. When no other material precedes, the verb itself hosts the clitics, as in (41):

(41) a. Poznavam go
    know-PRES-1SG him/it-CL-ACC
    “I have met him.”/ “I know it.”
b. Viždam gi
    see-PRES-1SG them-CL-ACC
    “I can see them (viz. either the boys, or the books).”
c. Gledam se
    look-PRES-1-SG REFL-CL-ACC
    “I am looking at myself.”

If the finite verb is a clitic form of the auxiliary to be, the clitics are adjacent to the past participle form, as in (42) and (43):

(42) a. Nie sme ti gi izpratili
    We be-1PL you-CL-DAT them-CL-ACC send-PAST PART PL
    “We have sent them to you.”
b. Izpratili sme ti gi
    send-PAST PART PL be-1PL you-CL-DAT them-CL-ACC
    “We have sent them to you.”

(43) a. Ti si mu se obadil
    you be-2SG him-CL-DAT REFL-CL-ACC call-PAST PART MASC SG
    “You (viz. a man, a boy) have called him.”
b. Obadil si mu se
    call-PAST PART MASC SG be-2SG him-CL-DAT REFL-CL-ACC
    “You have called him.”

Within the clitic cluster there is a strict order. Dative clitics always precede accusative clitics. Clitic auxiliaries, on the other hand, precede pronominal clitics, except for the third person singular form e, which follows them. Compare the sentences in (42), where the clitic string is:

(44) Aux.Cl < Dat. Cl < Acc.Cl
(45) presents another example of the ordering in (44), where the clitic doubling phenomenon also obtains:

(45) Knigata sigurno sùm mu ja dala
    book-the surely be-1SG him-CL-DAT it-CL-ACC give-PAST PART-FEM
    “I must have given him the book.”

When the third person singular of be, the form e appears, the order of clitics is:

(46) Dat.Cl < Acc.Cl < e (3sg.Aux)

Other examples illustrating the ordering of elements presented in (46), are given in (47) and (48) below:

(47) Pročel im go e
    read-PAST PART-MASC SG them-DAT-CL it-ACC-CL be-3SG
    “He has read it to them.”

(48) Prodal mu gi e knigi
    sell-PAST PART MASC-SG him-DAT-CL them-ACC-CL be-3SG book
    the
    “He has sold the books to him.”

Apart from verbal agreement, semantic appropriateness and context, Bulgarian has clitic doubling constructions. Pronominal clitics can coexist with a DP or a coreferential full pronoun, which in most of the cases is optional. It occurs either before or after the doubling clitic form.

(49) a. Marija gi vidja tjax
    Marija them-CL-ACC see-PAST-3SG them-FULL PRON-ACC
    “Marija saw them.”

b. Vidja gi tjax Marija
    see-PAST-3sg them-cl-Acc them-full pron-Acc Marija
    “Marija saw them.”
c. %Marija gi $tj$ax vidja
   Marija them-CL-ACC them-FULL PRON-ACC see-PAST-3SG
   “Marija saw them.”

d. $Tj$ax Marija gi vidja
   them-full pron-ACC Marija thtm-cl-Acc see-past-3sg
   “Marija saw them.”

e. $Tj$ax gi vidja Marija
   them-FULL PRON-ACC them-CL-ACC see-PAST-3SG Marija
   “Marija saw them.”
f. %$Tj$ax vidja gi Marija
   them-full pron-ACC see-past3sg them-cl-Acc Marija
   “Marija saw them.”

It has been argued in the traditional linguistic literature (Mirčev 1978, Mladenov 1979, Momčilov 1865) that clitic doubling in Bulgarian contributes to disambiguating case roles.¹⁰ The disambiguating function of pronouns in Bulgarian originates from the fact that they show case, gender and number.

(50) a. Čuvam ja (neja) mucho dobre
   hear-PRES-1SG her-CL-ACC her-FULL PRON-ACC very    well
   “I can hear her very well.”

b. Popita gi (tjax) njakolko pǔti
   ask-PAST-3SG them-CL-ACC them-full PRON-ACC several   times
   “He/She asked them several times.”

Although the presence of the full pronouns neja (her) and tjax (them) is optional in sentences (50a) and (50b), these sentences occur both in written and spoken language.

Nouns in Bulgarian do not have case endings. Full pronouns have nominative, accusative and dative forms and (object) clitics distinguish accusative and dative forms. The clitic doubling phenomenon is conceived as a result from the loss of case endings in nouns (Mirčev 1978, Mladenov 1979). Clitic doubling is viewed there as a way of minimizing ambiguity. Apparently, this is not its sole function in the modern language.

In (51a), the clitic form go shows that the direct object is masculine singular, so Ivan (which is a man’s name), should be the object and Marija the subject. In (51b), the opposite happens: Marija is a woman’s name and it is the object in order to match the clitic ja and Ivan presents the subject of the sentence. But the clitic pronoun cannot help if the two NPs are identical in gender and number:

(52) Georgi go vidja Ivan

Georgi him-CL-ACC see-PAST-3SG Ivan

a. “Ivan saw Georgi.”

b. “Georgi saw Ivan.”

As also pointed out by Rudin (1986), the (a) reading is preferred but informants always accept (b), after it has been brought to their attention.

There are certain cases when clitic doubling in Bulgarian is obligatory, namely with “oblique” subjects. Note the following sentences:

(53) a. Eto gi decata

here/there them-CL-ACC children-the

“Here/There are the kids.”

b. Njama gi decata

lack-PRES-3PL-IMPERS them-CL-ACC children-the

“The kids are not here.”

(54) a. Toplo im e na decata

warm them-CL-DAT-3PL be-3SG to children-the

“The kids are/feel warm.”

b. Lesno mu e na Petûr

easy him-CL-DAT-3SG be-3SG to Petûr

“Things are easy for Petûr.”
In (53) and (54), we see impersonal constructions, while in (55) post-verbal nominatives are observed.

1.6. Bulgarian is Head-Initial

Lexical projections in Bulgarian appear with a head-initial structure in all constructions which they allow:

(56) a. razrušenieto na Rim NP
destruction-the of Rome
“the destruction of Rome”
b. žaden za znanija AP
thirsty for knowledge-PL
“thirsty for (more) knowledge”
c. . . na Rim PP
. . . of Rome
d. . . če Marija predskaza če šte vali VP
that Marija predict-PAST-3SG that will rain
“that Marija has predicted that it will rain”

None of the lexical projections in (56) can surface as head-final. In Bulgarian, prepositions always precede the NP, as illustrated in (56a) - (56c). Most prepositions are phonologically clitics and in this way they are practically stuck in front of the NP. Prepositions following the NP (post-positions) are not allowed in Bulgarian. In the same way complements always follow the head noun, which is shown in (57a).
Determiners and complementizers in Bulgarian occupy different places with respect to their complements. Determiners follow their complements (with only the demonstrative pronoun being an exception), while complementizers precede them:

(58) a. dete-to
    child-the
    “the child” (NEUTER)
b. *to-[NP dete]
    the-child

(59) a. prikazka-ta
    fairy tale-the
    “the fairy tale” (FEMININE)
b. *ta-[NP prikazka]
    the-fairy tale

(60) a. Toj kazva, če [AgrSP vali]
    he say-PRES-3SG that rain-PRES-3SG
    “He says that it rains.”
b. Toj kazva *[AgrSP vali] če
    he say-PRES-3SG rain-PRES-3SG that
    he says rains that

(61) a. Toj pita dali [AgrSP vali]
    he ask-PRES-3SG whether rain-PRES-3SG
    “He asks whether it rains.”
b. Toj pita *[AgrSP vali] dali
    he ask-PRES-3SG rain-PRES-3SG whether
    he asks rains whether

(58) and (59) show that the determiners -to and -ta occur after the NP and they can never precede the NP (the complement). In (60) and (61), the
complementizers če (that) and dali (if, whether) precede their complement [AgrSP vali], they cannot follow it.

Let us regard next the various environments where the determiner (or more generally the definite element) occurs in the DP. It is illustrated in (62) – (65) that the definite element always appears on the head of the AP:

(62) a. interesna-ta statija
    interesting-the article
    "the interesting article"
    b. *interesna statija-ta

(63) a. mnogo interesna-ta statija
    very interesting-the article
    "the very interesting article"
    b. *mnogo interesna statija-ta

(64) a. kupeni-te včera knigi
    bought-the yesterday books
    "the books bought yesterday"
    b. *kupeni včera-te knigi
    c. *kupeni včera knigi-te

(65) a. nova-ta i interesna kniga
    new-the and interesting book
    "the new and interesting book"
    b. *nova i interesna-ta kniga
    c. *nova i interesna kniga-ta

The examples in (66) – (68) exhibit the different configurations in which the long forms and short forms of the possessive pronouns combine with the definite element:

(66) svoja-ta kniga/kniga-ta si/
    *kniga si
    "my/your/his/her/its/our/your/their (own) book"
The long forms of the possessive pronouns behave as adjectives in Bulgarian. On the basis of (66), (67) and (68), the following observations can be made:


---

(67) moja- ta kniga/kniga-ta mi/ *kniga mi
"my book"

(68) svoja kniga/moja kniga
POSS REFL book/POSS PRON book
"one of my/your/his/her/its/our/your/their books/one of my books"

The long form of the possessive pronouns can occur with or without the definite element accompanying it. Consequently, the following sequences are grammatical:

[DP long possessive pronoun + definite article + NP]
[DP long possessive pronoun + zero article + NP]

The possessive clitic is always attached to the article. The possessive clitic can occur only after the definite article. It can never follow an indefinite noun. Hence, two grammatical and one ungrammatical sequences obtain, i.e.

[DP definite article + possessive clitic + NP]
[DP NP +definite article + possessive clitic]
*[ zero article + possessive clitic + NP]

It is evident that the Bugarian DP is strongly dependent on a higher, adjectival/possessive/adverbial phrase. I assume the existence of an Agreement Adjectival/Possessive/Adverbial phrase higher than the DP as shown in (70):
If we take into account the examples in (66) – (68), the structure presented in (70) applies in the following way: svojata kniga and mojata kniga have the features [+poss,+def] which are strong and trigger the movements of svoja and moja from SpecDP via SpecPossP to SpecAgrPossP; the determiner ta is in D and via head movement ends up in the head AgrPossP; the NP kniga moves to SpecAgrDP; for knigata si and knigata mi (the features [+poss,+def] are strong), I assume that the clitic forms are generated in the head position AgrD and come together there with the definite article ta which has moved from D; in svoja kniga and moja kniga the feature matrix consists of [+poss,-def] where [+poss] is strong and svoja and moja, respectively aim at the highest SpecAgrPossP; as for the ungrammatical *kniga si and *kniga mi the harmony of the picture is destroyed by the fact that there is no trigger for the movement to AgrD where the clitics form are generated, i.e., clitics are overruled without the support of the determiner.

On the basis of the facts described in Sections 1.2. - 1.5., I argue that the tree diagram in (71) presents the structure of the Bulgarian clause.
There are minimum two CP positions, the highest one hosting the [+wh] feature and the other (or others) being “reserved” for the focused and/or topicalized phrases in “Yes-No” questions or preposed topicalized or discourse items. Bulgarian, like English, does not have verb movement and object movement in main clauses. As a result, the N-feature of T in both languages is strong and this triggers an overt movement of the subject NP to SpecTP. But the N-feature of Agr is weak in the two languages, otherwise object shift would be allowed. In Bulgarian, as different from English, in addition to the strong N-feature of T, the N-feature of Asp is also strong.
The free word orders (see Section 1.2.) are derived by raising of the verb to C while the topicalized/discourse elements (containing the morphological feature [+topic]) move to SpecCP. I assume the Aspectual Phrase AspP is higher than T. Negation comes even higher than AspP. Let us take for illustration the following example:

(72) Ivan ne trjabva da igrae futbol
    Ivan not must to-play-PRESENT-3SG football
    "Ivan mustn't/needn't play football."

As I have just stated, the N-feature of T is strong which triggers the movement of the subject Ivan to SpecT and the strong N-feature of Asp takes care of its higher raising to SpecAspP. It seems that the subject and the verb in main clauses with the SVO neutral order never raise to AgrSP or AgrS, respectively. The AgrSP phrase with the AgrS head is only an intermediate stop for the lower AspP and TP phrases and their corresponding Asp and T heads. Let us regard the "freedom" of the word order of (72). The remaining possible orderings are: 12

(73) a. Ivan futbol ne trjabva da igrae futbol
    b. Ivan ne trjabva futbol da igrae futbol
    c. Ivan futbol da igrae ne trjabva futbol
    d. Ivan da igrae futbol ne trjabva futbol
    e. #Ivan da igrae ne trjabva futbol
    f. Futbol Ivan ne trjabva da igrae futbol
    g. Futbol ne trjabva da igrae Ivan
    h. Futbol da igrae ne trjabva Ivan
    i. Futbol Ivan da igrae ne trjabva Ivan
    j. Futbol da igrae Ivan ne trjabva Ivan
    k. Futbol ne trjabva Ivan da igrae Ivan
    l. Ne trjabva Ivan da igrae futbol
    m. Ne trjabva da igrae futbol Ivan
    n. Ne trjabva futbol da igrae Ivan
    o. Ne trjabva Ivan futbol da igrae Ivan
    p. Ne trjabva futbol Ivan da igrae Ivan
    q. Ne trjabva da igrae Ivan futbol

12 The translations of the various orderings are not given since they basically repeat that of (72), with practically all constituents being focused or topicalized.
r. Da igrae ne trjabva Ivan futbol
s. Da igrae Ivan futbol ne trjabva
t. Da igrae Ivan ne trjabva futbol
u. Da igrae ne trjabva futbol Ivan
v. Da igrae futbol Ivan ne trjabva
w. Da igrae ne trjabva Ivan futbol

All orderings in (73) are either focused or topicalized material which fills up the SpecCP. In (73f), for example, either futbol or Ivan is in focus while in (73u) da igrae is the topic. As also pointed out in Section 1.2., the intonation effects must also be taken into account. All topicalized or focused, discourse elements trigger movement to the first CP.

1.7. Conclusion

In this chapter, I have given an overview of the syntactic structure of Bulgarian with a view to present it as a background to the study of binding. The structures of the sentence and the DP in Bulgarian are particularly relevant in the process of defining local domains. A special attention is drawn to the clitic forms since reflexive clitics (see chapter 3) are 'true' anaphors in Bulgarian.