

# Morphological and Prosodic Alignment of Bulgarian Clitics\*

Géraldine Legendre  
Johns Hopkins University  
January 1998

*Optimality Theory: Syntax, Phonology, and Acquisition*  
Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (eds.)  
Oxford University Press

## 1. Introduction

Bulgarian clausal clitics include not only familiar object pronouns, but also tense/aspect auxiliaries, modal particles, negative particles, as well as interrogative particles. They display well-known properties associated with clitic elements cross-linguistically. For example, they are typically stressless and they display positional restrictions. In simple sentences, most Bulgarian clitics cluster in second position.

- (1) a. Az *sŭm* mu go dal.  
I be-1 him-dat it-acc given  
'I have given it to him'.  
b. Dal *sŭm* mu go.  
'(I) have given it to him'.

In the absence of a subject pronoun (pervasive in this null-subject language), the non-finite verb must appear in clause-initial position, as shown in (1b). Moreover, Bulgarian clitics cannot be separated from the verb, except by another clitic. Finally, they typically appear in a sequence or cluster whose internal ordering is fixed.

This brief summary underlines the well-behaved character of Bulgarian clitics from a Balkan as well as universal perspective. What makes them of particular interest to a theory of clitics however is the fact that these properties, while typical of Bulgarian, are not obligatory. First, any clitic (other than the question particle *li*) must carry stress in a negative context, if it immediately follows the negative particle *ne* (Hauge, 1976). Throughout this paper, clitics are in italics while stressed clitics are also in upper case.

- (2) a. Ne GO razbiram.  
neg it-acc understand-1  
'(I) don't understand it'  
b. Ne ŠTE *sŭm* pročel knjigata.  
neg fut be-1 read book-the  
'(I) will not have read the book'

Second, some Bulgarian clitics -- *šte* (future auxiliary), *da* (modal/complementizer-like particle) and *ne* -- may appear in clause-initial position in violation of the second-position requirement characteristic of other clitics.

- (3) a. Šte go vi □ daš li?  
fut him-acc see-2 Q  
'Will (you) see him?'  
b. Da ne MI ja dadeš.  
mod neg me-acc it-acc give-2

- ‘Don’t give it to me!’
- c. *Ne ŠTE li go vi□daš?*  
neg fut Q him-acc see-2  
‘Will (you) not see him?’
- d. *Vi□dal li go e?*  
seen Q him-acc be-3  
‘Has (he) seen him?’

Third, additional properties displayed in (3) highlight the uniqueness of the interrogative particle *li*, used to mark yes-no questions: (a) it appears in a variety of positions, including second and further down but never first; (b) it does not necessarily cluster with other clitics; (c) in neutral yes-no questions, *li* cannot immediately follow *ne*; as a result it never carries stress.

Finally, even second-position clitics may fail to occur in the overall second position in a clause. This is the case when a clause contains multiple wh-phrases, topicalized phrases, parentheticals, and the like. Examples are from Tomić (1996).

- (4) a. *Koj kakvo mu e dal?*  
who what him-dat be-3 given  
‘Who gave what to him?’
- b. *Knigata, Penka ja e dala na Petko.*  
book-the Penka her-acc be-3 given to Petko  
‘As for the book, Penka gave it to Petko’

The challenge, therefore, is to provide a unified account of the unmarked as well as the marked properties of all Bulgarian clitics. In this paper I shall defend the view that clitics are not syntactic atoms generated and moved around in the syntax. Rather, clitics instantiate functional features which are realized morphologically as phrasal affixes, as independently proposed in Anderson (1992, this volume). On a par with lexical affixes, phrasal affixes are subject to alignment constraints (McCarthy and Prince, 1993a,b) which favor their realization at the edge of some domain. A competition for realization at the edge of a domain ensues among functional features, which is resolved by ranking the constraints in a language-particular order. Second-position clitics result from the interaction of an alignment constraint whose domain is V’ and one whose domain is the Intonational Phrase.

Additionally, I shall make two other main claims. (a) The question particle *li* is a second-position clitic; the fact that its true nature is often hidden follows from the interaction of morphological and prosodic constraints. (b) Despite the complexity of the interaction, the constraint ranking of Bulgarian is modular. I will propose that modularity of constraint ranking be elevated to the status of a meta-constraint of the theory.

The paper is organized as follows. Section 2 is devoted to presenting evidence that Bulgarian clitic auxiliaries and pronouns are syntactically inactive. This leads to the conclusion that they are not in the syntax but rather in the morphology. Section 3 focuses on demonstrating that the existing OT theory of morphology (Prince and Smolensky, 1993; McCarthy and Prince, 1993a,b) can explain all the Bulgarian facts by exploiting the optimality-theoretic mechanism of ranking applied to simple alignment constraints. The issue of the domain within which clitics are left-aligned is discussed at length. It is concluded, based on Bulgarian, Serbo-Croatian, as well as Tagalog evidence, that the domain of one alignment constraint is prosodic. The other, it is argued, has a phrasal domain. Bulgarian clitics which evade second-position requirements are also shown to fall out of the same formal optimization. Section 4 examines the negative particle *ne* and the question particle *li*, with the conclusion that their conditioning is, to some extent, prosodic. *Li* will be revealed to be a second-position clitic, despite the vagaries of its surface distribution. The interaction of syntactic, prosodic, and morphological constraints in Bulgarian leads to considering the issue of modularity in OT. It is demonstrated that the Bulgarian constraint ranking is in fact modular and leads to the formulation of a meta-constraint, the Constraint Intermixing Ban. Section 5 summarizes the main results of the paper.

## 2. Are Clitics in the Syntax?

One enduring feature of generative grammar is syntactic movement. For example, elements which by virtue of their thematic properties occupy a right periphery position in a clause may instead surface at the left periphery because they have undergone movement to a higher (leftward) position. This is the case for *wh*-phrases in many languages. This has also been claimed for object clitics in Romance following Kayne (1975). While some scholars have since argued against a movement analysis of object clitics (e.g. Borer, 1984; Jaeggli, 1986; Suñer, 1988), one important assumption behind Kayne's original analysis has gone virtually unchallenged among generative syntacticians. It is the view that clitic elements are generated in the syntax and as such obey syntactic constraints.

This is especially clear in the line of research initiated by Pollock (1989). Word order is taken to result from verb movement around fixed categories such as negation and adverbs, and much tree structure in Balkan languages is built upon functional  $X^0$  categories which are clitics, i.e. non-standard syntactic elements (Rivero, 1994a,b; Rudin, 1997, among others). The assumption that clitics are generated in the syntax has gone largely unquestioned despite the fact that its consequences have been theoretically costly. The treatment of Balkan clitics alone has, for example, led to substantial weakening of various versions of the generative theory of syntax, including Government-Binding and the Minimalist Program. Among the adjustments needed, we find the following: lowering of affixes in the syntax (Rivero, 1993), adjunction of heads to maximal projections in violation of the Structure Preserving Principle (Halpern, 1995), right adjunction of clitics (Rivero, 1993; Izvorski et al., 1997) generally disallowed in a restrictive theory of phrase structure, such as Kayne (1994), violations of the Head Movement Constraint (Rivero, 1994a), proliferation of distinct functional projections such as WP (WackernagelP) and TMP (Tobler-MussafiaP) (Rivero, 1994b), as well as movement in the phonology (Halpern, 1995).

The present paper challenges the syntactic status of clitics, arguing instead that they instantiate functional features which are realized morphologically as phrasal affixes. On a par with lexical affixes -- an alternative way in which functional features may be instantiated -- phrasal affixes are subject to alignment constraints which favor their realization at the edge of some domain. The argument is made first for clitic auxiliaries, then for pronominal clitics.

### 2.1. Clitic Auxiliaries

Many properties of clitic auxiliaries are completely unexpected if they head syntactic projections. Most importantly, clitic auxiliaries systematically differ in syntactic behavior from their non-clitic counterparts, both lexical verbs and auxiliaries. As we will see, the latter are syntactically active while the former are syntactically inert.

Bulgarian auxiliary 'be', for example, is well-known for its differing status depending on tense (Hauge, 1976). Present tense 'be' (e.g. first person *sŭm*, third person *e*) is phonologically weak and must procliticize or encliticize to a host; past tense 'be' (e.g. first person *bjax*, third person *beše*) carries stress and has no positional restrictions. In the following discussion past tense forms of 'be' are referred to as lexical auxiliaries. Contrary to lexical verbs and auxiliaries, clitic auxiliaries do not permit subject-auxiliary (SA) inversion in questions. This is shown in (5).

- (5) a. Kakvo pročete Ivan?  
       what read-3 Ivan  
       'What did Ivan read?'  
       b. ?Kakvo beše Ivan pročel?  
           what was-3 Ivan read  
           'What had Ivan read?'  
       c. \*Kakvo *e* Ivan pročel?

- what be-3 Ivan read  
 ‘What has Ivan read?’  
 d. Kakvo *e* pročel Ivan?

In (5a, b) the overt subject follows the lexical verb and auxiliary, respectively *pročete* and *beše*. Other word orders of (5b) are possible and even preferred, having to do with the general preference for sentence-final subjects in neutral contexts. But (5b) is not ungrammatical, it is only stylistically marked. (5c), however, is absolutely rejected by all native speakers, under any discourse conditions. The grammatical counterpart of (5c) is one in which the subject *Ivan* follows the complex clitic auxiliary and past participle: (5d). In other words, the overt subject cannot intervene between a clitic auxiliary and a past participle despite the fact that Bulgarian constituent order is very free (subject only to discourse factors). The ungrammaticality of (5c) does not make sense if *e* heads a functional projection, as is standardly assumed (e.g. Rivero, 1994a; Tomić, 1996). A similar pattern can be observed in yes-no questions:

- (6) a. Pročete *li* Ivan knjigata?  
 read-3 Q Ivan book-the  
 ‘Did Ivan read the book?’  
 b. Beše *li* Ivan pročel knjigata?  
 was-3 Q Ivan read book-the  
 ‘Had Ivan read the book?’  
 c. Pročel *li e* Ivan knjigata?  
 read Q be-3 Ivan book-the  
 ‘Has Ivan read the book?’

In particular, the overt subject *Ivan* intervenes between the lexical auxiliary and the past participle in (6b) while it follows the past participle-clitic auxiliary complex in (6c). As we will see shortly, the ungrammaticality of SA inversion with clitic auxiliaries is not an isolated fact but part of a general pattern in Bulgarian. Moreover, this striking feature is shared by other Balkan languages (Legendre, 1997; 1998). I propose that clitic auxiliaries do not permit SA inversion because they do not have the status of syntactic head which SA inversion requires. An analysis which does not take the clitic status into consideration must rely on generating various auxiliaries under different functional heads, as proposed in Krapova (1997). This leads to empirical and theoretical problems briefly reviewed below. The impossibility of SA inversion also leads to unsatisfactory analyses of second position clitics which rely on Morphological Merger. For example, Embick and Izvorski (1994) propose that Bulgarian clitics are stranded in clause-initial position in the syntax but appear in second position as a result of obligatory affixation to the adjacent lexical head to their right. They stipulate Short Participle Movement in the syntax to account for the adjacency of the clitic auxiliary and past participle in wh-questions (5c) (though they acknowledge, in a footnote, that it has no independent motivation). But why Short Participle Movement does not feed into Morphological Merger in (5c) is unclear. The formulation of the Morphological Merger rule must consequently stipulate the left clausal boundary and thus merely restates the second-position facts.

An important feature of the theory of head movement which underlies that of functional categories is the Head Movement Constraint (HMC, Travis, 1984). Put simply, the HMC prevents a given head from moving across another head dominating it, predicting blocking effects. The existence of Long Head Movement (LHM) in Bulgarian (Rivero, 1994a) across a functional head -- that is, the absence of blocking effects -- is a problem for the HMC. Rivero’s solution is to claim that LHM violates the HMC but complies with the ECP under Relativized Minimality. Under Rivero’s characterization of the Bulgarian facts, (7a) is grammatical because LHM (a type of A-bar movement) crosses an A-head *sŭm* while (7b,c) are ungrammatical because LHM crosses A-bar heads, i.e. modal and negative operator-like elements.

- (7) a. Pročel *sŭm* knjigata.  
 read be-1 book-the  
 ‘(I) have read the book’  
 b. \*Pročel *šte sŭm* knjigata.  
 read fut be-1 book-the  
 ‘(I) will have read the book’  
 c. \*Pročel *ne sŭm* knjigata.  
 read neg be-1 book-the  
 ‘(I) have not read the book’

This account faces the empirical problem that (7b) is not ungrammatical but only awkward. As independently reported in Embick and Izvorski (1994) and Todorova (1995), (7b) is possible in certain stylistically marked environments. Thus, *šte* must be considered an A-head in Rivero’s system despite its semantic affinity to *ne*. The status of A vs. A-bar head becomes a mere stipulation.<sup>1</sup>

In our terms, Bulgarian displays an absence of blocking effects with clitic auxiliaries, not because it violates the HMC, but simply because clitic auxiliaries are not present as syntactic heads. Hence they are not subject to the HMC. LHM with *ne*, despite its clear clitic status, will be argued to result from economy considerations and the fact that *ne* violates the second position requirement (see section 3.2 for details).<sup>2</sup> Note that the non-syntactic status of clitic auxiliaries brings together the absence of subject-clitic auxiliary inversion and the existence of LHM in Bulgarian and other Balkan languages.

There is yet further evidence for the systemacity of the syntactic contrast between clitic and lexical auxiliaries, as documented in Krapova (1997). Lexical, but not clitic, auxiliaries can be separated from their past participle by VP adverbs and floating quantifiers.

- (8) a. Ivana beše naburzo pročela knjigite.  
 Ivana was-3 quickly read books-the  
 ‘Ivana had quickly read the books’  
 b. Studentite bjaxa vsički pročeli knjigite.  
 students-the were-3 all read books-the  
 ‘The students had all read the books’  
 (9) a. \*Ivana *e* naburzo pročela knjigite.  
 Ivana be-3 quickly read books-the  
 ‘Ivana has quickly read the books’  
 b. \*Studentite *sa* vsički pročeli knjigite.  
 students-the be-3-pl all read books-the  
 ‘The students have all read the books’

The preferred grammatical counterparts of (9a,b) have the adverb or floating quantifier preceding the clitic auxiliary. The contrast between (8a) and (9a) leads Krapova (1997) to generate clitic auxiliaries under TP and lexical auxiliaries under a separate lower AuxP projection. The difference in adverb placement is for her the result of raising lexical auxiliaries to Agr in the syntax vs. at LF for clitic auxiliaries. From a theoretical perspective, this means that the weak/strong feature distinction is being parametrized within a language, weakening its explanatory value. From an empirical one, it fails to explain why both types of auxiliaries behave alike with respect to a small class of adverbs like *veče* ‘already’ and *ošte* ‘still’. These adverbs can intervene between any type of auxiliary and the past participle (10a,b) though the preferred word order (or less stylistically marked) for both auxiliaries is (10c):

- (10) a. Ivana *e* veče pročela knjigata.

- Ivana be-3 already read book-the  
 ‘Ivana has already read the book’
- b. Ivana beše veče pročela knjigata.  
 Ivana was-3 already read book-the  
 ‘Ivana had already read the book’
- c. Ivana veče *e*/beše pročela knjigata.  
 ‘Ivana has/had already read the book’

That is, in (10a) the clitic auxiliary *e* appears to have raised across the adverb, despite its weak feature. On the other hand, the adverb appears to be adjoined to TP in (10c) (since *e* heads TP in Krapova’s analysis). It is unclear how the adverb order in (10a) can be accommodated within her approach, short of an additional stipulation or giving up the cornerstone of the Pollockian approach to functional categories, i.e. the assumption that adverbs do not move.<sup>3</sup>

From the present perspective which relies on distinguishing clitic from non-clitic elements, *e* is a clitic hence not present in the syntax. The adverb *veče* is also a clitic with the same fate; semantically related adverbial clitics have been documented in another Balkan language, Romanian (Rivero, 1994a).

The final contrast among auxiliaries pertains to clitic placement.<sup>4</sup> The position of clitic pronouns is fixed in the case of clitic auxiliaries but flexible in the case of lexical auxiliaries. Object pronouns follow clitic *sŭm*; but they preferably precede lexical *bjax*, although Hauge (1976) reports (11c) as possible.

- (11) a. Az *sŭm mu go* dal.  
 ‘I have given it to him’.
- b. Az *mu go* bjax dal.  
 ‘I had given it to him’
- c. Az bjax *mu go* dal.

The question particle *li* shows a related pattern. This time, *li* precedes clitic auxiliaries but follows lexical auxiliaries:

- (12) a. Pročel *li e* knjigata?  
 ‘Has (he) read the book?’
- b. Beše *li* pročel knjigata?  
 ‘Has (he) read the book?’
- c. Pročete *li* Ivan knjigata?  
 ‘Did Ivan read the book?’

As originally discussed in Hauge (1976), *li* must immediately follow the first stressed element in a simple sentence. On the other hand, *e* cannot appear in first position. The past participle *pročel* in first position allows to satisfy both requirements on *e* and *li* simultaneously. How this is formally accomplished will be discussed in sections 3 and 4. *Beše*, however, does carry stress; hence *li* may immediately follow it. The same pattern is found with lexical verbs in simple tenses (12c).<sup>5</sup>

To sum up, there is a systematic contrast between two classes of auxiliaries in Bulgarian, exemplified by unstressed *sŭm* and stressed *bjax*. The latter syntactically behaves like a lexical verb; the former, on the other hand, is syntactically inert: in particular, it does not allow SA inversion nor block LHM. This behavior is straightforwardly explained if auxiliary clitics are not heads of functional phrases. This, in turn, suggests that a morphological approach deserves serious consideration.

## 2.2. Pronominal Clitics

In colloquial Bulgarian, [+specific] NPs -- proper names, emphatic non-clitic object pronouns, and

NPs containing the definite article suffix or the specificity-marking indefinite article -- appear with a clitic pronoun which agrees with the object NP in number, person, gender, and case (Vakareliyska, 1994). While this phenomenon of clitic doubling is obligatory in closely related Macedonian, it appears to be largely optional in Bulgarian. As Vakareliyska remarks, a topicalized non-clitic object pronoun or a topicalized object NP preceding a subject NP is usually doubled by a clitic. Clitic doubling may then serve to disambiguate the direct object from the subject in the latter case (Bulgarian does not morphologically differentiate nominative from accusative NPs).

- (13) a. Nego/Ivan *go* vi□dam.  
 him/Ivan him-acc see-1  
 ‘(I) see him/Ivan’ (= ‘as for him/Ivan, (I) see him’)
- b. Georgi *go* gleda Marija. (Rudin 1986:17)  
 Georgi him-acc watch-3 Marija  
 ‘Marija is watching Georgi’
- c. Georgi *ja* gleda Marija.  
 Georgi her-acc watch-3 Marija  
 ‘Georgi is watching Marija’

Note that doubling generally places emphasis on the object NP; thus in cases where the object and subject NPs are identical in gender and number, clitic doubling favors OVS interpretation (Rudin, 1986).

We need not concern ourselves with an exhaustive description of clitic doubling here. It is important, however, to be aware that clitic doubling is alive and well in Bulgarian. The other property of clitic doubling of more immediate relevance to the clitic status issue is case: the object NP appears without a preposition; hence it has case (under standard assumptions) and it violates Kayne’s generalization.

Following Suñer (1988), I take clitic doubling to be the mirror image of subject-verb agreement, i.e. a subcase of object-verb agreement. Moreover, keeping in line with OT’s economy of structure expressed here and elsewhere (Legendre et al. 1995, 1988; Grimshaw, 1997a; Legendre, 1997, 1998) I wish to adopt the more traditional view of agreement as a relation between an event and its arguments. Concretely, this means two things. One, the input to optimization includes theta roles which get expressed via case (a constraint CASE requires every nominal element to have case). Two, features of arguments (including person, number, case, etc) get realized twice in the morphology, i.e. on arguments and V/VP as an expression of this relation. In other words, object agreement (including clitic doubling) is the default. Where object agreement/clitic doubling does not occur, a constraint -- call it EXPRESS(F) -- is violated in order to satisfy a constraint like DROP GIVEN (which may turn out to be identical to the constraint DROPTOPIC proposed in Samek-Lodovici (1996)). It is well-known that pronominal elements like *him*, *them*, etc. are used for old, given information. Thus, if an input contains the lexical item *John* and the output contains only a double, say a clitic version of *him*, EXPRESS(F) is violated but higher-ranked DROP GIVEN is satisfied.

This brief excursion into the technical implementation of clitic doubling serves to highlight the fact that the case properties satisfied by object clitics are not uniquely interpretable as evidence that clitic pronouns are syntactic objects. That conclusion depends on general assumptions concerning case and agreement which, historically, have focused on languages without clitic doubling. In fact, all linguists who have examined languages with clitic doubling have concluded that object clitics could not be handled by movement in the syntax (Borer, 1984; Jaeggli, 1986; Suñer, 1988). The syntactic approach then must rely on alternative treatments for languages with clitic doubling and those without. This does not seem to me to be a theoretically satisfying conclusion.

Returning to Bulgarian clitic doubling, the fact that clitics are indeed agreement markers supports the conclusion reached previously on the basis of clitic auxiliaries: clitics belong to the morphology rather than to the syntax. Moreover, if pronominal clitics were active in the syntax, we would expect blocking

effects such as strong crossover effects. No such effects, however, are observed in Macedonian, which has obligatory clitic doubling (Legendre, 1998).

- (14) Čovekot<sub>i</sub> kogo<sub>i</sub> što go<sub>i</sub> vidov t<sub>i</sub>.  
 man whom that him saw-1  
 ‘The man whom I saw’

In Bulgarian, *deto* relative clauses are preferred with a clitic double, according to my main informant.<sup>6</sup>

- (15) Čovekut<sub>i</sub> deto go<sub>i</sub> vidjax t<sub>i</sub>.  
 man-the that him saw-1  
 ‘The man that (I) saw’

Under the assumption that clitics occupy specifier positions (Rivero, 1994b), movement of the antecedent ought to be blocked, for example by Relativized Minimality. But it’s not. Alternatively, if clitics are heads of functional projections (Tomić, 1996; Rudin, 1997), they ought to block movement by virtue of the HMC. But they don’t, as numerous examples have shown. One might counter-argue that these constraints do not apply to adjoined elements and hence take pronominal clitics to be adjoined to heads (or to XPs). But if adjunction of clitics is just a strategy to evade syntactic constraints, we have, in fact, an argument against treating clitics as syntactic objects.

To sum up, there is strong evidence that Bulgarian clitic auxiliaries and pronominals are syntactically inert, a conclusion which is inconsistent with the claim that clitics are syntactic atoms. Rather, it supports the view, most prominently put forward in Anderson (1992), that clitics are elements of morphology. In the next section, I argue that a particular theory of morphology -- the alignment-based theory of Prince and Smolensky (1993) and McCarthy and Prince (1993a,b)-- is exactly the theory we need to explain the complex distribution of clitics in Bulgarian.<sup>7</sup>

### 3. Alignment-based Morphology

#### 3.1. Clitics as Instances of Phrasal Morphology

If we take seriously the fact that the clustering of unrelated categories (auxiliaries, pronouns, etc.) in a single position universally distinguishes clitics from standard syntactic elements and consider the parallels that exist between clitics and word-level affixes, then an alternative view immediately comes to mind. Cross-linguistically, clitics display, within a phrase, the same range of placement options as word-level affixes, including initial, final, second-position, and penultimate position (Klavans, 1985; Anderson, 1992). Even infixation has a counterpart in clitics which break up the constituency of a phrase. This is possible in Croatian dialects of Serbo-Croatian (Anderson, 1996, this volume). This is also possible for the Bulgarian question particle *li* (see section 4) . Clitics and affixes are both positioned relative to an anchor point (proclitics vs. prefixes, enclitics vs. suffixes). Both are located in the scope of some constituent which constitutes its domain. Finally, the internal order within a clitic sequence is invariant, as is the order of affixes attached to a word.

As Anderson (this volume) also notes, the rigid ordering of clitics does not generally reflect the relative scope of functional categories. For example, first person *sim* precedes clitics pronouns but third person *e* follows them in Bulgarian and other South Slavic languages. No syntactic explanation has ever been proposed for this idiosyncrasy because there is none: irregularities of this sort are a hallmark of morphology, not syntax. As a result, in all syntactic analyses of South Slavic clitics I am aware of, the behavior of *e* requires a distinct mechanism, typically some sort of PF movement.

Finally, syntax is not equipped to deal with the rigidity of clitic order in languages which are

characterized by free order constituent order (subject to discourse constraints) and fairly free head movement (given the existence of LHM). Though typically ignored in the literature, the fundamental word order distinction between clitics and non-clitic elements requires an additional stipulation under the syntactic approach. This is however naturally captured under a morphological approach because morphology, as opposed to syntax, is characterized by fixed order of morphemes. This is, I suggest, directly linked to the type of constraints that operate in both components. Morphological constraints are, to a large extent, alignment constraints (McCarthy and Prince, 1993a,b) while syntactic constraints are typically constraints that force or restrict movement (Legendre et al., 1995, 1998; Grimshaw, 1997a).

To give just one example, the question particle *li* exhibits a property which does not make much sense if *li* is in C, as is commonly assumed (Rivero, 1993; King 1995; Tomić, 1996; Rudin et al., 1997). A yes-no question may contain the particle *li*, or alternatively the interrogative complementizer *dali* ‘whether’, which is argued to be a true complementizer in Rudin (1986) -- hence in C under standard assumptions.<sup>8</sup>

- (16) a. Dali Ivan *e* otišul?  
wether Ivan be-3 left  
‘Has Ivan left?’  
b. Otišul *li e* Ivan?  
left Q be-3 Ivan  
‘Has Ivan left?’

Unlike the complementizer *dali*, which can be used only to question a clause, *li* can question single constituents in isolation (Rudin, 1986:65).

- (17) a. Az *li*?            Kūštata *li*?        Dnes *li*?            Na masata *li*?  
‘me?’                ‘the house?’        ‘today?’            ‘on the table?’  
b. \*Dali az?            \*Dali kūštata?        \*Dali dnes?        \*Dali na masata?

It’s hard to conceive of a complete clausal structure with *li* in C when all that is present in (17a) is a single NP or PP constituent, followed by *li*. Consequently, this distribution leads to recognizing two (accidentally similar) question particles, *li* in C for clausal questions and *li* modifying a phrase (possibly adjoined to the right of the head or the phrase in syntactic analyses). If *li*, however, is a phrasal affix, then it can attach to any phrasal constituent, whether it be one headed by V or some other head.

Thus, I proceed with the assumption that clitics are morphological categories which instantiate properties of phrases, on a par with (word-level) affixes which instantiate properties of words. The functional properties in question are features of the node dominating the relevant phrase.

### 3.2. Second Position

Prince and Smolensky (1993:35ff) define the notion prefix and suffix as morphemes respectively subject to EDGEMOST(LEFT) and EDGEMOST (RIGHT) constraints which align a morpheme edge with a word edge. They demonstrate that Tagalog infixes such as *um* ‘Actor trigger’ result from the interaction of EDGEMOST(LEFT) with a general constraint against closed syllables (NOCODA), where NOCODA outranks EDGEMOST(*um*, LEFT). They further demonstrate how penultimate stress systems also make use of gradient EDGEMOST. That is, second position from a right edge follows from the interaction of EDGEMOST with a non-gradient NONFINALITY constraint: if NONFINALITY outranks EDGEMOST, a candidate with penultimate stress will violate EDGEMOST once to avoid violating NONFINALITY. A candidate with antepenultimate stress will violate EDGEMOST twice and hence be sub-optimal. EDGEMOST and NONFINALITY are construed as universal constraints. In addition, EDGEMOST is subject to directional parameters (Left, Right), domain parametrization (word, stem, etc.) and parametrization to individual morphemes in the first case. NONFINALITY is also subject to domain parametrization. Thus, EDGEMOST and NONFINALITY represent both universal alignment schemas

and families of constraints potentially individualized for particular morphemes.

I propose, along with Anderson (1996, this volume), that phrasal affixes are analyzable in terms of Prince and Smolensky's original alignment constraints. For Bulgarian (and, more generally, Balkan languages), the relevance of the left edge of the clause suggests retaining Prince and Smolensky's gradient EDGEMOST(LEFT) as well as making use of a domain-initial counterpart of their NONFINALITY constraint, which I shall refer to as NONINITIAL. In the interest of clarity of presentation, the discussion of the domain of these two constraints is postponed until the basic technical interaction of EDGEMOST and NONINITIAL has been introduced. For the moment, we may assume that the domain is that of the smallest constituent corresponding to the simple null-subject sentences we have exemplified so far: V'. Further refinements will follow.

It is important to keep in mind that no matter how theoretically desirable it is to make use of general alignment constraints in the placement of clitics (Jacobs, 1996), one cannot align a sequence of clitics without the use of individualized alignment constraints. This is not a weakening of the theory of alignment, this is just its natural implementation.

That is, clitics not only seek placement at the edge of a domain but they also cluster, a fundamental property typically ignored in syntactic and prosodic discussions of clitics -- notwithstanding some important exceptions, including Perlmutter (1971). Clitics cluster simply because they compete for the same EDGEMOST position. The competition is resolved by ranking individual EDGEMOST constraints (for each feature F) relative to one another. Hence it is not surprising that the ranking of EDGEMOST constraints, to a large extent, reflects the clitic cluster.<sup>9</sup> It is not correct, however, to assume that EDGEMOST constraints may only mirror the surface distribution of clitics. An important demonstration to the contrary is made in Legendre (1998), on the basis of the competition among EDGEMOST constraints in Macedonian.

A strict ranking of EDGEMOST constraints makes an important prediction: the relative order of clitics in the cluster remains the same regardless of where the cluster appears, relative to its anchor point. There is considerable cross-linguistic evidence that this prediction is correct: the change from pro- to encliticization, or vice-versa, does not involve rearranging the clitics involved.

Where clitics cluster -- e.g. clause-initial vs. second position -- results from the relative ranking of EDGEMOST and NONINITIAL. If NONINITIAL outranks EDGEMOST, second-position clitics follow. This is the main Bulgarian pattern, repeated in (18).

- (18) Pročel *sŭm* knigata.  
'(I) have read the book'

The competition is made most explicit, by using standard OT tableaux.<sup>10</sup>

T1. NONIN >> E(PERF<sub>LC</sub>)

	NONIN	E(PERF <sub>LC</sub> )
a. [ <sub>V</sub> <i>sŭm</i> pročel knigata] [perf]	*!	
☞ b. [ pročel <i>sŭm</i> knigata]		⊗
c. [ pročel knigata <i>sŭm</i> ]		**!

The input to optimization is assumed here to include the lexical items *pročel* and *knigata* and their basic propositional structure, plus functional features like [tense], [perfect], [negation], etc. In Macedonian, [tense] competes with other functional features for second position with the result that the placement of clitics is sensitive to the finiteness of the clause (Legendre, 1998). In Bulgarian, however, [tense] does not crucially

interact with other functional features, hence I am leaving it out of the discussion for the sake of simplicity. (See Legendre (1998) for an explicit comparison between the two languages). The first candidate in each tableau is annotated with the relevant features to facilitate interpretation of the forms. The constraint EDGEMOST(PERF) and its subscript ‘lc’ will be further discussed below.

In T1, candidate (a) loses because it violates higher-ranked NONINITIAL while candidate (c) loses because it incurs two violations of EDGEMOST, vs. one for the winner, candidate (b). That is, it is more important in Bulgarian to satisfy NONINITIAL than EDGEMOST(PERF). As discussed further below, EDGEMOST violations are determined by the number of morphemes which separate a target clitic from the left edge of the V’ domain.

Note some important properties of this analysis. There is no syntactic movement or LHM of the past participle *pročel*. There is thus no HMC violation, contrary to Rivero (1994a). There is no Prosodic Inversion (Halpern, 1995) or Morphological Merger (Embick and Izvorski, 1994) either -- these are PF movements designed to directly achieve second position.<sup>11</sup> No movement exists because none is needed. By economy, the past participle simply remains in situ. The alignment constraints responsible for the placement of clitics belong to the family of constraints independently needed for affixation.

Consider the first prediction made by optimality-theoretic constraints, that is re-ranking. If, contrary to the Bulgarian ranking, EDGEMOST outranks NONINITIAL, clause-initial clitics obtain. This is, in fact, the basic pattern in some other Balkan languages, in particular Romanian and Macedonian (Legendre, 1997, 1998).

- (19) R a. *L-am văzut.*  
 him have seen  
 ‘(I) have seen him’  
 M b. *Ti go dade.*  
 you it gave  
 ‘(He) gave it to you’

Tableau T2 displays the competition corresponding to the Romanian example. As T2 shows, it is more important in Romanian to satisfy EDGEMOST than NONINITIAL.

T2. Romanian: E(ACC) >> E(PERF<sub>lc</sub>) >> NONIN

	E(ACC)	E(PERF <sub>lc</sub> )	NONIN
☞ a. [ <sub>v</sub> <i>l-am văzut</i> ] [acc] [perf]		⊗	⊗
b. [ <i>văzut l-am</i> ]	*!	**	
c. [ <i>am îl văzut</i> ]	*!		*

An important prediction immediately follows from the analysis of a position as resulting from the interaction of two alignment constraints, EDGEMOST and NONINITIAL. A ranking is possible whereby some EDGEMOST constraints outrank NONINITIAL which, in turn, outranks other EDGEMOST constraints. The result is a language where second-position is required of some clitics but not others. This is in fact true of Bulgarian. While the perfect auxiliary and object clitics are strictly second position, the future auxiliary *ște*, the modal particle *da*, and the negative particle *ne* may freely occur in clause-initial position.

- (20) a. *Ște sîm pročel knigata.*  
 fut be-1 read book-the

- ‘(I) will have read the book’
- b. *Da ne MI ja dadeš.*  
mod neg me-acc it-acc give-2  
‘Don’t give it to me!’
- c. *Ne ŠTE sŭm pročel knjigata.*  
neg fut be-1 read book-the  
‘(I) will not have read the book’

As tableau T3 shows, EDGEMOST(FUT) must outrank NONINITIAL, which in turn must outrank EDGEMOST(PERF) (see T1). Thus, a subset of Bulgarian clitics behave like Romanian clitics: their preference for being at the clausal edge outweighs their desire to not be in initial position.

T3. E(FUT) >> NONIN >> E(PERF<sub>LC</sub>)

	E(FUT)	NONIN	E(PERF <sub>LC</sub> )
a. [ <sub>V</sub> <i>šte sŭm</i> pročel knjigata] [fut] [perf]		⊗	⊗
b. [pročel <i>šte sŭm</i> knjigata]	*!		**
c. [ <i>sŭm šte</i> pročel knjigata]	*!	*	
d. [ <i>šte</i> pročel <i>sŭm</i> knjigata]		*	**!

The reader might be skeptical of the clitic status of *šte*, *ne* and *da* in view of their conspicuous violation of the basic second-position requirement on Bulgarian clitics. In fact, some scholars have considered them not to be clitics, for that very reason (Dimitrova-Vulchanova, 1993; Rivero, 1994a). This cannot be correct, however, because *šte* and *da* observe all other basic properties of Bulgarian clitics: (a) they cluster with other clitics, (b) their order is invariant, and (c) they do not carry stress. Note that the latter property alone is evidence against Prosodic Inversion: as (20) shows, a clitic may serve as the host of an enclitic, despite the fact that it itself does not constitute a prosodic word.<sup>12</sup>

Moreover, *šte*, *ne*, and *da* cannot be separated from the verb by a VP adverb.

- (21) a. Ivan burzo *šte* pročete knjigata. (Krapova, 1997)  
‘Ivan will quickly read the book.’
- b. \*Ivan *šte* burzo pročete knjigata.
- c. Ivan burzo pročete knjigata.  
‘Ivan quickly read the book.’
- d. Ivana beše naburzo pročela knigite.  
‘Ivana had quickly read the books.’

The contrast in (21) remains unexplained under Rivero’s claim that *šte* heads a modal phrase. In light of (c) and (d), which respectively involve a tensed verb and a non-clitic auxiliary, one would have to stipulate that manner adverbs must precede *šte* or invoke adverb movement. The adverb pattern extends to *ne* and the particle *da*.<sup>13</sup>

- (22) a. Ivana *ne* burzo beše pročela knigite.  
‘Ivana had not quickly read the books.’
- b. \*Ivan *ne* burzo *šte* pročete knjigata.

‘Ivan will not quickly read the book.’

- (23) a. \*Obeštavam *da ne go* otnovo ostavjam sam.  
 promise-1 to not him-acc again leave-1 alone  
 ‘I promise to not leave him alone again’  
 b. Obeštavam *da ne go* ostavjam otnovo sam.

It shouldn’t come as a surprise that Balkan irrealis particles (which I take to subsume *šte* and *da*) and negation often violate second-position requirements. This is presumably linked to the scopal properties of these modal-like elements, though the issue will not be further examined here.

Consider how the present approach naturally handles some other properties of Bulgarian. As noted in Rivero (1994a), LHM is impossible in the presence of *ne*.

- (24) a. *Ne sŭm* pročel knjigata.  
 ‘(I) haven’t read the book.’  
 b. \*Pročel *ne sŭm* knjigata.

There is a very simple explanation for this and it does not require any additional mechanism, such as assuming that NegP is a barrier. The presence of *ne* in clause-initial position sanctioned by the grammar or constraint ranking ensures that auxiliary (and pronominal) clitics automatically satisfy NONINITIAL. There is absolutely no need to front the past participle. In fact, doing so would result in a crucial violation of EDGEMOST(NEG), now further away from the left edge. Additional violations would include \*STRUCTURE (economy of structure) and \*TRACE (economy of movement) though these are omitted in tableau T4. As shown in T4, candidate (b) can never be an optimal candidate in Bulgarian.

T4. E(NEG) >> NONIN

	E(NEG)	NONIN	E(PERF <sub>LC</sub> )
☞ a. [ <sub>V</sub> <i>ne sŭm</i> pročel knjigata] [neg] [perf]		⊗	⊗
b. [pročel <i>ne sŭm</i> knjigata]	*!		**

The same account extends to imperatives. In prohibitives, *ne* provides the morpheme needed to satisfy EDGEMOST(ACC). See Legendre (1997) for a cross-linguistic analysis of imperatives and prohibitives.

- (25) a. Četi *ja!*  
 read-imp it-acc  
 ‘Read it!’  
 b. *Ne JA* četi!  
 neg it-acc read-imp  
 ‘Don’t read it!’

Negative gerunds, however, differ from prohibitives. Here, the presence of *ne* does not cause a repositioning of object clitics.

- (26) a. Davaiki *mu go*...  
 giving him-dat it-acc  
 ‘Giving it to him...’

- b. *Ne davaiki mu go...*  
 ‘Not giving it to him...’

In Legendre (1997, 1998), I argue that Balkan gerunds carry a [gerund] feature, as indicated by the unique morphology of these forms. The pattern in (26) results from the fact that EDGEMOST(GER) outranks EDGEMOST(DAT).

Turning to wh-questions, note that the wh-phrase precedes *ne* in wh-questions, as shown in (27).

- (27) *Kakvo ne pročete Ivan?*  
 what neg read-3 Ivan  
 ‘What didn’t Ivan read?’

Borrowing two constraints from Grimshaw (1997a), we can easily make sense of (27). One constraint is OPERATOR IN SPECIFIER (OPSPEC) requiring syntactic operators to move to a higher specifier position. The other is OBLIGATORY HEADS (OBHD) which requires the head of each syntactic projection to be filled. (27) shows that *ne* violates its left edge requirement to allow the wh-phrase to be in a scopal position. Hence OPSPEC outranks EDGEMOST(NEG). The fact that Bulgarian questions take post-verbal subjects, on the other hand, shows that the verb *pročete* has moved to C, to satisfy high-ranked OBHD. In fact, everything falls into place in this OT analysis of (27) if, as argued above, *ne* is not present in the syntax. If *ne* were to head, say, a NegP necessarily positioned above the projection headed by the verb, then a fatal violation of OBHD would ensue. This theory-internal argument constitutes independent evidence that *ne* is best analyzed as a clitic, subject to a high ranked EDGEMOST constraint. This analysis of (27) also reveals two important aspects of the overall analysis. One, the position of clitics cannot be identified with a unique syntactic position. Clitics appear wherever V goes, e.g. in VP or in CP. Two, syntactic constraints systematically outrank morphological alignment constraints. We will return to the latter point in section 4.

In frameworks relying on inviolable constraints, the clause-initial position of clitics *šte* and *da* in null subject sentences is truly idiosyncratic and hard to derive without additional stipulation. In OT, however, this is the kind of behavior that one expects, and it derives from the same mechanism of constraint interaction which results in regular behavior: constraint ranking. There is nothing idiosyncratic about *šte* or *da*. They cluster with other clitics because, like all Bulgarian clitics, they are subject to EDGEMOST constraints strictly ranked with respect to one another. Their overall position in a clause, however, results from the relative ranking of EDGEMOST and NONINITIAL. Bulgarian is an interesting language because NONINITIAL breaks the block of EDGEMOST constraints into two blocks. We will also return to this point in section 4.

As discussed earlier, the distribution of the present perfect auxiliary in Bulgarian and other South Slavic languages depends on its person and number; third person singular ‘be’ follows object clitics while all other forms of clitic ‘be’ precede dative and accusative clitics. Under an analysis of ‘be’ as heading an AuxP projection, it is completely unexpected to have this kind of exception. A special PF mechanism must be appealed to. From a morphological perspective however, this pattern is not idiosyncratic; rather, it reflects the relative universal flexibility of affix ordering. The single mechanism of constraint ranking readily accounts for this kind of flexibility.

Note that ‘be’ forms, in fact, instantiate several features, including [perfect], [person], and [number]. For example, *sŭm* is the first person singular form of the perfect auxiliary ‘be’. Thus, EDGEMOST(PERF), though a convenient shortcut, is in fact an incomplete characterization of the relevant constraints. The forms are portmanteaux, that is to say, the constraints apply to the same local domain, a morpheme. As argued in Smolensky (1993, 1995), constraints applying to a common local domain can be conjoined. ‘Local conjunctions’ (Smolensky’s term) are, by definition, violated only when the conjuncts are both violated within a common local domain. This is, I propose, what underlies the distribution of ‘be’ forms in Bulgarian. Specifically, EDGEMOST(PERF) is decomposed into three EDGEMOST constraints, with equally ranked EDGEMOST(PERF)&EDGEMOST(1,2PERS), EDGEMOST(PERF)&EDGEMOST(PLU) outranking EDGEMOST(PERF). Note the relationship between the higher ranked, more specific constraints and the lower default one. This

reflects Pānini's Theorem on Constraint Ranking (Prince and Smolensky, 1993).

Consider how the local conjunction applies to first person singular *sŭm*.

#### T5. Local Conjunction of EDGEMOST Constraints

	NONIN	E(PERF)&E(1,2)	E(PERF)&E(PL)	E(ACC)	E(PERF)
a. [ <sub>V</sub> <i>sŭm go</i> pročel] [perf <sub>1sg</sub> ] [acc]	*!			*	
b. [ pročel <i>sŭm go</i> ]		⊗		⊗⊗	⊗
c. [ pročel <i>go sŭm</i> ]		**!		*	**
d. [pročel <i>go e</i> ] [perf <sub>3sg</sub> ] [acc]				⊗	⊗⊗
e [pročel <i>e go</i> ]				**!	*
f. [ <i>go e</i> pročel ]	*!				*

In, say, second position in V' (candidate (b)), *sŭm* violates the conjunction EDGEMOST(PERF) & (1,2PERS) ranked above EDGEMOST(ACC) as well as the lower ranked default EDGEMOST(PERF). A first person plural form would violate EDGEMOST(PERF) & EDGEMOST(PLU) in addition (the dotted line represents equal ranking). On the other hand, as candidate (d) shows, third person singular *e* violates only the default E(PERF) constraint -- the other two being irrelevant. EDGEMOST(PERF) is, of course, ranked below EDGEMOST(ACC). The local conjunction version of EDGEMOST(PERF) is identified throughout the paper by the subscript 'lc', for simplicity's sake.<sup>14</sup>

To summarize this section, all positional effects result from one mechanism, constraint ranking. The alignment constraints that are ranked (including local conjunctions) are constraints at work elsewhere in morphology and morphophonology. The constraints refer to functional features rather than individual morphemes. That is, they are not specific to Bulgarian, nor even to Balkan or Slavic languages. Clitics universally cluster because they compete for the same position, at the edge of some domain.

### 3.3. The Domain of Alignment Constraints

At this point, the question of the precise characterization of the domain of the two alignment constraints needs to be resolved. There are several issues to be addressed. One pertains to EDGEMOST: what is the domain at the left edge of which a given feature is realized?

Bulgarian clitics have long been characterized as syntactic proclitics, meaning they attach to the main verb. This presumably reflects the fact that clitics express properties of the head of the proposition, V. On the other hand, *šte*, *ne*, and *da* satisfy EDGEMOST despite the fact they do not carry stress. This suggests that the left edge requirement is not prosodic in nature. Rather it pertains to a phrase structure constituent.

The precise characterization of this syntactic constituent largely depends on one's assumptions about clausal structure. A central concept in much current work in linguistics is that of economy. In OT, this means economy of structure -- governed by \*STRUCTURE -- and economy of movement -- governed by \*TRACE (Legendre et al., 1995, 1998; or STAY Grimshaw, 1997a). In concrete terms, this means that under the VP-internal subject hypothesis, a simple clause need not involve more than a VP, if morphological properties are handled without head movement. In fact, Legendre (1998) argues that EDGEMOST applies to [tense] by documenting crucial competitions between tense and other features in Macedonian. Of course, some clauses do involve verb movement to a higher head position, but only when movement is motivated for other than morphological reasons, e.g. wh-movement.

Another relevant assumption concerns null subjects. If they are instances of *pro*, then a minimal SVO clause containing a null subject is a VP. If, on the other end, the existence of null subjects is denied (Grimshaw and Samek-Lodovici, 1995, 1998), then the same clause is a V'. I shall here adopt Grimshaw and Samek-Lodovici's view that null subjects result from the interaction of constraints pertaining to the syntactic realization of subjects and their discourse status and propose that the domain of the EDGEMOST constraint is V' in Bulgarian and other South Slavic languages.

The crucial evidence comes from simple clauses containing an overt subject. As (28) shows, domain-initial clitics like *šte* follow the subject.

- (28)           Az *šte sŭm* pročel knigata.  
                  'I will have read the book'

The domain cannot be the maximal projection containing the subject. Hence it cannot be the prosodic constituent corresponding to root clauses either, which presumably is the Intonational Phrase. Otherwise, *šte* would precede neutral subjects. On the other end, the fact that clitics may appear within the C' projection -- under V to C movement in questions -- raises the issue of whether the domain should be generalized to X'. But if CPs are only extended projections of V, as proposed in Grimshaw (1991), then V' is still the relevant domain of clausal clitics.

There is more to EDGEMOST than simply determining its domain. Because being at the edge is a matter of degree, EDGEMOST is a gradient constraint. Hence, degree of violation of EDGEMOST is measured in terms of units which, a priori, could be morphological or prosodic. As it turns out, both are relevant to Bulgarian. Two subclasses of clitics can be identified with respect to the units they are sensitive to. Clitic pronouns and auxiliaries count morphemes, as is clear from the fact that they may encliticize to another (stressless) clitic, *šte* or *da*, in initial position: see, for example, (20). In the next section, the question particle *li*, the modal particle *da*, and negative particle *ne* will be shown, however, to count prosodic units generally referred to as prosodic words.

To sum up, the domain of EDGEMOST is V' for all clitics, the gradiency of EDGEMOST is measured in morphological units for clitic auxiliaries and pronouns but prosodic units for *li*, *da*, and *ne*.

We turn now to the question of the domain of NONINITIAL. At first glance, the generalization appears to be in terms of the notion clause. This is because in phrase structure terms, a second-position clitic like the clitic auxiliary cannot be initial in V' (in null subject sentences), in VP (in sentences with overt subjects), or in CP (in wh-questions). Relevant examples are repeated here for convenience.

- (29)   a.       Dal *sŭm mu go*.  
                  '(I) have given it to him'.  
       b.       Az *sŭm mu go* dal.  
                  'I have given it to him'.  
       c.       Kakvo *e* pročel Ivan ?  
                  'What has Ivan read?'

Now, consider object topicalization in Bulgarian:

- (30)           Knigata, Penka *ja e* dala na Petko.       (Tomić, 1996)  
                  book-the Penka it-acc be-3 given to Petko  
                  'As for the book, Penka gave it to Petko'

Note that the fronted topicalized object in (30), *knigata*, is separated by an intonational break from the rest of the sentence. Note also that no word order change affects the subject, the verb, and the clitics. For the purpose of counting second-position, it is as if the topicalized NP didn't count. Yet, the whole sentence forms

a clause.<sup>15</sup> But, in terms of the whole clause, the clitics *ja* and *e* do not cluster in second position.

The same pattern can be observed in wh-structures with object topicalization, and, more generally speaking, in sentences containing sentence-internal pauses.

- (31) a. Knigata, koj *mu* *e* dal? (Tomić, 1996)  
 book-the who him-dat be-3 given  
 As for the book, who gave it to him?
- b. Ivan včera *se* obadi, vūrna *mu* *gi*, i *si* otide. (Hauge, 1976)  
 Ivan yesterday refl-acc called-3 returned-3 him-dat them-acc and refl-dat went-3  
 ‘Ivan called yesterday, returned them to him, and went home’

Note in particular that clitics occur in second position after each pause marked by a comma in (31b). The presence of intonational breaks marked by commas makes it very clear that the relevant domain is the Intonational Phrase rather than a phrase structure constituent.

Note, however, that the first reflexive clitic, accusative *se*, is in third position in (31b). This is not a counterexample, though. Rather, this follows from the analysis if the present alignment constraints are outranked by constraints which pertain to the placement of adverbs. Given the view of minimal structure outlined above, there is no evidence for a higher functional projection in the absence of verb movement in (31b). Hence the topic subject is presumably in SpecVP and the adverb *včera* is presumably adjoined to V'. A thorough study of adverb positions is clearly needed within the present framework to independently support this assumption. But assuming this can be done, *se*, in fact, satisfies both EDGEMOST and NONINITIAL under the present proposal: the former by its position at the left edge of V', the latter by virtue of not being the first element in the Intonational Phrase.

The fact that syntactic constraints outrank morphological ones also explains a similar effect in multiple wh-questions.

- (32) Koj kakvo *mu* *e* dal?  
 who what him-dat be-3 given  
 ‘Who gave what to him?’

In Bulgarian, all wh-phrases must move to SpecCP (Rudin, 1988). Note that (32) satisfies all relevant constraints if OPSPEC outranks the prosodic and morphological constraints. OPSPEC is satisfied by both wh-phrases. Clitics are at the left edge of V' and non-initial within the Intonational Phrase. Placing the dative clitic *mu* after the first wh-phrase, on the other hand, would in fact violate EDGEMOST(DAT), because the clitic would not be within V'.

Finally, in embedded clauses with a null subject, Bulgarian clitics immediately follow the complementizer or wh-phrase. The data is from Rudin (1986).

- (33) a. Mislja [<sub>CP</sub> če *e* isljazŭl].  
 think-1 that be-3 gone out  
 ‘(I) think that (he) went out’
- b. Ne znaja [<sub>CP</sub> dali *e* isljazŭl].  
 neg know-1 wether be-3 gone out  
 ‘(I) don’t know wether (he) went out’
- c. Tova a knigata [<sub>CP</sub> deto *ja* kupix].  
 this is book-the that it bought-1  
 ‘This is the book that (I) bought’
- d. Ne znam [<sub>CP</sub> kogo *si* vidjal].

neg know-1 who be-2 seen  
 ‘(I) don’t know who (you) saw’

The prosodic structure literature is typically mute about the prosodic structure of syntactic clauses unless they are separated by commas, in which case they are invariably identified as Intonational Phrases. According to Ladd (1996), boundaries of Intonational Phrases need not be defined by a clear pause; other manifestations include a subtle local slowing or pitch change, or a local fundamental frequency fall or rise. Until better understanding of complex prosodic structure is gained, it seems reasonable to assume that syntactic clauses map onto a prosodic constituent bigger than the Phonological Phrase, which might be the Intonational Phrase.<sup>16</sup> My main informant reports that a clear pause is optional before the complementizer/wh-phrase in (33) though one is in fact recommended by the rules of proper enunciation.

With this caveat in mind, I conclude, from the evidence presented in (30)-(33), that the domain within which a clitic must be non-initial is the Intonational Phrase. Clitic placement in simple sentences can obscure this generalization because simple sentences constitute Intonational Phrases.

Evidence that the domain of NONINITIAL is the Intonational Phrase is not limited to Bulgarian. For example, substantial evidence in Serbo-Croatian is discussed in Radanović-Kocić (1996).

- (34) SC a. Ja, tvoja mama, obećala *sam ti* igračku.  
 I your mom promised be-1 you-dat toy  
 ‘I, your mom, promised you a toy’  
 b. Ja *sam ti* obećala igračku.  
 ‘I promised you a toy’  
 c. \*Ja *sam ti*, tvoja mama, obećala igračku.  
 d. \*Ja, tvoja mama, *sam ti* obećala igračku.

Clitics are normally placed after the subject in Serbo-Croatian (34b). But when an appositive is inserted, the placement of the clitic must be delayed: (34a) vs. (34c) vs. (34d). This shows that, just like in Bulgarian, the domain within which a clitic cannot be initial is the Intonational Phrase. Radanović-Kocić (1996) provides substantial additional evidence, including the following contrast between restrictive and unrestrictive relatives.

- (35) SC a. Ona moja sestra koja *je* u Sarajevu *vas se* sjeća.  
 that my sister who is in Sarajevo you refl remembers.  
 ‘My sister who is in Sarajevo remembers you’  
 b. Moja sestra, koja *je* u Sarajevu, sjeća *vas se*.  
 ‘My sister, who is in Sarajevo, remembers you’  
 c. \*Moja sestra, koja *je* u Sarajevu, *vas se* sjeća.

In (35a) the clitics *vas se* cluster in second position after the complex subject *ona moja sestra koja je u Sarajevu* consisting of a restrictive relative clause. In (35b), however, the unrestrictive relative results in intonational breaks, the second of which marks the edge of the domain within which *vas se* cluster in second position. This is confirmed by the ungrammaticality of (35c). Under the assumption that embedded clauses constitute separate Intonational Phrases, the clitic *je* in the relative clause in (35a) obeys the same constraint.

Anderson (this volume) mentions the interesting case of Tagalog clitics, most of which bear their own stress but nonetheless appear in second position.<sup>17</sup> As Anderson comments, there does not seem to be any phonological reason why these could not be initial. Yet the conclusion that Anderson draws from this -- that the domain is syntactic rather than prosodic -- is incorrect. Kroeger (1993) provides evidence that the domain of NONINITIAL in Tagalog is also the Intonational Phrase.

In Tagalog, Adjunct Fronting is distinct from Topicalization: the former does not result in an

intonation break after the fronted element while the latter does. As the Intonational Phrase domain hypothesis would predict, clitics must immediately follow the fronted constituent in Adjunct Fronting but clitics follow the verb in Topicalization structures. This is exemplified by the third person singular subject clitic *siya* in (36).

- (36) T a.      Bukas   *siya* aalis.  
           tomorrow he   leave-fut  
           ‘It’s tomorrow that he is leaving’  
       b.      Bukas,   aalis *siya*.  
           tomorrow leave-fut he  
           ‘Tomorrow, he’s leaving’

To sum up, the formal characterization of alignment constraints is a complex issue because their domain may, a priori, be prosodic or syntactic; and so may be the units that underlie the gradience of EDGEMOST. I have argued that Bulgarian, in fact, exploits all the options available. The domain of EDGEMOST is defined in phrase structure terms. Recall the evidence: domain-initial clitics like *šte* follow rather than precede neutral subjects in root sentences. Hence, they are not positioned at the left edge of the Intonational Phrase which coincides with the sentence. Rather, they are realized at the left edge of V’. For most clitics, violations of EDGEMOST are counted on the basis of the number of morphemes which separate the clitic from the left edge of V’. The domain of NONINITIAL, however, is the Intonational Phrase, hence prosodic. Crucial evidence comes from complex sentences with sentence-internal pauses, in Bulgarian and other languages. Simple sentences, in fact, provide poor evidence for the domain of NONINITIAL because the Intonational Phrase coincides with the clause.

This double characterization -- morphological and prosodic -- formally captures the traditional claim that Bulgarian clitics are syntactically proclitics but phonologically enclitics (Tomić, 1996). Second-position clitics immediately precede V because they are at the left edge of V’ and Bulgarian is a VO language. They cannot be in initial position of an Intonational Phrase. Hence they must be preceded either by a single element within V’ or any number of elements outside of V’ as long as they belong to the same Intonational Phrase. The present analysis is not merely an implementation of the traditional claim. The morphological theory of clitics these constraints are part of and the optimality-theoretic character of these constraints have been argued to provide together a more exhaustive and satisfying explanation for the marked and unmarked properties of Bulgarian clitics than previous studies have offered.

The unique optimality-theoretic character of these alignment constraints raises some further issues relevant to modularity. As the reader may be aware, the candidate representations which undergo optimization are global structures in the sense that these representations combine all traditional levels of representations (syntactic, morphological, and prosodic) into one single level of representation. Furthermore, constraints may refer to any sub-level of representation: some refer to syntax (OBHD, OPSPEC), others to morphology (EDGEMOST), yet others to prosody (NONINITIAL). These constraints interact, hence the question: is the interaction modular? That is to say, do rankings of constraints freely intermix or do rankings consist of blocks of constraints referring to the same sub-level of representation? In section 4, I will demonstrate that rankings are indeed modular. But first, we turn to the prosodic conditioning of *li* and *ne*.

#### 4. Further Prosodic Alignment

The fact that Bulgarian *ne* and *li* obey prosodic constraints has been known since (at least) Hauge (1976). His detailed description of clausal clitics includes the following statements: ‘*ne* always moves its stress over to the following word, also when this word is a clitic’ (p. 18); ‘*li* is placed immediately to the right

of the first stressed element within the verb constituent' (p. 20). While these generalizations are conceived as highly specific rules in Hauge's standard transformational approach, they are, I argue below, best viewed as resulting from the interaction of several constraints. These include already established NONINITIAL and EDGEMOST, as well as prosodic constraints independently needed to handle Bulgarian prosody in general. I proceed by examining the negative particle and then turn to the question particle.

#### 4.1. The Negative Particle *ne*

In general, the clitic *ne* directly affects only the prosodic structure of its environment. It's only indirectly that it affects the position of *li*. Consider (37).

- (37) a.  $\check{S}te$   $go$   $\{_{PrWd} VI \square das\}$ .  
 fut him-acc see-2  
 '(You) will see him'
- b.  $\{_{PrWd} Ne \check{S}TE\}$   $go$   $\{_{PrWd} VI \square das\}$ .  
 neg fut him-go see-2  
 '(You) will not see him'

Curly brackets  $\{ \}$  will be used to represent Prosodic Words (PrWd) which, I assume, typically do not include clitic material.<sup>18</sup> Clitics, unlike lexical heads, are in general stressless; this follows if clitics do not head a PrWd. I propose that the general parsing of clitics outside of a PrWd follows from a constraint which requires them to be parsed in the higher unit of prosodic structure called Prosodic Phrase (PrPh): PARSE(F, PRPH), where [F] stands for any feature (realized as a clitic). In addition, a general alignment constraint on PrWd will be needed: ALIGN(LEXHEAD, L; PRWD, L); it requires the lexical head to be left aligned with the left edge of a PrWd. This constraint is called ALIGN rather than EDGEMOST simply to keep prosodic alignment from morphological alignment apart in our discussion. The reader should keep in mind, however, that they are both instantiations of the same general alignment schema, the only difference here being their domain (prosodic for ALIGN; V' for EDGEMOST).

Obviously, Bulgarian clitics do get parsed into a PrWd in a negative context, as (37b) above shows. Stress on the future clitic  $\check{S}TE$ , I propose, results from the interaction of two constraints. One is PARSE(F, PRPH) which both *ne* and *šte* violate. Given Hauge's generalization, the other constraint pertains to *ne*. It can be stated as another instantiation of the prosodic alignment constraint, ALIGN(NEG, L; PRWD<sup>o</sup>, L): align the left edge of *ne* with the left edge of the head of a PrWd.<sup>19 20</sup> This constraint is further discussed below.

Putting these proposals together for (37a,b), we can turn to the competition itself, represented in tableau T6. Note that it is a double tableau, with positive context at the top and negative context at the bottom. This format allows to easily compare the two separate optimizations.

T6. Prosodic structure of *ne*

	A(NEG)	A(LEXHD)	P(F, PRPH)
☞ a. [ <sub>V</sub> <i>šte go</i> {VI□das}] [fut] [acc]			
b. [{ <i>šte go</i> VI□das}]		**!	**
c. [{ <i>ŠTE go</i> } {VI□das}]			**!
☞ d. [{ <i>ne ŠTE</i> } <i>go</i> {VI□das}] [neg] [fut] [acc]			⊕⊕
e. [{ <i>ne ŠTE go</i> } {VI□das}]			***!
f. [{ <i>ne ŠTE</i> } { <i>go</i> VI□das}]		*!	***
g. [{ <i>ne šte go</i> VI□das}]	**!	***	***
h. [{ <i>NE šte</i> } <i>go</i> {VI□das}]	*!		**
i. [{ <i>ŠTE ne</i> } <i>go</i> {VI□das}]	*!		**

To avoid unnecessary crowding, T6 contains only candidates which are optimal with respect to morphological alignment. For example, an alternative to candidate (a) with the order of the two clitics reversed would be sub-optimal because it violates the higher-ranked of the relevant EDGEMOST constraints, EDGEMOST(FUT). The three alternative candidates (a), (b), (c) in T6 represent alternative prosodic parsings of the input. Candidate (a) is optimal because it is perfect: it violates no prosodic requirement. Other candidates violate one or both of ALIGN(LEXHD, L;PRWD, L) and PARSE(F, PRPH).

The bottom competition illustrates how *ne* affects the outcome. The order of clitics results from the ranking of EDGEMOST constraints (not represented), with EDGEMOST(NEG) outranking EDGEMOST(FUT). Note that both prosodic alignment constraints, ALIGN(NEG, L;PRWD<sup>o</sup>, L) and ALIGN(LEXHD, L;PRWD, L), can be simultaneously satisfied if the sequence is parsed into two prosodic words (candidates (d) and (e)) rather than one (candidate (g)). The decision then falls to PARSE(F, PRPH) which favors parsing the clitic *go* outside of either PrWd. Tableau T6 shows that each of the proposed prosodic constraints is in fact fatal to at least one candidate. For example, ALIGN(NEG, L; PRWD<sup>o</sup>, L) eliminates candidate (i) in which the clitic order within the first PrWd is reversed (candidate (i) also involves an additional EDGEMOST(NEG) violation but the latter is not fatal if prosodic constraints outrank morphological ones, as argued below). Candidate (h) also violates ALIGN(NEG, L; PRWD<sup>o</sup>, L) though in a different way. *Ne* is at the left edge of the PrWd but it is stressed. This is where reference to PrWd<sup>o</sup> rather than PrWd captures the violation incurred by (h). If ALIGN(NEG) referred to PrWd, then it would be satisfied by (h) and (h) would be as optimal as (a), contrary to fact. Because the stresslessness of clausal *ne* is an absolute requirement (Hauge, 1976) we could, alternatively, invoke a \*NE constraint ('don't stress *ne*') and keep the domain of ALIGN(NEG) as PrWd. From a theoretical perspective, it is preferable, however, to derive stress properties of various clitics from the interaction of general constraints like ALIGN and PARSE rather than specific constraints like \*NE.

4.2. The Question Particle *li*

Recall Hauge's generalization according to which *li* must immediately follow the first stressed element. Reference to the 'first element' immediately rings the second-position bell. In other words, *li*, like most Bulgarian clitics, is a second-position clitic. But here is the twist: while other second-position clitics satisfy EDGEMOST on the basis of the number of morphemes that separate them from the left edge of V', *li*

satisfies EDGEMOST on the basis on the number of PrWds that separate it from the left edge of V'. An explicit formalization of the two EDGEMOST constraints is given in (38).

- (38) a. EDGEMOST(ACC) = \*<sub>[V' ( )<sub>m</sub> (acc)]</sub> where ( ) represent morpheme boundaries  
 b. EDGEMOST(Q) = \*<sub>[V' { }<sub>PrWd</sub> (Q)]</sub>

Once again, it is important to realize that these constraints simply instantiate the general alignment schema by specifying its two main parameters (the other is directional and it is directly incorporated in (38)). One is the domain at the edge of which the feature favors realization. It is V' for all features, as argued earlier. The other relates to the units in terms of which violations of EDGEMOST are to be measured.

EDGEMOST(Q), in fact, needs further refinement. Rudin et al. (1995) report that *li* may optionally break up a syntactic constituent when, for example, a modifying adjective is the focus of the question.

- (39) a. [<sub>NP</sub> NOVATA kola] *li* prodade (ili starata)?  
           new-foc car Q sold-2 or old-the  
           ‘Did you sell your NEW car (or the old one)?’  
 b. [<sub>NP</sub> NOVATA *li* kola] prodade (ili starata)?

This phrasal analog to infixation is reminiscent of the one found in mostly Croatian dialects of Serbo-Croatian.<sup>21</sup> Assuming that this patterns violates INTEGRITY(XP) (Anderson, this volume), then the EDGEMOST(Q) constraint in (38) is in need of a slight modification to reflect a wider option of domains: X<sup>n</sup> instead of V' (where X<sup>n</sup> = V' or XP).<sup>22</sup>

As it turns out, this refinement is needed independently of the pattern in (40). As we have already seen, *li* is used to question any constituent -- see (17) above and (42) below.

The competition underlying the surface distribution of *li* is displayed in the double tableau T7. NONINITIAL is omitted because it does not affect the outcome: it is violated by all candidates.

T7. Prosodic structure of *li*

	A(Q)	A(NEG)	A(LEXHD)	P(F, PRPH)	E(Q)	E(FUT)	E(ACC)
☞ a. [ <sub>v</sub> <i>šte go</i> {VI□das} <i>li</i> ] [fut] [acc] [Q]					⊗		⊗
b. [ <i>šte</i> {VI□das} <i>li go</i> ]					*		***!
c. [ <i>šte li go</i> {VI□das}]	***!						**
d. [ <i>šte go li</i> {VI□das}]	**!						*
e. [{ <i>šte go</i> VI□das} <i>li</i> ]			**!	**	*		*
f. [ <i>go šte</i> {VI□das} <i>li</i> ]					*	*!	
☞ g. [{ <i>ne ŠTE</i> } <i>li go</i> {VI□das}] [neg] [fut] [Q] [acc]				⊗⊗	⊗	⊗	⊗⊗⊗
h. [{ <i>ne ŠTE</i> } <i>go</i> {VI□das} <i>li</i> ]				**	**!	*	**
i. [{ <i>ne ŠTE</i> } { <i>li go</i> VI□das}]			**!	****	*	*	***
j. [{ <i>ne šte li go</i> VI□das}]	***!	***	****	****		*	***
k. [ <i>ne šte</i> {VI□das} <i>li go</i> ]		*!			*	*	****
l. [{ <i>ne GO</i> } <i>li šte</i> {VI□das}]				**	*	***!	*

The competitions of particular interest in T7 are between candidates (a) vs. (b) (top rows) and candidates (g) vs. (h) (bottom rows). They provide evidence for the proper formulation of EDGEMOST. If the units relevant to violations of EDGEMOST(Q) were the same as those of EDGEMOST(ACC), then candidate (b) would win the top competition -- *li* is closer to the left edge of the domain in terms of morphemes (two morphemes away) than candidate (a) (three morphemes away). Note that the pattern cannot be handled by simply re-ranking the two constraints. If EDGEMOST(ACC) outranked EDGEMOST(Q), then candidate (g) would lose to (h): two violations of EDGEMOST(ACC) for (h) vs. three for (g). It is beyond doubt that EDGEMOST(ACC) violations are measured in terms of morphemes -- recall that an unstressed clitic such as *šte* or *da* can provide the support they need to their left. Once the prosodic structure is made explicit, it is also very clear from the minimally different candidates (g) and (h) that PrWds, rather than morphemes, are relevant to *li*.

Note that the prosodic alignment on *li* differs from that of *ne* in one respect: the directionality of the alignment. The left edge of *ne* must be aligned with the left edge of a PrWd<sup>o</sup>, hence it is parsed inside the PrWd. The left edge of *li* must be aligned with the right edge of a PrWd, hence it is parsed outside the PrWd. This follows from the standard assumption that a PrWd minimally consists of a binary foot (see further discussion below tableau T8). Thus, there remains an irreducible difference between *ne* and *li*, naturally captured in terms of two alignment constraints, ALIGN(Q, L; PRWD, R) and ALIGN(NEG, L; PRWD<sup>o</sup>, L).

Like EDGEMOST, ALIGN is gradient. In T7, I am assuming that its violations are measured in terms of syllables. (See, for example, sub-optimal candidates (c), (d), (j) and (k)). Alternatively, the units could be construed as morphemes. I leave this matter open for now.

Is the placement of *li* sensitive to the particular clitics involved? That is, does the presence of domain-

initial *šte* in T7 affect the competition in any crucial way? The answer is negative, as shown in (40) and corresponding tableau T8.

- (40) a. Izpratix *li mu* kniga?  
sent-1 Q him-dat book  
'Did (I) send him a book?  
b. *Ne MU li* izpratix kniga?  
not him-dat Q sent-1 book  
'Didn't (I) send him a book?

T8. P(F, PRPH) >> E(Q) >> E(DAT)

	A(Q)	A(NEG)	A(LEXHD)	P(F, PRPH)	E(Q)	E(DAT)
☞ a. [ <sub>V</sub> {izPRAtix} <i>li mu</i> ]					⊗	⊗⊗
b. [ <i>li mu</i> {izPRAtix}]	**** !					*
c. [{izPRAtix <i>mu</i> } <i>li</i> ]				*!	*	*
d. [{izPRAtix} <i>mu li</i> ]	*!				*	*
☞ e. [{ <i>ne MU</i> } <i>li</i> {izPRAtix}]				⊗⊗	⊗	⊗
f. [{ <i>ne MU</i> } {izPRAtix} <i>li</i> ]				**	**!	*
g. [{ <i>ne LI</i> } <i>mu</i> {izPRAtix}]	*!			**		**
h. [{ <i>ne izPRAtix</i> } <i>li mu</i> ]			*!	*	*	***
i. [ <i>ne</i> {izPRAtix} <i>li mu</i> ]		*!			*	***

The closest competitions are between candidates (a) and (c) (top rows) and candidates (e) and (f) (bottom rows). PARSE(F, PRPH), which favors, parsing clitics outside of the PrWd is fatal to (c). On the other end, EDGEMOST(Q) is fatal to (f). Note that in both competitions, the fatal constraint is relatively low-ranked in T8.

Candidate (i) deserves further comments. It is sub-optimal because *ne* is outside the PrWd headed by the verb. If it were not for the fact that PrWds minimally consist of a binary foot (McCarthy and Prince, 1996), then *ne* could be placed outside the PrWd and right-aligned with the left edge of the PrWd°. This alternative would yield the right result in T7, as the reader may verify. But the alternative does not work in T8: under right-alignment, ALIGN(NEG) is satisfied and candidate (i) becomes optimal, contrary to fact. Thus, *ne* must be left-aligned within the PrWd. This, in turn, constitutes empirical evidence for McCarthy and Prince's claim.

Finally, the placement of *ne* and *li* in the absence of any other clitics falls out of the present analysis.

- (41) {<sub>PrWd</sub> *Ne ZNAEŠ*} *li*?  
neg know-2 Q  
'Don't (you) know?'

Note that (41), as represented, only violates ALIGN(LEXHD,L; PRWD, L). It confirms the ranking ALIGN(NEG, L; PRWD°, L) >> ALIGN(LEXHD,L; PRWD, L). Alternative prosodic parsings violate additional constraints,

including \*STRUCTURE if a second (unnecessary) PrWd is added.

In the analysis of *li* proposed above, prosody plays an important role but prosody alone cannot account for *li*'s distribution. The interaction between EDGEMOST constraints also plays a decisive role (as we just saw in T8). But movement plays no role at all and no constraint of syntax is violated. This contrasts with an analysis like Rivero (1993) who argues for a type of movement generally banned by the theory she assumes. In her analysis, *li* lowers and left-adjoins to the finite verb in the presence of *šte* or *ne*. It is also proposed as a language-specific rule of Bulgarian. Note that (41) forces her to assume that *li*, upon lowering, may also right-adjoin to the verb. It is unclear how the directionality of adjunction can be predicted, short of an additional stipulation.

Returning to tableaux T7 and T8, note that they provide evidence for the relationship among blocks of constraints. Hence, they can shed light on the modularity of constraint ranking. In section 3, it was argued that syntactic constraints (e.g. OPSPEC and OBHD) outrank morphological alignment constraints (see examples (27) and (32)). The fact that optimal candidates in tableaux T7 and T8 may violate morphological EDGEMOST constraints but not purely prosodic constraints shows that prosodic constraints must outrank morphological constraints. But what about EDGEMOST(Q), which is a mixed constraint? As a clausal clitic, its domain is V' but its violations count PrWds rather than morphemes. T7, in fact, provides crucial evidence that EDGEMOST(Q) outranks all morphological EDGEMOST constraints: under the alternative ranking the competition would yield the wrong optimal candidate.

One other constraint to consider is EDGEMOST(NEG) whose status has been left unspecified so far. As it turns out, either formulation as a purely morphological or a mixed constraint will work. Consider why. Because *ne* is not a second position clitic, it does not violate EDGEMOST(NEG), hence the units relevant to counting violations of EDGEMOST are irrelevant. There is at least one theoretical reason, however, to prefer the mixed formulation. Besides *li*, *ne* is the only Bulgarian clitic subject to strict prosodic alignment constraints. A generalization would fail to be captured if the EDGEMOST constraint on *ne* didn't mirror that on *li*.

Which brings us to the remaining constraint, NONINITIAL. Recall that its domain was argued to be the Intonational Phrase. If our hypothesis about the modularity of the constraint ranking is correct, then the prosodic constraint NONINITIAL must outrank EDGEMOST(Q). Direct evidence for the ranking comes from clauses containing intonation breaks.

- (42) Knigata, Petko *li mu e dal?* (Tomić, 1996)  
 book-the-TOP Petko-FOC Q him-dat be-3 given  
 'As for the book, is it Petko who gave it to him?

*Li* occurs in second position following the focused subject *Petko*; the topicalized object is irrelevant. In (42), the question is no longer a neutral yes-no question. Rather, it takes on a cleft-like interpretation: the question is not about the event but about *Petko* (Rudin et al., 1995). In our terms, this means that the NP *Petko* is marked with the features [foc] and [Q] in the input. Economy of structure and absence of evidence for verb movement suggest that *Petko* is in situ (SpecVP). If the hypothesis that clitic placement follows from alignment is correct, then the constraints pertaining to *li* are in fact blind to which element carries Q. The reader may verify that sentence (43) is grammatical because it corresponds to the optimal candidate in T9.

## T9. Prosodic constraints &gt;&gt; E(Q)

	A(Q)	A(LEX HD)	P(F, PRPH)	NONIN	E(Q)	E(DAT)	E(PERF)
☞ a. [ <sub>V</sub> {PETko} <i>li mu e</i> {DAL}] [foc] [Q]					⊗	⊗⊗	⊗⊗⊗
b. {PETko <i>mu</i> } <i>li e</i> {DAL}			*!		*	*	***
c. {PETko} <i>mu e</i> {DAL} <i>li</i>					**!	*	**
d. <i>li</i> {PETko} <i>mu e</i> {DAL}	**!			*		**	***

Consider the fact that in T9, all candidates simply correspond to an Intonational Phrase subconstituent of (42). If we now consider another optimization that takes the whole sentence (42) into consideration, there is a counterpart to (d), say (d') preceded by the topicalized object *knigata*. In that context, ALIGN(Q, L; PRWD, R) is satisfied -- the topicalized object provides the necessary host. Candidate (d') still loses to its optimal competitor (a'), but only because it violates NONINITIAL in its Intonational Phrase (as does (d) in T9). This establishes the partial ranking NONINITIAL >> EDGEMOST(Q) and the claim that prosodic constraints outrank mixed ones. This ranking also yields a surprise result already hinted at at the beginning of the paper. Recall that this ranking is synonymous with second position. Hence, *li* is indeed a second position clitic.

The optimality-theoretic analysis thus achieves two goals. One is to explain the vagaries of *li*'s placement, the other is to reveal the deep regularity of the Bulgarian system. The former has been shown to result from the interaction of prosodic and morphological constraints. The latter results from the basic ranking NONINITIAL >> EDGEMOST which underlies most Bulgarian clitics, including *li*.

Finally, recall that EDGEMOST(NEG) outranks NONINITIAL -- since it is not a second-position clitic. Hence the global ranking in Bulgarian can be summarized as in (43). The focus is on the interface between the prosodic and morphological constraints, hence other constraints are globally referred to as Syntactic or Prosodic:

- (43) Syntactic constraints >>  
 Prosodic constraints >>  
 Morpho-prosodic constraints (EDGEMOST(NEG) >> NONINITIAL >> EDGEMOST(Q))  
 >>  
 Morphological constraints (EDGEMOST(FUT) >> EDGEMOST(DAT), etc.)

What (43) shows is that even in a language which displays a complex interaction of an impressive number of constraints the ranking itself is modular. That is, there is no evidence whatsoever in Bulgarian that constraints may freely intermix. Thus, I interpret the solidity of the Bulgarian case to be indicative of a meta-constraint at work. This constraint on constraint rankings is stated in (44).

- (44) CONSTRAINT INTERMIXING BAN: Constraints belonging to different modules of the grammar may not intermix.

The architecture of standard OT does permit reference to any level of representation. That is, a given alignment constraint may relate structures from one module to another directly, as was demonstrated for the EDGEMOST family. This simply means that some of the alignment constraints argued for in this paper are interface conditions. As the growing literature on the syntax-phonology interface attests, the need for interface conditions is independent of OT. But it finds a natural implementation in standard OT because, by their very

nature, candidate structures are global structures. As we have seen, optimization proceeds in parallel rather than serially.

The alternative approach is one in which the phonology repairs what the syntax produces. This is in fact the prevailing one, under many guises which have been examined throughout this paper. The repair approach has two major drawbacks: one is empirical, the other is theoretical. As we have seen, clitics do not, in general, share the syntactic properties of their non-clitic counterparts. Reported claims to the contrary have overlooked a great deal of evidence, some of which has been brought to light here. Moreover, from the perspective of the grammar, there is in principle no reason why phonology should undo what the syntax does -- and do it, using the same mechanism, i.e. movement. This duplication, alone, raises doubts that the traditional view on the relationship between syntax and phonology is correct.

## 5. Conclusion

To summarize, I have argued in this paper that the complex distribution of Bulgarian clitics is not the result of syntactic movement. Nor does it result, even partially, from a post-syntactic re-ordering of elements at PF. Instead, I have put forward substantial arguments in favor of analyzing clitics in morphological terms. That is to say, clitics are phrasal affixes whose positional effects mirror those of word-level affixes (see also Anderson, this volume). I have argued in detail that the position of clitics is the product of a complex interaction of alignment constraints. I have shown that an optimality-theoretic constraint ranking captures their unique properties. This includes their invariant sequences as well as their different anchors, the latter mostly resulting from the fact that constraints are assumed to be violable.

A large inventory of Bulgarian structures have been examined. This has led, among other things, to some novel conclusions. For example, I have argued that second-position clitics result from the interaction of EDGEMOST with NONINITIAL, despite the fact that the alignment domain of the former is defined in terms of phrase structure while that of the latter is defined in terms of prosodic structure. In root sentences, the two domains coincide, but in complex sentences with internal intonation breaks they do not. The interaction of these two constraints explain why Bulgarian clitic auxiliaries and pronouns are in second position in simple sentences but not necessarily in complex ones.

I have also claimed that the Bulgarian question particle *li* is in fact a second-position clitic though the interplay of constraints thoroughly masks its true nature. Finally, I have argued that, despite the complexity of the interaction, the constraint ranking of Bulgarian remains modular. I have proposed that modularity of constraint ranking be elevated to the the status of a meta-constraint of the theory.

## REFERENCES

- Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge University Press.
- Anderson, S. R. (1996). How to put your Clitics in their Place. *The Linguistic Review* 13, 165-191.
- Anderson, S. R. (this volume). Towards an Optimal Account of Second Position Phenomena.
- Beckman, M. E. and J. Pierrehumbert. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook* 3, 255-309.
- Borer, H. (1984). *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.
- Dimitrova-Vulchanova, M. Clitics in Slavic. (1993). *Working Papers in Linguistics* 18, University of Trondheim.
- Embick, D. and R. Izvorski. (1994). On Long Head Movement in Bulgarian. In J. Fuller, H. Han, and D. Parkinson, eds. *ESCOL '94*. DMLL Publications, Cornell University, 104-115.
- Grimshaw, J. (1991). Extended Projections. Ms, Brandeis University.
- Grimshaw, J. (1997a). Projection, Heads, and Optimality. *Linguistic Inquiry*, 28:3, 373-422.
- Grimshaw, J. (1997b). The Best Clitic and the Best Place to Put it. Handout to talk presented at the Hopkins

- Optimality Theory Workshop/University of Maryland Mayfest, Baltimore, MD.
- Grimshaw, J. and V. Samek-Lodovici. (1995). Optimal Subjects. In J. Beckman, L. Walsh Dickey, and S. Urbanczyk (eds.), *Papers in Optimality Theory*, GLSA, UMass, 589-605.
- Grimshaw, J. and V. Samek-Lodovici. (1988). Optimal Subjects and Subject Universals. In P. Barbarosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds), *Is the Best Good Enough? Proceedings of the Workshop on Optimality in Syntax*. MIT Press and MIT Working Papers in Linguistics.
- Halpern, A. L. (1995). *On the Placement and Morphology of Clitics*. CSLI Publications.
- Hauge, K. R. (1976). *The Word Order of Predicate Clitics in Bulgarian*. Meddelelser N0. 10. University of Oslo.
- Izvorski, R. (1995). Wh-movement and Focus-movement in Bulgarian. In R. Eckardt and V. van Geenhoven (eds.), *Console II Proceedings*. The Hague: Holland Academic Graphics, 235-254.
- Izvorski, R., T. Holloway King, and C. Rudin. (1997). Against *li* Lowering in Bulgarian. *Lingua* 102:2/3, 187-194.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press.
- Jacobs, H. (1996). An Optimality-Theoretic Analysis of Phonological and Syntactic Aspects of Enclisis and Proclisis in Old French, Brazilian, and European Portuguese. Rutgers Optimality Archives.
- Jaeggli, O. (1986). Three Issues in the Theory of Clitics: Case, Double NPs, and Extraction. In H. Borer (ed), *Syntax and Semantics 19: The Syntax of Pronominal Clitics*. Academic Press, New York, 15-42.
- Kayne, R. (1994). *The Antisymmetry of Syntax*. MIT Press.
- King, T. H. (1995). *Configuring Topic and Focus in Russian*. CSLI Publications.
- Klavans, J. (1985). The Independence of Syntax and Phonology in Cliticization. *Language* 61.1: 95-120.
- Krapova, I. (1997). Auxiliaries and complex tenses in Bulgarian. In Wayles Browne, Ewa Dornisch, Natasha Kondrashova & Draga Zec (eds.). *Annual workshop on Formal approaches to Slavic linguistics. The Cornell meeting, 1995*. (= Michigan Slavic materials, no. 39.) Ann Arbor: Michigan Slavic Publications, 320-344.
- Kroeger, P. (1993). *Phrase Structure and Grammatical Relations in Tagalog*. CLSI Publications.
- Ladd, D. R. (1986). Intonational Phrasing: The Case for Recursive Prosodic Structure. *Phonology Yearbook* 3, 311-340.
- Ladd, D. R. (1996). *Intonational Phology*. Cambridge University Press.
- Lambova, M. D. (1996). *Optional Participle Auxiliary Orders: A Study of the Syntax of Verbal Inflection in Bulgarian*. M.A. Thesis, Temple University.
- Legendre, G. (1996). Clitics, Verb (Non)-Movement, and Optimality in Bulgarian. Johns Hopkins University Technical Report JHU-CogSci-96-5.
- Legendre, G. (1997). Optimal Romanian Clitics: A Cross-linguistic Perspective. Johns Hopkins University Technical Report JHU-CogSci-97-9.
- Legendre, G. (1998). Second Position Clitics in a V2 Language: Conflict Resolution in Macedonian. To appear in Proceedings of ESCOL 1997. CLC Publications, Cornell University.
- Legendre, G., C. Wilson, P. Smolensky, K. Homer, and W. Raymond. (1995). Optimality and Wh-Extraction. In J. Beckman, L. Walsh Dickey, and S. Urbanczyk (eds.), *Papers in Optimality Theory*, GLSA, UMass, 607-636.
- Legendre, G., P. Smolensky, and C. Wilson. (1998). When is Less More? Faithfulness and Minimal Links in Wh-Chains. In P. Barbarosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds), *Is the Best Good Enough? Proceedings of the Workshop on Optimality in Syntax*. MIT Press and MIT Working Papers in Linguistics.
- McCarthy, J. and A. Prince. (1993a). *Prosodic Morphology I; Constraint Interaction and Satisfaction*. To appear, MIT Press.
- McCarthy, J. and A. Prince. (1993b). Generalized Alignment. *Yearbook of Morphology*, 79-153.
- McCarthy, J. and A. Prince. (1996). *Prosodic Morphology 1986*. Technical Report #32, Center for Cognitive Science, Rutgers University.

- Nespor, M. and I. Vogel. (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Perlmutter, D.M. (1971). *Deep and Surface Structure Constraints in Syntax*. Holt, Reinhart, and Winston.
- Pollock, J.-Y. (1989). Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20: 365-424.
- Prince, A. and P. Smolensky. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. To appear in MIT Press.
- Radanović-Kocić, V. (1996). The Placement of Serbo-Croatian Clitics: A prosodic Approach. In A.L. Halpern and A.M. Zwicky (eds.), *Approaching Second: Second Position Clitics and Related Phenomena*. CSLI Publications.
- Richardson, M. 1997. Czech Clitics: 2P or not 2P that is the Question. In L. Kaiser (ed.), *Yale A-Morphous Linguistics Essays*. Yale Department of Linguistics. 131-150.
- Rivero, M. L. 1993. Bulgarian and Serbo-Croatian Yes-No Questions: V<sup>0</sup>-Raising to -li vs. -li Hopping. *Linguistic Inquiry* 24:567-575.
- Rivero, M. L. 1994a. Clause Structure and V-Movement in the Languages of the Balkans. *Natural Language and Linguistic Theory* 12:63-120.
- Rivero, M. L. (1994b). On Two Locations for Complement Clitic Pronouns: Serbo-Croatian, Bulgarian, and Old Spanish. in A. van Kemenade and N. Vincent (eds.), *Inflection and Syntax in Language Change*. Cambridge University Press. **check date**
- Rudin, C. 1986. *Aspects of Bulgarian Syntax: Complementizers and Wh Constructions*. Slavica Publishers, Inc.
- Rudin, C. 1997. AgrO and Bulgarian pronominal clitics. In Martina Linseth & Steven Franks (eds.). Annual workshop on Formal approaches to Slavic linguistics: The Indiana meeting, 1996. (= Michigan Slavic materials, no. 42.) Ann Arbor: Michigan Slavic Publications, 224-252.
- Rudin, C., C. Kramer, L. Billings, and M. Baerman. 1999. Macedonian and Bulgarian *li* questions: Beyond Syntax. *Natural Language and Linguistic Theory* 13, 3: 541-586.
- Samek-Lodovici, V. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. Ph.D. Dissertation, Rutgers University.
- Selkirk, E. O. 1995. The Prosodic Structure of Function Words. In J. Beckman, L. Walsh Dickey, and S. Urbanczyk (eds.), *Papers in Optimality Theory*, GLSA, UMass, 439-469.
- Smolensky, P. (1993). Harmony, Markedness, and Phonological Activity. Handout to talk presented at the Rutgers Optimality Workshop 1, New Brunswick, N.J.
- Smolensky, P. (1995). On the Internal Structure of the Constraint Component *Con* of UG. Handout to talk presented at UCLA.
- Suñer, M. 1988. The Role of Agreement in Clitic-Doubled Constructions. *Natural Language and Linguistic Theory* 6:3, 391-434.
- Todorova, M. (1995). Infinitival Clauses in Bulgarian and the Status of 'da'. McGill University ms.
- Tomić, O. M. (1996). The Balkan Slavic Clausal Clitics. *Natural Language and Linguistic Theory* 14, 811-72.
- Travis, L. (1984). *Parameters and Effects of Word order Variation*. Ph.D. Dissertation, MIT.
- Vakareliyska, C. (1994). Na-drop in Bulgarian. *Journal of Slavic Linguistics* 2(1): 121-50.

\* This paper is a substantially revised version of a talk given at the 1996 Formal Approaches to Slavic Linguistics Meeting in Indiana and of my 1996 JHU Technical Report. I wish to thank the FASL audience, Catherine Rudin, Loren Billings, Andrew Spencer for their comments, as well as Adiamandis Gafos for asking the right question. Special thanks go to my invaluable informants: Marina Todorova, Mariana Lambova, and Boris Nikolov in the United States, Elena Andonova, Gergana Lalova, and Vassil Nikolov in

Bulgaria. I am particularly grateful to Paul Smolensky for his suggestions and his feedback on the final outcome. Research for this paper was partially funded by NSF grant IRI-9596120.

1. Another problem in Rivero (1994a) is that *bjax* is phonologically strong hence a lexical auxiliary in her characterization but it allows LHM just like her other functional auxiliaries. See further discussion below.

2. Optional LHM across a non-clitic auxiliary is analyzed in Legendre (1996).

3. There is in fact no agreement in all previous analyses on what the syntactic positions of various clitics should be. To give a few examples, all clitics but *li* are adjoined to VP in Halpern (1995). In Rivero (1993, 1994a) *li* is in C, *sŭm* is in Aux. Pronominal clitics (= XPs) are in WackernagelP in Rivero (1994b). Rudin (1997), however, takes pronominal clitics to be heads of a sequence of Agr-P projections with AgrO-P dominating AgrS-P. For Tomić (1996), all clitics head functional projections, including *li* in C and object clitics in AgrO. Finally, Izvorski (1995) argues that *li* is not in C but heads a separate phrase FocP. The very fact that there is no agreement on clitic positions in the syntax is, in my opinion, further evidence that the assumption that they are in the syntax at all is on shaky ground.

4. Krapova (1997) further notes that the third person clitic auxiliary *e* may delete under special conditions, e.g. in some renarrated constructions: (i) *Ivana -- pročela knjigata* ‘[They say] Ivana read the book’. This is not the case for its lexical counterpart *beše* which can never be deleted. In our terms (i) is not an instance of deletion; rather the feature combination [perfect] [third person] fails to be realized at PF. It is unclear what the explanation for this phenomenon is though it fits into the universal tendency for [third person] to be tied to distinct morphosyntactic properties (e.g. clitic clustering in Romance, split morphological case systems based on person hierarchies, etc.). Note that it cannot simply be related to the fact that *beše* cannot appear in the renarrated construction: (ii) \**Az bjax bila pročela knjigata* ‘[They say] I had been reading the book’. Corresponding past participles -- *bil* -- show all the properties of lexical auxiliaries; contrary to *bjax*, they may appear in the renarrated construction: (ii) *Az sŭm bil pušil tri kutii na den* ‘[They say] I have smoked three packs a day’.

5. The alternative order *pročel li beše knjigata?* is also possible. Marina Todorova (personal communication) reports that such differences in the position of *li* in the presence of a lexical auxiliary most often relate to differences in the discourse context (though they may be quite subtle). She offers an example of a clear contrast. If someone said the equivalent of ‘Yesterday I saw Ivan in the pub’, then it would be appropriate for the interlocutor to follow with the question *Beše li se napil* ‘Had he gotten drunk?’. If, however, the first comment was something like ‘Yesterday I saw Ivan in the pub and he acted strange’ then the follow-up question would be formulated as *Napil li se beše* ‘Why? Had he gotten drunk?’. In the second case the second speaker is almost offering a guess as to Ivan's state, whereas in the first (s)he is simply asking about it. In the present analysis, these two questions are not discourse-equivalent, hence they correspond to different inputs. They are not part of the same competition.

6. Rudin (1986:135) reports that wh-relatives, in contrast to *deto* relatives, normally may not contain what she calls a resumptive pronoun: (i) \**Tova e deteto koeto go vidjax včera* ‘this is the child who(m) (I) saw yesterday’.

My main informant does not reject these as completely ungrammatical though she prefers dropping the clitic double.

7. The validity of the alignment-based OT theory of clitics is independently demonstrated in Balkan (Legendre, 1997, 1998), South Slavic (Anderson, 1996, this volume), Czech (Richardson, 1997), and Romance (Grimshaw, 1997b).

8. Asking a yes-no question with the complementizer *dali* adds a nuance of ‘wondering out loud’ (Rudin, 1986). Hence the structures with *li* and *dali* do not compete in one single optimization. They correspond to different inputs and tableaux.

9. Any syntactic analysis of clitics must deal with the clustering problem as well and either build it in the structure -- either by stipulating the hierarchical order of projections containing the clitics or stipulating the order of adjunction operations -- or alternatively, reorder them at PF. In fact, all these solutions have been proposed. But they all involve some additional mechanism besides placing clitics in a certain position in the first place (regardless of whether this is accomplished by base-generation or movement).

10. Standard OT conventions include the following:  $\sigma$  = optimal candidate; \* = individual violations of a given constraint; \*! = fatal violations;  $\otimes$  = violations incurred by optimal candidates.

11. Prosodic Inversion moves a clitic (generated in the leftmost position in its syntactic domain) immediately to the right of its prosodic host at PF. The host is crucially assumed to be a prosodic word because Serbo-Croatian clitics may appear in second position after the first prosodic word. Morphological Merger is a similar operation but movement is to the right of a lexical head (rather than a prosodic word). As we will see in section 4, Bulgarian effects similar to the Serbo-Croatian ones are naturally handled in the present approach without recourse to movement.

12. Tomić (1996) independently claims that *ne* and *šte* are clitics but her analysis does not recognize violable constraints. Hence, she is led to stipulate a PF mechanism of cancellation of prosodic subcategorizations to reconcile the fact that clause-initial *ne* and *šte* are (in her terms) phonologically proclitics while second-position clitics are phonologically enclitics.

13. Rivero (1994a) assumes that Bulgarian *da* occupies the head of a Modal Phrase, otherwise occupied by the future clitic *šte*. But note the relative order of *ne* and the modal element: *da* precedes *ne* while *šte* follows *ne*.

(i) *Ivan da ne xodi tam!* ‘Ivan should not go there’ (Dimitrova-Vulchanova, 1993) (ii) *Ne šte sŭm mu go dal.* ‘I will not give it to him’ (Hauge, 1976). This shows that *da* and *šte* cannot both be exponents of the same Modal node. Both are clitics and the relevant constraint ranking is: EDGEMOST(MOD) >> EDGEMOST(NEG) >> EDGEMOST(FUT) >> NONINITIAL.

14. Local conjunctions operate in all components of the grammar. See Legendre et al. (1998) for an example in syntax and Legendre (1997) for a further example in Romanian morphosyntax.

15. I am assuming that topicalized phrases are adjoined to the highest maximal projection by virtue of the absence of subject-verb inversion and the effect of high-ranked constraint \*STRUCTURE mentioned previously.

16. Beckmann and Pierrehumbert (1986) in fact recognize two types of intonational phrases, roughly small ones and big ones. Ladd (1986) argues that Intonational Phrases constitute recursive structures.

17. The fact that Tagalog clitics are prosodically autonomous should not raise eyebrows either. As it turns out, any Bulgarian clitic (though a single one at a time) may also bear stress, provided it is in a very specific context -- immediately after *ne* (Hauge, 1976). This is one of the issues addressed in the next section.

18. Nespor and Vogel (1986) propose that clitics belong to a larger constituent called the Clitic Group. Selkirk (1995), however, argues against the existence of the Clitic Group.

19. I am also assuming that lexical heads must be stressed, though this is hardly controversial. The corresponding constraint may be undominated in Bulgarian or in Gen. I leave the matter open for now. That all syllables are parsed into Prosodic Phrases is presumably in Gen.

20. Thanks to Paul Smolensky for his suggestions in developing this proposal.

21. Rudin et al. interpret this pattern as the ‘remnant of the second position effects in the placement of *li*’. The present paper claims, however, that Bulgarian *li* is fundamentally a second-position clitic. Its surface distribution to some extent masks this reality, due to the complex interaction of relevant constraints.

22. Because *ne* is also used to negate constituents, it is likely that the domain of EDGEMOST(NEG) will also need to be refined along the same lines as that of *li*. I leave this matter open for now.