

## **The Role of Prosody in the Linearization of Clitics: Evidence from Bulgarian and Macedonian\***

*Boris Harizanov*  
University of California, Santa Cruz

### **1 Introduction**

Macedonian and Bulgarian are genetically related Eastern South Slavic languages and geographically related members of the *Balkan Sprachbund*. Yet, Macedonian and Bulgarian exhibit drastically different patterns of clitic placement. While verb-adjacent clitics in both languages may either precede or follow the verb, only Macedonian clitics can appear in an initial position:<sup>1</sup>

- (1) a. dade            mu            ja  
      (s)he.gave 3.SG.M.DAT 3.SG.F.ACC  
      ‘(s)he gave it to him’                    (Bulgarian: ok, Macedonian: \*)
- b. mu            ja            dade  
          3.SG.M.DAT 3.SG.F.ACC (s)he.gave  
          ‘(s)he gave it to him’                    (Bulgarian: \*, Macedonian: ok)

The traditional explanation of this contrast relies on a constraint against initial clitics, which happens to be present in Bulgarian but absent in Macedonian (see e.g. Franks 2008, p. 102, and references therein).

---

\* Many thanks to Karlos Arregi, Ryan Bennett, Gabriela Caballero, Sandra Chung, Emily Elfner, Vera Gribanova, Junko Ito, Paul Kiparsky, Jim McCloskey, Armin Mester, Jaye Padgett, Draga Zec, the members of the UC Santa Cruz Phonetics & Phonology Group, and the anonymous FASL reviewers for feedback on various aspects of this work.

<sup>1</sup> Clitics are glossed as combinations of person (1, 2, 3), number, gender (when applicable), and case. Abbreviations: ACC – accusative, DAT – dative, F – feminine, IMP – imperative, M – masculine, PRESP – present participle, Q – question particle, REFL – reflexive, SG – singular. Transliteration standard: ISO 9 (1968).

Different implementations of this constraint exist but according to this general approach the basic parametric difference between the two languages is the directly observable (in)tolerance of initial clitics.

Here I ask whether the difference in clitic placement between Bulgarian and Macedonian correlates with any other properties of these languages. If so, such correlations should point the way to an elaboration of the traditional analysis that attains a higher degree of explanatory adequacy by theoretically relating systematic differences between Bulgarian and Macedonian. I pursue such an analysis, demonstrating that the difference in clitic placement is predictable from independently observable differences in the prosodic behavior of clitics in the two languages. This result has the desirable theoretical consequence that it eliminates the need for a language specific constraint which prohibits clitics from the initial position within some prosodic constituent. My proposal is based on the idea that only clitics that are attached high in the prosodic hierarchy can potentially interact with constraints on prosodic well-formedness that make reference to prosodic constituents which are high in the hierarchy (e.g. the utterance). Since Macedonian clitics are adjoined below the Prosodic Word level but Bulgarian clitics are adjoined higher, only the latter are linearized in non-canonical positions, as a response to the requirements of a prosodic constraint which prohibits phonologically deficient elements at the left edge of the utterance.

The approach to the linearization of clitics in Bulgarian and Macedonian which I pursue here informs a number of theoretical issues concerning the syntax-phonology interface. First, indirect-reference theories of the interface posit a prosodic level of representation as a mediator between syntax and phonology; a question of central importance in this context is how prosody interacts with other components of the interface—in particular, linearization. The present findings suggest, in agreement with much previous literature, that prosodically deficient elements (clitics) are linearized in non-canonical positions to avoid prosodically deviant structures. Another question then concerns what structures qualify as deviant, or equivalently, what is the nature of constraints on prosodic well-formedness? The present investigation leads to the conclusion that, at least in this case, constraints on prosodic well-formedness are subject to a certain kind of visibility restriction, so that they may target specific prosodic constituents to the exclusion of others, which remain “invisible” to them.

This paper is organized as follows. Section 2 provides an overview of the central contrast in clitic placement between Bulgarian and Macedonian. The analysis of this contrast is outlined in section 3, and various details and consequences of the analysis are discussed in section 4. Finally, section 5 summarizes the proposal and concludes with some of the questions that it raises.

## 2 The Contrast in Clitic Placement

It is well-known that while Macedonian pronominal and auxiliary clitics appear in initial positions, their Bulgarian counterparts never do and, instead, exhibit variable placement sensitive to prosodic context. This section recapitulates the relevant empirical details.

### 2.1 *Macedonian Clitics*

Verbal clitics in Macedonian immediately precede their host if it is tensed:<sup>2</sup>

- (2) a. kučeto ja kasa mačkata  
           the.dog 3.SG.F.ACC bites the.cat  
           ‘the dog bites the cat’  
       b. \*kučeto kasa ja mačkata (Friedman 2001, p. 37)
- (3) a. neizinata stara majka ti go dala  
           her old mother 2.SG.DAT 3.SG.M.ACC given  
           ‘her old mother has given it to you’  
       b. \*neizinata stara majka dala ti go (Tomić 1996, p. 830)

In the absence of topicalized or focused constituents in the left periphery of the clause and when subjects are dropped, the clitics occupy the initial position in the clause and the clitic-host order is preserved:

- (4) a. go vikna deteto  
           3.SG.M.ACC (s)he.called the.child  
           ‘(s)he called the child’  
       b. \*vikna go deteto (Vidoeski 2005, p. 16)

---

<sup>2</sup> For present purposes, tensed hosts include lexical verbs that bear tense morphology and *-l* participles (Joseph 1983).

- (5) a. Ti ja dade li Ana vaznata?  
 2.SG.DAT 3.SG.F.ACC gave Q Ana the.vase  
 ‘Did Ana give you the vase?’  
 b. \*Dade li ti ja Ana vaznata? (Tomić 1996, p. 825)

The clitics immediately follow their host if the host is non-tensed—i.e. an imperative, a participle (except an *-l* participle; see fn. 2), or a non-verbal predicate (Tomić 1996, p. 824, 862):

- (6) a. daj mi ja knjigata  
 give.IMP 1.SG.DAT 3.SG.F.ACC the.book  
 ‘give me the book’  
 b. \*mi ja daj knjigata
- (7) a. davajk'i mi ja knjigata, me bakna  
 give.PRESP 1.SG.DAT 3.SG.F.ACC the.book 1.SG.ACC kissed  
 ‘giving me the book, (s)he kissed me’  
 b. \*mi ja davaik'i knjigata, me bakna
- (8) a. mil mi e Petko  
 dear 1.SG.DAT is Petko  
 ‘Petko is dear to me’  
 b. \*mi e mil Petko

Thus, the position of Macedonian clitics of the kind discussed above can be predicted entirely on the basis of morphosyntactic information like the tenseness of their host. The placement of these clitics, therefore, does not appear to be affected by prosodic requirements.

## 2.2 Bulgarian Clitics

Verbal clitics in Bulgarian immediately precede their host by default:<sup>3</sup>

- (9) a. Petko vinagi mi go dava  
 Petko always 1.SG.DAT 3.SG.M.ACC gives  
 ‘Petko always gives it to me’  
 b. \*Petko vinagi dava mi go

<sup>3</sup> They do not occupy the second position in the clause under any definition of “second”, since, as long as they remain verb-adjacent, they can be arbitrarily distant from the left edge of the clause, as (9a) shows.

This default behavior, however, is disrupted in case the clitic-host order would leave the clitics in the initial position. For example, if the subject in (9a) is right-dislocated and there is no other material to the left of the verb, the clitics must appear to the immediate right of their host, the verb:

- (10)a. dade mi go Petko  
 gave 1.SG.DAT 3.SG.M.ACC Petko  
 ‘Petko gave it to me’  
 b. \*mi go dade Petko

By contrast with Macedonian, tenseness of the host does not affect clitic placement in Bulgarian. For example, clitics must follow an imperative verb if it is leftmost within the clause, as in (11a). However, they will precede it in the presence of a clause-initial adverbial, as in (11c) (cf. section 2.1 and examples (6)-(8) in particular).<sup>4</sup>

- (11)a. daj mi ja knjigata  
 give.IMP 1.SG.DAT 3.SG.F.ACC the.book  
 ‘give me the book’  
 b. \*mi ja daj knjigata  
 c. b̄rzo mi ja daj knjigata  
 quickly 1.SG.DAT 3.SG.F.ACC give.IMP the.book  
 ‘give me the book quickly’

Other non-tensed hosts, such as participles and non-verbal predicates, behave similarly.<sup>5</sup> The conclusion then is that while syntax might be a necessary determinant of the position of the Bulgarian clitic cluster, it is not sufficient. The process of linearization of clitics in this language must, in addition, make reference to prosodic context (cf. Bošković 2001).

---

4 In Macedonian, on the other hand, clitics remain postverbal in the presence of additional preverbal material (example (20b) from Franks 2009, p. 206):

- (i) utre kupuvaj go penkaloto  
 tomorrow buy.IMP 3.SG.N.ACC the.pen  
 ‘buy the pen tomorrow’

5 Adverbial participles (e.g. *četejki knjigata* ‘reading the book’; also referred to as gerunds or verbal adverbs) constitute a possible exception, as they invariably precede the clitics that they host.

### 3 Prosodically Driven Clitic Placement

My treatment of clitic placement in Bulgarian and Macedonian, as well as the contrasts between them, relies on the intuition that postverbal clitic placement in Bulgarian results from an interaction between the phonological properties of clitics and principles of prosodic well-formedness, defined over prosodic constituents. Preverbal clitic placement is, thus, avoided when it yields prosodic deviance, with a prosodically deficient element pronounced in a position reserved for prosodically strong elements.

The analysis is developed on the assumption that prosodic structures are organized hierarchically in constituents, which are based on, but not isomorphic to, syntactic constituents (Selkirk 1995, among others).<sup>6</sup>

- (12) Prosodic constituents above the level of the Foot
- a.  $\iota$ : Intonational Phrase
  - b.  $\varphi$ : Phonological Phrase
  - c.  $\omega$ : Prosodic Word

Next, I adopt a purely phonological conception of what clitics are and define them as phonological forms (realizing morphosyntactic elements) which are deficient in prosodic structure at the level of the Prosodic Word (see Anderson 2011 for a similar view). In other words, clitics are phonological strings whose segmental content may be organized into syllables and possibly feet, but which are not lexically assigned the status of Prosodic Words.

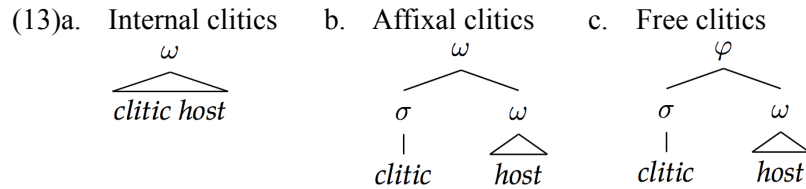
With these two central assumptions in place, the main ingredients of the analysis can be given as follows. First, Macedonian clitics are adjoined below the Prosodic Word level, while Bulgarian clitics are adjoined higher (see section 3.1). Second, a *STRONG START* constraint prohibits prosodically deficient material from appearing at the left edges of certain prosodic constituents (see section 3.2). The immediate consequence of these conjectures is that only clitics adjoined above the Prosodic Word level—as in Bulgarian—can potentially violate *STRONG START*. In other words, clitics in Bulgarian appear postverbally in order to avoid prosodic deviance.

---

<sup>6</sup> I follow Ito and Mester 2012 in assuming an impoverished inventory of prosodic constituents, in which the Utterance is replaced by the *maximal* Intonational Phrase. For present purposes nothing hinges on this choice.

### 3.1 *The Prosodic Organization of Clitics*

The first main component of the analysis concerns the prosodic structures that clitics in Bulgarian and Macedonian participate in. Work on the prosodic typology of function words has established that clitics may adjoin at different prosodic levels (see e.g. Selkirk 1995 and Ito and Mester 2009):



Internal clitics are parsed inside the Prosodic Word defined by their host. Affixal clitics, on the other hand, attach outside of the minimal Prosodic Word that immediately contains the host but are still dominated by a Prosodic Word, giving rise to a recursive Prosodic Word structure. Finally, free clitics are immediately dominated by the next higher prosodic constituent, the Phonological Phrase, and are completely external to the Prosodic Word of the host.

The working hypothesis here—motivated at length in section 4—is that the clitics in Macedonian are internal and/or affixal (cf. Zec 2005, section 3 on Serbian, and Peperkamp 1997 and Anderson 2011 on Neapolitan and Lucanian; see also Rudin et al. 1999). On the other hand, clitics in Bulgarian are free (cf. Zec 2005, section 2 on Serbian, Peperkamp 1997 and Anderson 2011 on Standard Italian; Selkirk 1995, p. 198 on English; see also Rudin et al. 1999).

### 3.2 *STRONG START*

The second component of the analysis has to do with the crosslinguistic tendency for edges of prosodic constituents that are relatively high on the Prosodic Hierarchy to be positions of prosodic strength, so that no weak elements are ever found there (see e.g. Selkirk 2010). This is certainly true in Bulgarian, where, informally, the left edge of maximal Intonational Phrases (i.e. Utterances; see fn. 5) does not tolerate material that is not parsed inside a Prosodic Word. Following recent work on other languages (Selkirk 2010, Elfner 2011, Bennett et al. to appear), I hypothesize that the following formal constraint is behind such effects in Bulgarian:

(14) STRONG START

The leftmost constituent of a maximal Intonational Phrase should not be a prosodically deficient element (i.e. such an element must be parsed inside a Prosodic Word).

The effects of STRONG START in Bulgarian amount to the requirement that the left edge of a maximal Intonational Phrase must coincide with the left edge of a Prosodic Word. Section 4.2 provides evidence that the relevant prosodic domain is, in fact, the maximal Intonational Phrase.

A constraint similar to STRONG START can be assumed to be active at the right edge of Phonological Phrases in English and to account for the examples in (15) (Selkirk 1995, p. 200).<sup>7</sup> In this position, only the phonologically strong form of a function word is allowed:

- (15)a. I don't know whether Ray is            [ɪz], \*[z]  
b. I can eat more than Sarah can        [kæn], \*[kən], \*[kn]

Likewise, a version of STRONG START has been argued to be active at the left edge of Phonological Phrases in Irish (Elfner 2011 and Bennett et al. to appear). Weak pronouns in Irish are not tolerated in this position, and the language exhibits at least two ways of avoiding potentially deviant prosodic structures: postposing, as in (16a), or in-situ strengthening, as in (16b) (examples (2) and (5c) from Bennett et al. to appear).

- (16)a. Fuair sé \_\_\_ óna dheartháir an lá cheana é.  
got he from-his brother the other day it  
'He got it from his brother the other day.'  
b. Thóg siad í ar bord.  
raised they her on board  
'They lifted her on board.'

As far as Bulgarian is concerned, the same ways of avoiding STRONG START violations should be available in principle. However, strengthening of the clitic as in English and Irish, which amounts to it being parsed as a Prosodic Word, or at least a foot, is not attested in Bulgarian: the language does not allow promotion of clitics to independent Prosodic

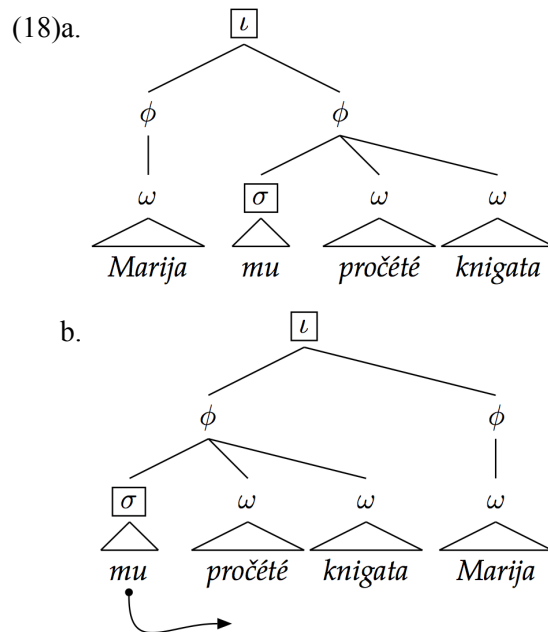
---

<sup>7</sup> Werle (2009), discusses a potential instance of Utterance non-finality in Serbian (p. 343ff.), citing data from Browne 1975 (p. 120). However, the effect is not exactly the mirror image of what is observed in Bulgarian because Serbian clitics are, in fact, able to appear Utterance-finally if there is no viable alternative way to linearize the structure.



Words. Instead, it exhibits variable clitic placement, whereby clitics are linearized in non-canonical positions, as in Irish (cf. (16a)), to avoid a violation of the prosodic constraint. Thus, the canonically preverbal clitic in (17a) with the prosodic structure in (18a) appears after the verb when the subject is right-dislocated, as shown in (17b) and (18b).<sup>8</sup>

- (17)a. Marija mu pročete knjigata  
 Maria 3.SG.M.DAT read the.book  
 ‘Maria read the book to him’  
 b. — pročete mu knjigata Marija  
 read 3.SG.M.DAT the.book Maria  
 c. \*mu pročete knjigata Marija  
 3.SG.M.DAT read the.book Maria



### 3.3 Summary

Given the prosodic structures posited in 3.1 and the conjecture from 3.2 that STRONG START does not tolerate clitics at the left edge of maximal Intonational Phrases, only the Bulgarian clitics are expected to

<sup>8</sup> This account is not intended to be interpreted as necessarily involving literal rightward movement of the clitic. It is, in fact, consistent with a number of different implementations of linearization (including copy pronunciation, prosodic inversion, OT competition, etc.).

potentially violate STRONG START. This is so because they are adjoined above the Prosodic Word level and are immediately dominated by a Phonological Phrase, while Macedonian clitics, being inside Prosodic Words, are in a sense invisible to this version of STRONG START.

#### 4 Analytical consequences

The account outlined above explains the clitic placement patterns in Bulgarian and Macedonian in terms of differences in prosodic structure, coupled with the effects of STRONG START. The prosodic behavior of clitics in the two languages and some aspects of STRONG START are further elaborated and motivated in sections 4.1 and 4.2, respectively. A notable characteristic of this account is that it shifts the burden of explaining the contrast in clitic placement between Bulgarian and Macedonian to the distinct attachment sites of the clitics in each language. As a result, it leaves open the possibility that STRONG START is actually active in Macedonian as well. Preliminary evidence that this is the case is discussed in section 4.3.

##### 4.1 *Motivating the Prosodic Attachment Sites of Clitics*

It was hypothesized in section 3.1 that Macedonian clitics are parsed inside Prosodic Words (i.e. they are immediately dominated by Prosodic Words), while Bulgarian clitics are immediately dominated by Phonological Phrases. To determine which of these structures is instantiated in each case I survey phonological phenomena that occur at the edges of Prosodic Words or across their boundaries (Booij 1987, Revithiadou 2008, Anderson 2005, p. 40ff).

4.1.1 Macedonian. Clitics in Macedonian participate in phonological processes which apply within Prosodic Words. First, they interact with stress assignment. Macedonian has regular antepenultimate stress and initial stress in mono- or disyllabic words. Relevant for the present purposes is that postverbal clitics affect stress placement, shifting it to the antepenultimate syllable (as has been reported to be the case in Indonesian, Latin, and Modern Greek):

- (19)a. doneSI<sup>9</sup>  
bring.IMP  
'bring'

---

9 Capital letters mark stressed syllables.

- b. doNEsi go  
 bring.IMP 3.SG.M.ACC  
 ‘bring it’
- c. doneSI mi go  
 bring.IMP 1.SG.DAT 3.SG.M.ACC  
 ‘bring it to me’

Preverbal clitics also interact with stress placement, albeit in a more limited set of contexts. For example, they are included in the three-mora stress window and can carry stress themselves in the context of sentential negation and *wh*-questions:<sup>10</sup>

- (20)a. ne GO vide  
 not 3.SG.M.ACC (s)he.saw  
 ‘(s)he didn’t see him’
- b. zošto mu GO dade  
 why 3.SG.M.DAT 3.SG.M.ACC gave  
 ‘why did you give it to him’

On the other hand, preverbal clitics that do not occur in these special contexts do not affect the position of stress, which remains on the verb:

- (21)a. go VIdE  
 3.SG.M.ACC (s)he.saw  
 ‘(s)he saw him’
- b. mu go DAde  
 3.SG.M.DAT 3.SG.M.ACC gave  
 ‘(s)he gave it to him’

However, there is independent evidence that preverbal clitics in general are parsed inside Prosodic Words. A phenomenon that diagnoses this behavior is a type of vowel deletion: preverbal clitics of the form CV lose their vowel V when followed by a vowel-initial stem (Vidoeski 2005, p. 21):<sup>11</sup>

10 An anonymous reviewer raises the question of why sentential negation and *wh*-questions affect stress placement in Macedonian in this way. Various aspects of this question have attracted attention in the literature (see e.g. Rudin et al. 1999, Tomić 1996) but the issue is beyond the scope of this paper. One possibility is that the sentential negation marker and *wh*-words are characterized by prosodic requirements of their own, which determine the kind of prosodic structures that these elements can appear in.

11 VV sequences are not generally illegal in the language across word boundaries, which indicates that the V-deletion phenomenon in (22) does only apply within Prosodic Words.

- (22)a. go ostai → [gostai] ‘(s)he left it’<sup>12</sup>  
 b. se utepa → [sutepa] ‘(s)he hurt herself/himself’  
 c. k'e igram → [k'igram] ‘I'll play’  
 d. ne izleze → [nizleze] ‘didn't come out’

Vowel deletion of this kind is a widespread phenomenon in the western dialects of Macedonian (Vidoeski 2005). It affects not only clitics but also other unaccented words, including lexical words, that happen to form an “accentual unit” with the vowel-initial word they immediately precede.

To capture these patterns, I assume that the domain of stress assignment in Macedonian is the *minimal* Prosodic Word, while the domain of morphophonemic alternations like vowel deletion of the kind described above is the *maximal* Prosodic Word. Then, pre- and postverbal clitics that affect stress placement must be parsed internal to the minimal Prosodic Word, as in (23a). On the other hand, preverbal clitics that do not affect stress, illustrated in (21) above, but do participate in vowel deletion must be part of a recursive Prosodic Word structure, as in (23b).



These structures result from the interaction between the prosodic deficiency of clitics, which is lexically specified (Zec 2005), and constraints on prosodic structuring, which in Macedonian must allow recursive Prosodic Words.

4.1.2 Bulgarian. Clitics in Bulgarian do not participate in phonological processes that apply within Prosodic Words. To begin with, stress in Bulgarian is lexical and some morphemes can affect stress placement within a word. In the following example the suffixal definiteness marker attracts the stem-initial stress to itself:

- (24)a. hubost → [hubost] ‘beauty’  
 b. hubost+ta<sub>DEF</sub> → [hubostTA] ‘the beauty’

<sup>12</sup> Square brackets do not imply IPA transcription.

On the other hand, neither pre- nor postverbal clitics affect the position of stress in this way (as has been reported to be the case in Italian, Spanish and Ancient Greek):

- (25)a. doneSI  
bring.IMP  
'bring'
- b. bǎrzo go doneSI  
quickly 3.SG.M.ACC bring.IMP  
'bring it quickly'
- c. doneSI mi go  
bring.IMP 1.SG.DAT 3.SG.M.ACC  
'bring it to me'

Voicing alternations also indicate that Bulgarian clitics are parsed outside of Prosodic Words. Some morphemes, such as the suffixal definiteness marker and the plural suffix, bleed regular word-final devoicing in certain contexts:

- (26)a. mǎž → [mǎš] 'man'  
b. mǎž+a<sub>DEF</sub> → [mǎža] 'the man'  
c. mǎž+e<sub>PLURAL</sub> → [mǎže] 'men'

Clitics do not bleed word-final devoicing in this manner. Consider, for instance, the clitic copula *e*, which happens to be segmentally identical to the plural suffix above, and the third-person feminine dative/possessive clitic *i*:

- (27)a. mǎž+e<sub>COPULA</sub> → [mǎše] 'it's a man'  
b. mǎž+i+e<sub>COPULA</sub> → [mǎšie] 'he is her husband'

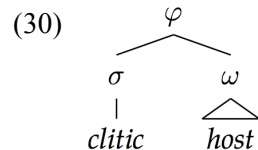
Finally, some morphemes interact with liquid/ǎ metathesis, another process limited to the domain of Prosodic Words, which occurs with certain roots:

- (28)a. grǎb → [grǎp] 'back'  
b. grǎb+a<sub>DEF</sub> → [gǎrba] 'the back'  
c. grǎb+ove<sub>PLURAL</sub> → [gǎrbove] 'backs'

Clitics, however, do not trigger liquid/ǎ metathesis in the same environments:

- (29)a. grăb+e<sub>COPULA</sub> → [grăpe] ‘it’s a back’  
 b. grăb+i → [grăpi] ‘her back’

Assuming that the domain of stress assignment and morphophonemic alternations like word-final devoicing and liquid/ă metathesis is the Prosodic Word, clitics in Bulgarian must attach above the level of the Prosodic Word and be immediately dominated by a Phonological Phrase:



The participation of Macedonian clitics in Prosodic Word-level processes and the non-participation of their Bulgarian counterparts in such processes follows from the prosodic attachment sites of the clitics in each language. It is this difference in their attachment sites, established in this section, that is taken to play a central role in the linearization of clitics and to correlate with (in)tolerance of initial clitics.

There are at least two potential sources of this difference. One possibility is that it is rooted in a syntactic difference between the structures in which clitics participate in Bulgarian and Macedonian. If this is the case, the principles that map syntactic structures to prosodic ones in each language may output distinct prosodic parses because they encounter distinct syntactic structures as input. Another possibility is that the difference in clitic attachment sites stems from a difference in the mapping principles themselves, while the underlying syntax remains more or less the same in both languages. In this case, the relevant difference might be that only Macedonian requires Phonological Phrases to be exhaustively parsed into Prosodic Words. This would have the effect that no Phonological Phrase could directly dominate a clitic, a state of affairs that is the norm in Bulgarian.

While I do not pursue many of the intriguing related questions here, preliminary evidence for the former approach may come from differences that have been postulated in the syntax of cliticization in Bulgarian and Macedonian. In particular, there is ample evidence that Bulgarian clitics behave as nominal phrases in a number of ways (Harizanov, to appear) while Macedonian clitics pattern with agreement markers (Franks 2009); see also Rudin 1997 for relevant discussion.

#### 4.2 *The relevance of STRONG START and its domain of application*

An attempt to model the effects of STRONG START in Bulgarian without reference to a prosodic well-formedness constraint might instead rely on lexically encoded information about the clitics themselves. For instance, it could be suggested that clitic placement in Bulgarian is a matter of prosodic subcategorization—i.e. the clitics are lexically specified as enclitics and as such, they need a host to their left and cannot appear clause-initially. Here I argue that this cannot be the case, following much previous work on the issue (Bošković 2001, Franks and Bošković 2001, Pancheva 2005, Franks 2008, Harizanov 2011, among others).

First, assuming that elements like *i* ‘and’, *no* ‘but’, and *če* ‘that’ in Bulgarian are unsuitable hosts for clitics, it is notable that clitics appear preverbally in the presence of any of these unstressed elements:<sup>13</sup>

- (31) ... *i* / *no* / *če* *mi*            *go*            *dade* *Petko* *včera*  
          and/ but / that 1.SG.DAT 3.SG.M.ACC gave Petko yesterday  
          ‘... and/but/that Petko gave it to me yesterday’

If the clitics were specified as enclitic, they would invert with the verb in (31), since they cannot find a suitable host to lean on to their left. Configurations like these demonstrate that Bulgarian clitics do not have a preference for their direction of phonological attachment and can be either proclitic or enclitic, as long they remain verb-adjacent.<sup>14</sup>

Second, the clitics in Bulgarian can immediately follow Intonational Phrase boundaries, signaled by commas below and introduced by parentheticals and certain types of adjuncts (see also Bošković 2001, p. 218, fn. 37; Pancheva 2005, p. 114, fn. 7; Franks 2008, p. 99ff.):

---

13 At least some of these elements can appear in isolation and in such cases they are, presumably, parsed as Prosodic Words (see Franks 2008, p. 99 and fn. 17 for relevant discussion). However, this does not affect the argument here because in (31) these elements are unstressed and, therefore, do not qualify as suitable clitic hosts.

14 If the inability of elements like *i* ‘and’, *no* ‘but’, and *če* ‘that’ to host clitics is to be analytically related to their own cliticness, the question arises as to why they can appear initially. The difference between them and the clitics that cannot appear initially might be lexical: for instance, they could be specified as internal to Prosodic Words so that they do not interact with STRONG START (like Macedonian clitics). Further interesting questions, which I leave open, arise about the interaction of the two types of clitics in Bulgarian. On the related question about the interaction between clitic placement and coordination, see Franks and Bošković 2001.

- (32)a. Portiera na cirka, šašto armenec,  
 the.gatekeeper at the.circus also an.armenian  
 go pita: ...  
 3.SG.M.ACC asked  
 ‘the circus gatekeeper, (who was) an Armenian too, asked him’
- b. Pritesnena, go popitah dali tova ne krie  
 worried 3.SG.M.ACC I.asked whether this not hides  
 njakakāv risk  
 some risk  
 ‘Worried, I asked him whether this didn't hide any risk’

In these examples, if the clitics required a phonological host to their left, they would not be able to appear right after an Intonational Phrase boundary and would invert with the verb. Both of the patterns just discussed are predicted if clitic placement in Bulgarian is, instead, driven by a constraint against clitics in initial position such as *STRONG START*.

In addition to demonstrating the irrelevance of prosodic subcategorization to the linearization of clitics in Bulgarian, the data in (32) above shows that the relevant domain of clitic non-initiality is the *maximal* Intonational Phrase (i.e. the Utterance; see fn. 5). This follows from the assumption that the boundary at the right edge of the parentheticals and adjuncts in (32) marks the left edge of an Intonational Phrase that immediately contains the clitics. This Intonational Phrase is, however, embedded in the maximal Intonational Phrase that, in this case, contains the whole utterance. If *STRONG START* is relativized to this larger prosodic constituent, it follows that the clitics will appear preverbally in the configurations in (32), since they are not at the left edge of the maximal Intonational Phrase.<sup>15</sup>

#### 4.3 *STRONG START* in Macedonian?

Given the account outlined in section 3, the prosodic attachment site of clitics in Macedonian makes it impossible to detect the potential effects of *STRONG START* in the language because even if the constraint is active, the clitics are invisible to it. The question then arises of whether the

<sup>15</sup> Franks and Bošković (2001) and Harizanov (2011) argue that the relevant domain of clitic non-initiality in Bulgarian is a syntactic one (a phase or a Spell-Out domain, essentially equivalent to a CP constituent). What remains unclear under such an approach, as pointed out by Franks (2008), fn. 21, is why elements defined in terms of their prosodic deficiency should be sensitive to syntactic context, as opposed to prosodic context. The trade-offs between these two approaches are worth comparing as part of a general theory of the syntax-phonology interface.



potential activity of STRONG START in Macedonian can be detected in any way at all. Assuming that clitics can have distinct prosodic attachment sites, and that STRONG START is active in both Bulgarian and Macedonian, we might expect to find clitics in Macedonian that are adjoined above the Prosodic Word and that, as a result, cannot be initial within maximal Intonational Phrases. In essence, such clitics would behave indistinguishably from their Bulgarian counterparts that I have been concerned with.

An element that potentially fits the bill is the question particle *li*, which cannot appear initially even though it occupies the highest syntactic position—the head C of the clause CP (Rudin et al. 1999, p. 543):<sup>16</sup>

- (33)a. Go            vide li?            b. \*Li go            vide?  
       3.SG.M.ACC saw Q                Q 3.SG.M.ACC saw  
       ‘Did (s)he (s)he.see him?’

As predicted by the present account, the sensitivity of *li* to STRONG START correlates with its prosodic attachment site, which differs from that of the other Macedonian clitics. For instance, unlike them, *li* does not affect stress placement in Macedonian (Rudin 1999, p. 552):

- (34)a. doNEsuvaš                      b. doNEsuvaš li?  
       you.bring                        you.bring Q  
       ‘you are bringing’            ‘are you bringing?’

The question particle, *li* must then be adjoined above the Prosodic Word level, just like the Bulgarian clitics, with which it shares the behavior illustrated in (34). Therefore, it is plausible that STRONG START is active in Macedonian as well.<sup>17</sup>

16 Linearizing *li* within the Prosodic Word that contains the pronominal clitic and the verb (*\*Go li vide?*) would satisfy STRONG START but is excluded by whatever principles ensure that Prosodic Words are atomic with respect to linearization and that force initial clitics to invert with the Prosodic Word to their right.

17 It is of course possible to analyze the inability of *li* to appear initially in terms of its lexical specification as an enclitic (cf. the discussion in section 4.2). However, such an approach requires additional stipulations and multiplies the types of clitics that need to be posited in the language. On the other hand, the analysis outlined here only relies on general prosodic principles and ties the behavior of *li* to its prosodic attachment site, which would be accidental under the lexical specification approach.

## 5 Concluding remarks

In sum, the position of clitics relative to their verbal host varies as a function of prosodic context in Bulgarian: clitics appear preverbally or postverbally. In Macedonian, on the other hand, as long as their host is tensed, clitics appear preverbally. Variability in clitic placement correlates with the adjunction site of clitics—in particular, with whether they are adjoined above the Prosodic Word level (Bulgarian) or below this level (Macedonian). The proposed analysis relates these facts causally by recognizing the role of principles of prosodic well-formedness in linearization. Specifically, the contrast in clitic placement between Bulgarian and Macedonian follows from an interaction between the prosodic adjunction site of clitics and a **STRONG START** constraint, which prohibits clitics that are not parsed inside Prosodic Words from appearing at the left edge of maximal Intonational Phrases. This account does not require reference to the lexical specifications of clitics, deriving this aspect of their behavior on the basis of general prosodic principles. It also connects clitic placement analytically to an independently observable fact of the prosodic organization of each language.

Central to the characterization of **STRONG START** is the assumption that it is restricted so that it interacts with clitics but cannot “see” inside Prosodic Words. If the constraint is universally restricted in this way, the analysis predicts that the clitics in a language that can be reliably shown to have this version of **STRONG START** may not be subject to variable placement (or any other way of sidestepping the otherwise ensuing prosodic deviance, like strengthening) if they participate in Prosodic Word-level phonological processes; they simply would not be visible to the constraint. If, instead, there is crosslinguistic variation with respect to the domains that are visible to **STRONG START**, the question arises of whether there is some language in which the constraint imposes requirements on prosodic structure below the Prosodic Word level. For instance, are iambs ever prohibited from appearing at the left edge of Phonological Phrases? A more general related question is whether we find any interactions between Prosodic Word-level phonological processes and prosodic constituents that are “too high” on the prosodic hierarchy (e.g. Utterance-final devoicing; see Nespors and Vogel 1986 for relevant discussion). Answers to these questions will undoubtedly bring us closer to a deeper understanding of the role of prosody in the mapping from syntax to phonology and of the syntax-phonology interface, in general.

## References

- Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Anderson, Stephen R. 2011. Clitics. In *The Blackwell Companion to Phonology*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 2002–2018. Blackwell Publishing.
- Bennett, Ryan, Emily Elfner, and James McCloskey. To appear. Lightest to the right: An apparently anomalous displacement in Irish. *Linguistic Inquiry*.
- Booij, Geert. 1988. Review article on Nespor and Vogel (1986). *Journal of Linguistics* 24:515–525.
- Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier.
- Browne, Wayles. 1975. Serbo-Croatian enclitics for English-speaking learners. In *Contrastive analysis of English and Serbo-Croatian I / Kontrastivna analiza engleskog i hrvatskog ili srpskog jezika I*, ed. Rudolf Filipović, 105–134. Zagreb: Institute of Linguistics.
- Elfner, Emily. 2011. Syntax-prosody interactions in Irish. Doctoral Dissertation, University of Massachusetts, Amherst.
- Franks, Steven. 2008. Clitic placement, prosody, and the Bulgarian verbal complex. *Journal of Slavic Linguistics* 16:91–137.
- Franks, Steven. 2009. Macedonian pronominal clitics as object agreement markers. In *A linguist's linguist: Studies in South Slavic linguistics in honor of E. Wayles Browne*, ed. Steven Franks, Vrinda Chidambaram, and Brian Joseph, 189–222. Bloomington, IN: Slavica Publishers.
- Franks, Steven, and Željko Bošković. 2001. An argument for multiple Spell-Out. *Linguistic Inquiry* 32:174–183.
- Friedman, Victor A. 2001. *Macedonian reference grammar*. Slavic and Eurasian Language Resource Center, Duke University.
- Harizanov, Boris. 2011. NonInitiality within spell-out domains: Unifying the post-syntactic behavior of Bulgarian dative clitics. In *Morphology at Santa Cruz: Papers in honor of Jorge Hankamer*, ed. Nick LaCara, Anie Thompson, and Matt A. Tucker. Santa Cruz, CA: UCSC Linguistics Research Center.
- Harizanov, Boris. To appear. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language & Linguistic Theory*.

- Ito, Junko, and Armin Mester. 2012. Recursive prosodic phrasing in Japanese. In *Prosody matters: Essays in honor of Elisabeth Selkirk*, ed. Toni Borowsky, Shigeto Kawahara, Takahito Shinya, and Mariko Sugahara. London: Equinox.
- Joseph, Brian D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general, and historical linguistics*. Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.
- Pancheva, Roumyana. 2005. The rise and fall of second-position clitics. *Natural Language & Linguistic Theory* 23:103–167.
- Peperkamp, Sharon. 1997. *Prosodic words*. HIL dissertations 34. The Hague: Holland Academic Graphics.
- Revithiadou, Anthi. 2008. A cross-dialectal study of cliticization in Greek. *Lingua* 188:1393–1415.
- Rudin, Catherine. 1997. Agr-O and Bulgarian pronominal clitics. In *Formal approaches to Slavic linguistics: The Indiana meeting, 1996*, ed. Martina Lindseth and Steven Franks, 224–252. Ann Arbor: Michigan Slavic Publications.
- Rudin, Catherine, Christina Kramer, Loren Billings, and Matthew Baerman. 1999. Macedonian and Bulgarian LI questions: beyond syntax. *Natural Language & Linguistic Theory* 17:541–586.
- Selkirk, Elisabeth. 1995. The prosodic structure of function words. In *Signal to syntax: bootstrapping from speech to grammar in early acquisition*, ed. J. Morgan and K. Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum.
- Selkirk, Elisabeth. 2010. The syntax-phonology interface. In *The handbook of phonological theory*, ed. John Goldsmith, Jason Riggle, and Alan Yu. Wiley Blackwell, 2 edition.
- Tomić, Olga Mišeska. 1996. The Balkan Slavic clausal clitics. *Natural Language & Linguistic Theory* 14:811–872.
- Vidoeski, Božidar. 2005. *Dialects of Macedonian*. Bloomington, IN: Slavica Publishers.
- Werle, Adam. 2009. Word, phrase, and clitic prosody in Bosnian, Serbian, and Croatian. Doctoral Dissertation, University of Massachusetts, Amherst.
- Zec, Draga. 2005. Prosodic differences among function words. *Phonology* 22:77–112.

bharizan@ucsc.edu