Syntax and prosody: *li* questions in Macedonian and Bulgarian

Loren Billings, Christina Kramer, and Catherine Rudin

Abstract. The distribution of the yes/no-interrogative clitic li in Macedonian and Bulgarian reveals a complex interaction of syntax with prosodic factors. The underlying syntactic uniformity of li questions in the two languages is obscured by a series of prosodic phenomena affecting one language or the other. In Macedonian two prosodic factors affect the placement of li: optional stressing of auxiliaries and optionally allowing certain sequences of words to have only one stress. In Bulgarian two different prosodic phenomena are relevant: stressing of clitics after the negative element ne and inversion of initial clitics with the following verb. When these four factors are controlled for, the syntax of li questions in the two languages is startlingly homogeneous. If no element is focused (i.e., moved to SpecCP), then, in both languages, the tensed verb head-incorporates into li in C. Usage differences complicate the picture somewhat as well.

1. Introduction

Much recent work in theoretical linguistics relies upon cross-linguistic comparison to elucidate the limits and causes of variation in Universal Grammar; that is, to establish just what structures and processes are possible in human languages. It has often proven especially useful to compare related languages, thus teasing out differences between languages which share much of their grammatical structure. The two languages considered in this article, Macedonian and Bulgarian, are very closely related genetically, both being South Slavic. In addition, they are areally related: both participate in the Balkan *Sprachbund*, sharing many of the contact-induced grammatical features common to the Balkan area. It is thus of particular interest to note cases in which the grammars of the two languages diverge. We expect that differences between such closely related languages will involve rather superficial factors, changes which could be introduced into a grammar in a relatively short time, even if their surface effects are considerable.

We examine here one construction, the yes/no question formed with the interrogative clitic li, which displays interesting similarities and differences in the two languages. Superficially, Macedonian and Bulgarian appear to have distinct conditions on the placement of li, particularly with respect to the other clitics. We argue, however, that the syntactic behavior of li is fundamentally identical in the two languages, as is their clitics' placement. Several prosodically realized differences between the two languages interact with their common syntax to

1

^{*}Thanks to the following colleagues for their assistance: P. Ambrosiani, W. Browne, D. Ćavar, L. Hammer, R. Izvorski, T. King, B. Nilsson, J. Toman, O. Tomić, M. Votruba, D. Zec and audiences at annual AATSEEL meetings (Billings 1994, Rudin 1993b, Rudin and Kramer 1994), Cornell University (Billings 1996a), FASL-5 (Rudin 1996), FLSM-4 (Rudin 1993a), and—in part—the University of Massachusetts (Rudin, King and Izvorski 1995). Any inconsistencies, mistakes, or misinterpretations of their worthwhile contributions, of course, are only our fault. Billings's contribution was supported primarily by a visiting fellowship at ZAS, Berlin, during the summer of 1996; thanks especially to Tracy Hall for making this visit possible. Additional support to Billings came from a Foreign Language Area Studies fellowship awarded by Cornell University.

produce contrasting surface orders. Differences in the usage of *li*, as well as other ways of forming yes/no questions in Macedonian, further differentiate the two languages. Thus, apparent syntactic differences turn out to be due to non-syntactic factors, both phonological (prosodic) and pragmatic.

This paper has the following organization: Following a description of some of the more striking surface differences between Macedonian and Bulgarian (in $\S 2$), we present an analysis (in $\S 3$) unifying the underlying syntactic distribution of li in the two languages. Then (in $\S 4$) we discuss four prosodic peculiarities in the two languages that obscure this syntactic uniformity. Following a brief background on the prosodic systems of the two languages ($\S 4.1$), we discuss four phenomena which obscure the placement of li: In Macedonian two lexical words, under certain circumstances, share a single word stress ($\S 4.2$). In the same language clitic forms of 'be' are sometimes stressed ($\S 4.3$). In Bulgarian the negative element ne causes the following element, even an otherwise-unstressable clitic, to be stressed ($\S 4.4$). Additionally, only in Bulgarian, pronominal and auxiliary clitics are prohibited from being clause-initial ($\S 4.5$). Finally (in $\S 5$), we show how pragmatic factors limit the acceptability of li in Macedonian.

2. The Problem

In both Macedonian and Bulgarian, as well as several other Slavic languages, yes/no questions may be formed by adding li to a declarative sentence. But at first glance, the position of li in some types of Macedonian questions appears to be radically different from its position in Bulgarian. Examples (1) and (2) display precisely opposite grammaticality judgments for positive yes/no questions with complementary word orders in Macedonian and Bulgarian, while (3) and (4) show the same disparity in negative questions, and (5) and (6) in conditional questions. (Throughout the paper we have boldfaced li in examples.)

(1) a. Go vide \mathbf{li} ? him_{ACC} saw_{3.SG} Q

(Macedonian)

- b. *Vide li go?'Did (s)he see him?'
- (2) a. *Go vidja **li**?

(Bulgarian)

b. Vidja **li** go? saw_{3.SG} Q him_{ACC} 'Did (s)he see him?'

(3)	a.	Ne go vide \mathbf{li} ? (Neg him _{ACC} saw _{3.5G} Q	(lacedonian)
	b.	*Ne go li vide?	
	c.	Ne li go vide? ¹	
		'Didn't (s)he see him?'	
(4)	a.	*Ne go vidja li ?	(Bulgarian)
	b.	Ne go li vidja?	
		NEG him_{ACC} Q $saw_{3.SG}$	
	c.	*Ne li go vidja?	
		'Didn't (s)he see him?'	
(5)	a.	Bi mi dal li pari? (Mould me _{DAT} given _{M.SG} Q money	Macedonian)
	b.	*Bi li mi dal pari?	
		'Would he give me money?'	
(6)	a.	*Bi mi dal li pari?	(Bulgarian)
	b.	Bi li mi dal pari? would _{3.SG} Q me _{DAT} given _{M.SG} money 'Would he give me money?'	

Examples (1) through (6) all involve pronominal clitics; go 'him $_{ACC}$ ' or mi 'me $_{DAT}$ '. In li questions without other clitics, the strikingly divergent pattern of grammatical word orders disappears; in fact, normal word order is often identical in the two languages; compare the (a) and (b) versions of (7) through (9):

(7)	a.	Zboruvate li angliski?	(Macedonian)
		speak _{2.PL} Q English	
	b.	Govorite li angliski? speak _{2.PL} Q English	(Bulgarian)
		'Do you speak English?'	
(8)	a.	Kniga li procita Anna? book Q read _{3.SG}	(Macedonian)
	b.	Kniga li procete Anna? book Q read _{3.5G}	(Bulgarian)
		'Did Anna read a book?'	

¹ Not all Macedonian speakers accept (3c). Olga Tomić has suggested to us that those who do may be confusing it with *neli* 'isn't it?'. See also our discussion of Enlarged Stress Domain in §4.2 below. Regarding the gloss of (3c), cf. §4.2 below, especially ex. (28b).

² Cf. Avgustinova 1994 and Hauge 1976 for a full discussion of the clitic system of Bulgarian. Cf. the end of §4.4 below regarding the differing accentuation of *bi* in (5)-(6).

- (9) a. Od dve poluistini stanuva **li** celinja? (Macedonian) from two half-truths becomes_{3.5G} Q whole
 - b. Ot dve poluistini stava **li** edna cjala? (Bulgarian) from two half-truths becomes_{3 SG} Q one whole

'Do two half truths make a whole one?'

At this point one might posit special rules in each language for the placement of li relative to other clitics, or different positions for clitics in Macedonian and Bulgarian. This is unnecessary, however. We argue below that the syntactic position of li is the same in both languages, not only in cases like (7) through (9) where it appears identical, but also in (1) through (6). The conditions on word order with clitics do differ between the two languages, but the relevant conditions are prosodic, not syntactic. We adopt the analysis proposed by Rudin, King, and Izvorski 1995 and Izvorski, King, and Rudin 1996 for Bulgarian li questions, and show that it accounts for the corresponding Macedonian construction as well. In fact, the previously proposed structure is arguably clearer in Macedonian, where it is not obscured by certain prosodically controlled word-order changes. Such an analysis of li questions in Bulgarian is strengthened by this cross-linguistic comparison.

Questions with li in Bulgarian have received quite a bit of attention in the recent theoretical literature (see especially Rivero 1993, King 1993: 146–55/1994: 113–18/1995: 156–63, Izvorski 1994, Penčev 1993 and Rudin 1992; 1993a). But to the best of our knowledge, Macedonian li questions have not been analyzed in any detail, although reference grammars contain brief descriptions, Friedman's (1993: 286–87) being the most complete.

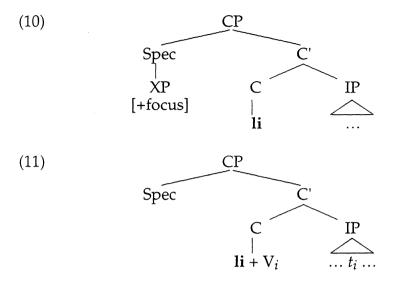
3. Syntactic Analysis

In both Macedonian and Bulgarian, li is a yes/no-interrogative particle which can also check a focus feature. As an enclitic, it is suffixed to a stressed element. We assume, with many recent analyses, that li is in C. (For arguments, see Rudin 1993a and Rivero 1993.) When C is [+focus], it checks a [+focus] feature on a fronted focus phrase in SpecCP (so-called specifier-head agreement), as in the tree in (10). Otherwise, when no focus phrase precedes li, the verb must raise and be

³ Englund 1977 includes a relatively large corpus study of yes/no questions in both Bulgarian and Macedonian. While empirically very complete, Englund's study stops short of any extensive analysis. Tomić (1996a: 511) discusses Macedonian *li* briefly. (We were not able to locate Englund 1979 or Tomić 1996b in time for this paper. From their titles, these works promise to provide such analysis.)

Much of the this analysis holds for Russian as well (King 1993: 134–44/1994: 92–110/1995: 137–53; Rudin, King and Izvorski 1995), which has both the structures in (10) and (11). (Cf. Billings 1996b, however, for a phrasal-affix analysis of Russian *li*.) In Serbo-Croatian (Wayles Browne p.c., Mihaljević 1996, Rivero 1993, Wilder and Ćavar 1994) and Czech (Toman 1996: 508–09), *li* is limited to the structure in (11); (10) is not attested in those languages. Unlike modern (Serbo-) Croatian, Mihaljević 1996 reports, Croatian Church Slavonic did allow non-verbal elements to precede *li*. The other remaining South Slavic language, Slovenian, apparently no longer uses *li* as an independent morpheme (cf., however, *SSKJ* 1975: 600, which reports archaic examples, as well as one apparently clause-initial example of *li*!).

head-incorporated into C (V-to-C movement) to check this [+focus] feature, as in the tree in (11).⁵



Structure (10) is exemplified by the sentences in (8), where kniga is the phrase under focus. Structure (11) is seen in both (7) and (9); the portion of the tree labeled "li + V" is realized in these sentences as verb **followed** by li ($zboruvate\ li$, for instance). For reasons which will become clear below, we analyze this as right-adjunction of V to li, followed by prosodic inversion (following Halpern 1992/1995). In (9), in addition to V-raising, a non-focused XP (a topic phrase, $od\ dve\ poluistina$) is fronted to a position higher than SpecCP (probably adjoined to CP). It is not in a specifier-head relation with li, and therefore has no features checked by li. V-to-C raising is necessary to support the interrogative feature of li here, showing that li is not simply a second-position clitic (or clausal affix), but rather an element in the phrase structure.

The problematic cases with clitics, seen in (1) through (6), also have the structure in (11). The structure here, however, is less obvious; *li* does not always appear suffixed to the verb. We believe, however, that the facts in all of these cases are accounted for automatically, assuming the structure in (11) and given the existence in one or the other language of four prosodic distinctions.

4. Prosodic Distinctions

Macedonian and Bulgarian each manifest certain prosodic phenomena which affect the position of *li*. Before discussing these four phenomena, however, however, we provide some background on the prosodic system of the two languages' clitics. (Numbered examples throughout this section show clitics *italicized*, lexical words <u>underlined</u>, stressed syllables in ALL-CAPS, and syllable breaks with dots.)

⁵ Slightly different proposals, such as those of Izvorski 1993, in which li is the head of FocusP rather than CP, would also be compatible with our analysis of Macedonian. We will not defend the details of the structures sketched in (10) and (11) here; the point is simply that in Macedonian, as in Bulgarian, li either incorporates a verb or checks a focus feature on an immediately preceding focused phrase in li's specifier (SpecCP).

4.1 Primer on the Languages' Clitic-Prosody Systems

The two languages' prosodic systems, especially the stress system within the word, differ significantly, requiring separate descriptions. We begin with Macedonian, showing how the antepenultimate-stress system is especially useful in elucidating the uniqueness of li as a clitic. We then follow with a point-by-point comparison with Bulgarian, showing the specific ways in which the two differ.⁶

4.1.1 The Macedonian Clitic-Prosody System

Our description here of the Macedonian system is drawn from several works, many of which deal with the theoretical problem of accounting for antepenultimate stress—a somewhat exotic pattern cross-linguistically. We refrain from entering into the somewhat lively debate about just how antepenultimate stress is to be formalized, assuming simply that such a mechanism exists. We concentrate instead on how to fit *li* into such a system, something that has been ignored to date in the literature, to our knowledge.

Macedonian's antepenultimate stress⁸ allows us to assess whether a particular clitic is part of the same prosodic word (PrWd) with the lexical word

See works listed in the preceding footnote for discussion of words with exceptional stress. These have either penultimate or final stress. As the following two representatives of each type of exception show (respectively), however, the addition of syllables (such as the plural marker or an article) to the end of the word regularize them to antepenultimate stress:

	INDEF.SG	DEF.SG	INDEF.PL	DEF.PL	Gloss
(i)	te.le.VI.zor	te.le.VI.zo.rot	te.le.VI.zo.ri	te.le.vi.ZO.ri.te	'television'
(ii)	de.le.GAT	de.le.GA.tot	de.le.GA.ti	de.le.GA.ti.te	'delegate'

That is, as shown by (i), whereas the INDEF.SG has exceptional penult stress, the addition of one syllable (in either the DEF.SG or INDEF.PL columns) keeps the stress peak on the same vowel, thus giving these forms non-exceptional antepenultimate stress. Moreover, the addition of a second syllable (shown in the DEF.PL column) actually shifts the stress peak to the next vowel, keeping this form antepenultimate in stress. Example (ii) shows this more gradually: The INDEF.SG form is two syllables out of place, while the DEF.SG and INDEF.PL columns are one syllable out of place. Finally, the DEF.PL form shows regular stress, on the antepenult. Neither of these exception types, therefore, involves stress *earlier* than the antepenult. There is one other type of exception of a different type, exhibited by the verbal-adverb (= gerund) form only, as shown in (iii) and (iv):

(iii) <u>no.SEJ.ki</u> <u>no.SEJ.ki</u> go <u>no.SEJ.ki</u> mu ja 'bringing' 'bringing him' 'bringing him it(fem)' [≈ exx. 56a-c in Franks (1987: 128)]

(iv) VI.vaj.ki go so.op.ŠTU.vaj.ki go so.op.ŠTU.vaj.ki mu 'calling' 'calling him' 'announcing' 'notifying him' [≈ exx. 57a-d in *Ibid*.]

As we show below, the addition of enclitics to non-finite verbal forms shifts stress rightward. Not so with this part of speech, which traditionally has fixed stress on the first syllable of the -ejki/-ajki suffix, as shown in (iii). Any added clitics fail to attract stress, even to the point of having preantepenultimate stress. In another style the verbal adverb has acquired regular antepenultimate

⁶ Comparative works of the Macedonian and Bulgarian nominal (type-II and-III) clitics include Tomić 1996a and Elson 1976. See these especially for discussion of their distribution. For example, Macedonian shows a tripartite proximal-neutral-distal distinction, while some Bulgarian noun classes show NOM/object distinctions.

These works on the theoretical issues surrounding Macedonian antepenultimate stress (some of which discuss the related issue of penultimate stress in Polish) include the following: Baerman (1996), Comrie (1976), Deevy (1995), Franks (1987; 1989; 1991), Garde (1968), Halle and Kenstowicz (1991), Hammond (1989), Kager (1993) and Kenstowicz (1991).

preceding it (= prosodically hosted by the preceding word) or not. That is, if the clitic is one of only two syllables following the stress, then it is within the PrWd, as shown in (12) through (15):

(12)	The yes/no-interrogative particle li : Not part of the PrWd a. SVE.kr.va. li 'Is it (a) mother-in-law?' b. *sve.KR.va. li	Type I
(13)	The definite article <i>-ta</i> (and other allomorphs): Part of PrWd a. <u>sve.KR.va.</u> ta 'the mother-in-law' b. * <u>SVE.kr.va</u> .ta	TypeII
(14)	Possessive clitics (homophonous with DAT clitics): Part of PrWd a. <u>sve.KR.va</u> .mi 'my mother-in-law' b. * <u>SVE.kr.va</u> .mi	TypeIII
/ - - \		

(15) Other clitics, incl. pronominals: Part of PrWd

a. <u>do.NE.si</u>.go 'Bring it_{ACC}!' b. <u>do.ne.SI</u>.mi.go 'Bring me_{DAT} it_{ACC}!'

Thus, (12) differs from (13) through (15) in that li is not part of the so-called trisyllabic stress window. That is, in each of (13) through (15) the addition of a monosyllabic clitic shifts the stress by one syllable rightward, to the antepenult of the lexical word plus the enclitics.

This distinction has been missed in several descriptions of Macedonian enclitic prosody in the literature. For example, Elson (1976: 276) claims, "For Macedonian, there is nothing of significance to be gained from this [comparing the accentual properties of type-II and other, undisputed enclitics/BK&R] because non-enclitics, enclitics, and the forms of the article all behave the same way with regard to the assignment of antepenult stress ..."; Elson (1976) unfortunately fails to consider *li*, just clitics of types II through IV.

(16)			Verbal Clitics	Nominal Clitics
	a.	No clitic:	DO.ne.si KNI.ga 'Bring (a) book!'	SVE.kr.va 'mother-in-law'
	b.	Type-I clitic:	do.NE.se.te.li 'Do you bring?'	SVE.kr.va.li 'Is it a mother-in-law?'
	C.	Type-II clitic:	N/A	sve.KR.va.ta 'the mother-in-law'
	d.	Type-III clitic:	N/A	sve.KR.va.mi 'my mother-in-law'
	e.	Type-IV clitic:	do.NE.si.go 'Bring it _{ACC} !'	N/A
			[(16a, c, e) from Elson	(1976); (16b, d) elicited/BK&R]

Clearly, (16b) shows that not all clitics affect stress.

speech, as in (iv). Note that in these examples additional syllables still do not shift stress rightward, again resulting in pre-antepenultimate stress, as shown in the rightmost example in (iv). The exceptions in (i) through (iv) are the only deviations from regular antepenultimate stress in Macedonian. We present only regular-antepenultimate-stress data below.

Elson (1976) does, however, consider **Bulgarian** *li*, but ignores its Macedonian counterpart.

Franks (1989) likewise assesses the accentuation of Macedonian clitics but fails to consider li altogether. Moreover, other studies discuss the accentual properties of li in Macedonian without actually mentioning on which syllable the stress is pronounced (cf. Englund 1977: 119-120; Friedman 1993: 287); Tomić (1996a) initially discusses the similarity of type I with types II and III, but then discontinues consideration of type I (where, citing Halpern 1992 extensively, she discusses junctural phonology at the word-clitic boundary).

Another distinction among these four types is whether these clitics can be initial. By this criterion, types I through III are distinct from type IV. If the verb is finite, then type-IV clitics precede the verb and can then be clause-initial in Macedonian. Those in types I through III must always be non-initial in a particular syntactic domain. ¹⁰

Type-II and -III clitics must be non-initial the **noun phrase**, as shown in examples (17) and (18). The use of type-III clitics in Macedonian is limited to kinship terms and lexified expressions, but their use is consistent.¹¹

- (17) $\underline{u.BA.vi.ot}$ \underline{MAS} [handsome+the]_{M.SG} man 'the handsome man' (NB: -ot is an allomorph of -ta 'the') [\approx Elson (1976: 279, n. 8), who uses phonetic transcription]
- (18) $\underline{MAJ.ka.mu}$ $\underline{STA.ra}$ (also: $\underline{sta.RA.ta.mu}$ $\underline{MAJ.ka}$) mother+his old_{F.SG} [old+the]_{F.SG}+his mother 'his old mother' (cf. * $\underline{MAJ.ka}$ $\underline{STA.ra.mu}$, * $\underline{STA.ra.ta}$ $\underline{MAJ.ka.mu}$) [Friedman (1993: 286)]

Any other placement of these clitics—i.e., proclitic to any word or enclitic to any word but the first one—is ungrammatical.¹²

Unlike type-II and -III clitics, *li* (in type I) must follow a word of the **clause** (specifically, "clause" here is the complementizer phrase, not counting any adjuncts to it), as in (19):

(19) <u>RAZ.bi.raš</u>.*li* <u>TI</u> <u>ma.KE.don.ski</u> understand+Q you_{NOM.SG} Macedonian 'Do you understand Macedonian?'

[Englund (1977: 93), stresses elicited/BK&R]

¹⁰ Ćavar (1996a, 1996b) and Ćavar and Wilder (1994) distinguish between Wackernagel's (1892) Law and the so-called Tobler-Mussafia Effect. Wackernagel's Law requires particular constituents to be in second position, while the Tobler-Mussafia Effect requires merely that certain elements be non-initial, based on observations about Romance. Cf. Mussafia 1898; we've been unable to locate Adolf Tobler's ca. 1880 work. Ćavar and Wilder also discuss the notion of prosodic subcategorization in Zec and Inkelas (1990: 369): "[[]]PrWd__]PrWd".

We distinguish between type II and III for reasons that aren't crucial to the main text of this paper: First, the two are not in complementary distribution, as the alternative form in (18) shows. Next, these two types have differing segmental-junctural properties, discussed in Elson 1976, Sadock 1991 and Tomić 1996a. See also the next footnote.

More specifically, type-II and -III clitics are encliticized to the first word in the nominal expression with nominal features (including adjectives and numerals). That is, an NP-initial adverb will be skipped over by the article and possessive clitics. Our description here is intended to be primarily descriptive.

(NB: in (19) ti is not a clitic, but rather the word-stressed, nominative-case pronoun.) We do not, however, assume that li is a clausal affix, merely prosodically adjoined to the clause's first word, as proposed for Russian li by Billings (1994, 1996b), for example. Instead, as shown in (10) and (11), li is in C; prosodic inversion takes place only if no element dominated by CP precedes li.

Type-IV clitics, unlike types I through III, are unique in being able to appear initially., as shown in (20a-c): ¹³

(20) a. mi go KA. \check{z} a.le (three-syllable verb) me_{DAT} it_{ACC} said_{PL} 'They said it to me.'

b. mi go DA.le (two-syllable verb) me_{DAT} it_{ACC} gave_{PL} 'They gave it to me.'

c. mi go \underline{DAL} (one-syllable verb) me_{DAT} it_{ACC} gave_{M.SG} 'He gave it to me.'

In other words, there exists no non-initiality requirement on type-IV clitics in Macedonian.

Note also that whereas post-verbal type-IV clitics affect the place of stress on the verb, *mi* and *go* in (20) do not. This stress asymmetry has been observed in Romance languages. ¹⁴ Macedonian type-IV clitics, as in some of these Romance languages as well, precede only finite verbs. ¹⁵

One (always) pre-verbal element, the negative particle *ne*, which in some cases looks like a clitic, is inherently accented in Macedonian (one of the main

(i) <u>ne</u> *ME* <u>da.vaj</u> <u>MAJ.ko</u> 'Don't give me (in marriage), mother!' NEG me(acc) give(imperative) mother(vocative) [Lunt (1952: 22, fn. 1)]

(ii) Ne go gledaj! 'Don't look at him!'

NEG him(acc) look(imperative) [= ex. 28 in Alexander (1994: 10)]

Lunt implies that this order is marked (but shows ESD diacritics!). Alexander marks (ii) with a question mark and shows the preferred order, *Ne gledaj go!* [= her ex. 29], without marking stress in either example. These forms are significant from a prosodic standpoint, as Alexander points out, because those speakers who use(d) them appear to merely have a non-initiality requirement on clitics of non-finite verb forms—Wackernagel's (1892) Law in Alexander's terminology—because non-finite verbal forms are frequently the first lexically accented word of the clause. Those who don't use forms like (i) and (ii) appear to order their verbal clitics as in Romance, discussed in the preceding fn.

 $[\]frac{13}{14}$ Examples (1a) and (5a) also show this point. See fn. 2 regarding the clitichood of bi.

Cf. Peperkamp 1995 for a summary of the literature on the dual-position clitics in Romance. Specifically, Macedonian appears to be very similar to certain Lucanian dialects of Italian. As Peperkamp (1995: 122, citing Lüdtke 1979: 31 in part) reports, Lucanian words show stress on one of the last three syllables and the addition of enclitics and suffixes both regularize the stress of the suffixed and/or encliticized stem to penultimate stress. (Macedonian, as a preceding footnote mentions, has words with stresses in the same final-trisyllabic window, with the addition of suffixes and non-li enclitics regularizing the stress, in the case of Macedonian, to antepenultimate position.) Garde (1968: 32) and Kenstowicz (1991) also point out this Macedonian-Italian similarity.

We have uncovered the following two exceptions, of the same kind, to this description:

assertions of Garde 1968). ¹⁶ As (21) shows, whereas the type-IV clitics *mu* and *go* are inherently unaccented, *ne* is accented.

(21) NE mi go DA.le NEG me_{DAT} it_{ACC} gave_{PL} 'they didn't give it to me'

Example (21) does not show clearly which lexical (underlined) word hosts the clitics. Two plausible prosodic organizations are shown in (22a-b):

There are a few reasons to favor the bracketing in (22a): First of all, these clitics are syntactically ordered preceding the finite verb; the null hypothesis is that the prosody matches the syntax. Another reason to favor (22a) is the behavior of such clitics when the verb is clause-initial, as in (20a-c). If these clitics are hosted by any word prosodically, it must be by the verb, perhaps by so-called stray adjunction.

Another argument in favor of (22b) is shown in (23), in which there is no overt verb: ¹⁷

In the case of (23), the only lexical word is the *wh*-interrogative *kamo*. Once there are three syllables worth of clitics, then the stress shifts off of the lexical word and on to the third (monosyllabic) clitic from the end. This suggests that the clitics in (21) are prosodically hosted by the **preceding** lexical word *ne*.

While the data in (20) and (23) suggest that clitics more naturally adjoin to a preceding prosodic host, these data are inconclusive. Example (23) also looks like stray adjunction, as in (20a-c). That is, when clitics have a lexical-word neighbor on only one side, then they will be hosted prosodically by that word. What we really need is an example like (21), but with at least **three** syllables worth of verbal clitics between the two lexical words, like (24): ¹⁸

Anderson (1996: 188–89) offers a different account for the unique properties of *ne* in Macedonian, considering it a clitic which is positioned differently from the other clitics.

Ex. 19d in Franks (1989: 561) glosses this same example as 'Where did that thing of yours get to on me?' (Elson 1993, incidentally, was apparently written unaware of Franks 1989.) Both Elson and Franks apparently got this example from Koneski [1957: 123/]1967: 163, which doesn't have glosses, because it's written in Macedonian. This example, furthermore, does not appear in Lunt or Garde (which would have glosses, at least in French in the latter case). We have added the word-for-word glosses and part-of-speech labels, which shows the type-IV and -III interpretations of *ti*, respectively.

Elson (1993: 152–53) reports the following pair of examples; see §4.2 regarding (ii).

⁽i) $\underline{\text{ne}}$ SME mu go $\underline{ZE.le}$ (ii) $\underline{\text{ne}}$ sme mu GO $\underline{\text{ze.le}}$ NEG $are_{1.PL}$ him_{DAT} it_{ACC} $took_{PL}$ We didn't take it from him.'

(24) $\underline{\text{NE}}$ sme mu go $\underline{\text{DA.le}}$ NEG are_{1.PL} him_{DAT} it_{ACC} given_{PL} 'We didn't give it to him.'

This example shows that three intervening clitics fail to draw stress off of *ne*, which indicates that the bracketing in (22a) is correct, with the proviso that only prosodic enclitics affect stress.

To summarize this subsection, Macedonian has clitics with four different kinds of properties: li, in type I, is always prosodically enclitic and never initial, but, unlike all other clitics, fails to affect stress. Like li, the definite article and possessive clitics (types II and III) are non-initial (within a noun phrase) and prosodically enclitic, but unlike li, always affect stress. Finally, the type-IV clitics precede finite verbs and follow non-finite ones and are always hosted by the verb.

4.1.2 The Bulgarian Clitic-Prosody System

In this subsection we show the aspects of the Bulgarian clitic system that differ form the corresponding Macedonian ones described in the preceding subsection. In more than one aspect we simply note that the Bulgarian system is different, without giving as detailed a description as we did of the Macedonian system. We defer instead to the very detailed account in Hauge 1976. ¹⁹

In Bulgarian, stress in largely unpredictable. That is, the location of stress (or its pattern of assignment) is lexically encoded. Stress may fall on any syllable of a word, and is usually unaffected by the addition of clitics. (A partial exception is the definite article, which in some stress paradigms takes stress on itself:

(Stress notation modified; word-for-word glosses added; sentential gloss unchanged.) Elson says clearly that the three clitics are hosted prosodically by *ne*. Our informants rejected example (i) , preferring instead the order in (50a). Elson (1993: 157, n. 1) glosses and attributes this and other data in the article with the following acknowledgment: "Items cited for illustrative purposes or their models, are from Lunt 1952: 21–25, Koneski 1967: 139–210, or Garde 1968." Indeed, example (ii)—the acceptability of which we don't dispute—appears in Garde 1968: 31 and Lunt 1952: 23. The source of Elson's assumption, that the clitics are hosted by *ne*, seems to be Garde (1968: 36):

(iii) — / NE — sme mu go / ZELE —
$$\rightarrow$$
 — / NE sme mu gó zele — [sic.]

Garde's abstract notation implies that whereas on the left side of the arrow [= some sort of underlying representation prior to application of ESD] the (all-caps) lexically accented stems *ne* and *zele*, with the clitics hosted by *ne*, the right side of the arrow [= the combined underlying and surface representations of the ESD form] has only *ne* with accent and everything hosted prosodically by it, with actual stress indicated by the acute accent on *gó*. Elson (1993: 152–53), while arguing against much of Garde's proposal, appears to espouse the left-hand side of (iii), assuming the right side of (iii) and the form in (i) to be attested. Our data on *li*-insertion below (in §4.2), specifically in (35a), show additional evidence that the clitics are hosted prosodically by the verb. Still, Macedonian—and Balkan in general—being a very diverse dialectal situation, we leave open the possibility that (iii) is attested for some speaker somewhere.

Anderson (1996: 188) incorrectly characterizes the ordering of type-IV languages in Bulgarian (and Macedonian) "follow gerunds, infinitives and imperatives." (i) neither Bulgarian nor Macedonian has an infinitive; (ii) as both Hauge (1976: 5) and Alexander (1994: 9) point out, such clitics precede non-initial imperative verbs as well in Bulgarian.

```
    (25) a. <u>GRAD</u> 'city'
    (26) a. <u>PA.met</u> 'memory'
    (27) a. <u>SE.dem</u> 'seven'
    b. <u>gra.dÂT</u> 'the city'
    b. <u>pa.met</u>.TA 'the memory'
    b. <u>se.dem</u>.TE 'the seven'
    [≈ exx. in Elson (1976: 276)]
```

The addition of other clitics does not shift stress.

Bulgarian has all four types described in the preceding subsection. Type I (li) is used far more frequently than in Macedonian, as we discuss (in §5) below. Type II has roughly the same distribution, with the stress peculiarity shown in (25) through (27), while type III is far more frequently used in Bulgarian than in Macedonian. Types I through III have the same prosodic properties in both languages—all required to be non-initial (li, in the non-adjoined-to CP; types II and III, in the NP).²⁰

Two major distinctions of Bulgarian we discuss below in separate subsections: Type-IV clitics in Bulgarian are, like li is in both languages, required to be non-initial in the clause (§4.4) and the negative element ne idiosyncratically stresses the following constituent (§4.5), even a clitic. Both of these phenomena drastically obscure the syntax and relatively simple prosody of li in Bulgarian. Before Proceeding to those two phenomena, however, we return to two distinctions in Macedonian.

4.2 Enlarged Stress Domain in Macedonian

We show in above (in §4.1.1) that *ne* in Macedonian is inherently accented (unlike *ne* in Bulgarian; cf. §4.5 below). It is also possible, however, for *ne* and the following verb to share a single PrWd stress; this possibility of stressing both lexical words as one PrWd accounts for the acceptability of both (3a) and (3c) above, and results in contrasts like (28):

```
(28) a. NE li SA.kaš da O.diš? (Macedonian)

NEG Q want<sub>2.SG</sub> to go<sub>2.SG</sub>

'Don't you really want to go?' (or 'Do you really not want ... ?')

b. NE sa.kaš li da O.diš? (Macedonian)

NEG want<sub>2.SG</sub> Q to go<sub>2.SG</sub>

'Don't you want to go?'
```

In (28a) the speaker is sure of a confirmative answer; the *ne* is being questioned. In (28b), however, the speaker is unsure whether you want to go or not; *sakaš* is being questioned. In Bulgarian, only *Ne iskaš li da otideš?*—the order corresponding to (28b)—is possible.

²⁰ Cf. other works, however, which discuss interesting interactions of these non-initiality domains. Halpern (1992/1995: 227–31), devotes an appendix to overlapping domains, citing Ewen (1979) and Hauge (1976) with Bulgarian examples. Halpern discusses what happens when the leading edge of a clause coincides with an NP's leading edge. He claims, contrary to our proposals here, that type-II and -III clitics must be peninitial in that NP, while type-I clitics must be clause-peninitial. Halpern also discusses briefly the apparent ability of possessive (type-III) clitics to appear out of their NP, within the cluster of type-IV clitics. Note also that within the nominal expression the article precedes the possessive clitics, as the alternative form in (18) shows.

This possibility of stressing two lexically accented words as a single PrWd has been called "enlarged stress domain" (Franks 1987, also referred to as "collocational stress" in Elson 1993 and "accentual units" in Alexander 1994), as shown in (29) through (32). In the (a) examples each (underlined) lexical word receives the predictable stress (antepenult if at least trisyllabic; otherwise, initial). In the (b) examples, however, the entire two-word domain receives a single PrWd stress, on the antepenult. Aside from adjective + noun, shown in (29), the other word pairs reported in the literature are preposition + noun, *wh*-interrogative + verb, and negation + verb, shown in (30) through (32), respectively. These four environments are elaborated as well in Elson (1993) and Franks (1989).²¹

Without ESD

- (29) a. <u>LE.va.ta NO.ga</u> left+the foot 'the left foot'
- (30) a. <u>O.ko.lu TR.lo</u> near sheep-pen
- (31) a. \underline{KOJ} $\underline{RE.\check{c}e}$ who_{NOM} said_{3.SG}
- (32) a. NE mi go DA.le $NEG me_{DAT} it_{ACC} gave_{PL}$ 'they didn't give me it'

With ESD

- b. <u>le.va</u>.*TA* <u>no.ga</u> left+the foot 'the left foot'
- b. <u>o.ko.LU</u> <u>tr.lo</u> near sheep-pen
- b. \underline{KOI} re.če who_{NOM} said_{3.SG}
- b. $\underline{\text{ne}}$ mi GO $\underline{\text{da.le}}$ NEG me_{DAT} it_{ACC} gave_{PL} 'they didn't give me it'

Additionally, as (29) and (32) show, clitics can appear between the lexical words. In (29) the definite article -ta is enclitic to the first word of the noun phrase; in (32) the clitics mi and go are syntactically ordered before the finite verb. The environment in (31) also allows medial clitics.²²

As we show above (at the end of $\S4.1.1$), the verb is the **prosodic** host of the clitics mu and go in (32a). We repeat example (24) as (33a), adding its ESD

Alexander (and we) cannot explain this ESD form, a fixed expression, in syntactic terms.

(i) $\underline{\text{ne}}$ sum mu GO $\underline{\text{zel}}$ 'I didn't take it from him.' [Lunt (1952: 23)] . NEG am him it took (cf. * $\underline{\text{ne}}$ sum MU go $\underline{\text{zel}}$)

(ii) <u>ne</u> BI <u>dal</u> or <u>NE</u> bi <u>dal</u> '(He) should not have given'

NEG should gave NEG should gave [= ex. 18a in Franks (1989: 559)]

The second peculiarity is an exception to Franks's monosyllabic-head effect just in case the entire ESDomain consists of exactly three syllables, as shown in (ii). Examples (i) and (ii) also have one non-ESD variant each: NE sum mu go ZEL, NE bi DAL. Kepeski and Pogačnik (1968) provide an opportune pair, similar to (ii): doBAR den 'good day', DObra nok' 'good night'.

The ESD forms in (9b) and (10b) are now quite marked in Contemporary Standard Macedonian, considered as either archaic or dialectal. We cite them just once, nonetheless to report the extent of ESD in the language. Alexander (1994: 11) also lists the following ESD example. Cf. Lunt (1952: 24–25) for further discussion.

⁽i) <u>dobre</u> TE <u>najdov</u> well you(acc) found(1.sg) 'Welcome!' (literally: 'I found you well')

Two peculiarities of ESD occur when the latter lexical stem is monosyllabic: The first—which Franks (1989) calls the "monosyllabic-head effect"—prevents the stress from preceding the beginning of the second stem by more than one syllable, as shown in (i):

counterpart in (33b). The bracketings in (34a-b) represent the prosodic organizations of (32a-b). We avoid the debate in the ESD literature about whether one of the two lexical words becomes a clitic.

Without ESD

With ESD

(33) a. NE sme mu go DA.le NEG are_{1.PL} him_{DAT} it_{ACC} gave_{PL} 'We didn't give it to him.' b. ne sme mu GO da.le

(34) a. [<u>NE</u>][sme mu go <u>DA.le</u>]

b. [ne sme mu GO da.le]

In addition to the three-syllables-worth-of-clitics test in (24), the bracketing in (34a) can be tested by turning the clause into a yes/no question using li, as shown in (35) and (36): ²³

- (35) a. NE li sme mu go DA.le b NEG Q are_{1.PL} him_{DAT} it_{ACC} gave_{PL}
 - b. *<u>ne</u> sme mu GO **li** <u>da.le</u>
- (36) b. *NE sme mu go li DA.le
 'Didn't we give it to him?'
- b. ne sme mu GO da.le li

The grammatical positions of *li*, depending on whether there is ESD, are in (35a) and (36b). There are three plausible hypotheses about the placement of *li* worth considering—assuming a prosodic-inversion account such as Halpern's (1992/1995)—shown in (37a-c):

- (37) a. Hypothesis A: *li* follows first PrWd **stress** [Wackernagel 1892].
 - b. Hypothesis B: *li* follows first PrWd **domain**, assuming the structure in (22b) above: [NE mi go] [DA.le].
 - c. Hypothesis C: *li* follows first PrWd **domain**, assuming the structure in (22a) above: [NE] [*mi go* DA.le].

Hypothesis A predicts that li will appear at the first available syntactic boundary following the first word stress; it correctly predicts the form in (35a), but incorrectly predicts (35b). Hypothesis B correctly predicts the form in (36b), but incorrectly predicts (36a). Hypothesis C is the only one to correctly predict both of the attested forms: (35a) and (36b). The clitic li, therefore, corroborates the correctness of our analysis above in (22a), as shown first by the datum in (33a).

90 DA.la li? (Macedonian, with ESD) (38) a. <u>ne</u> sum **NEG** $you_{DAT.SG}$ it_{ACC} given_{F.SG} Q am_{1SG} **li** sum DA.la? (Macedonian, without ESD) b. ne 90 SÂM li ti DA.la? c. ne (Bulgarian) 90 'Haven't I given it to you?'

 $^{^{\}rm 23}$ We discuss $\it neli$ vs $\it ne\, li$ in fn. 1 above. Cf. also the following ungrammatical forms:

⁽i) NE mi go DA.le li (ii) *ne li mi GO dale

In (i) *li* follows two stresses, while in (ii) *li* follows no stresses—both illicit inversion.

The only order allowed in Bulgarian is shown in (38c).

A similar example, from a folk song, is (39), in which the speaker believes the addressee **should** be ashamed.

(39) <u>NE li ti TE.be SRA.mo.ta?</u> (Macedonian)
NEG Q you_{DAT.SG} you_{DAT.SG} ashamed
'Aren't you ashamed?' [from *Acana mlado nevesto*]

The corresponding Bulgarian question, *Ne te li e sram?*, must place *li* after *te*, not right after *ne* (for reasons we present in §4.5 below). This example does not have to do with ESD, since no such option exists here (i.e., no verb after *ne* + clitics).

In (31b) above we show that ESD can take place between an initial wh-interrogative stem and the verb. As (the title of) Rudin 1992 shows for Bulgarian, li can appear in wh-interrogative clauses as well, lending an emphatic 'on earth' or 'the hell' meaning to the question. This is true as well for Macedonian. It should thus be possible for li to appear in a clause like (31b). Indeed, we've found only the following example, in (40a):

(40) a. Koj li ke bide toj? — ... (Macedonian) who Q MOD is_{3.SG} that_{M.SG}/he
b. KOJ li ke BI.de TOJ
'Who on earth would that/he be?'

In light of the widely accepted model that wh-interrogatives like koj 'who' are in SpecCP and li is in C, which we adopt here, then it should follow that there cannot be an ESD counterpart of this example (i.e., *Koj ke bide li toj?), since li already follows some stressed constituent, koj, at S-structure. Whereas ne + verb constructions (with intervening clitics) have both ESD and non-ESD variants (with li ordered accordingly), wh + verb constructions are attested in only this example in Englund's li corpus of Mac, which our informants pronounced only without ESD.²⁴

As Elson (1993: 158, n. 4) points out, while all other ESD pairs in (29) through (32) constitute a syntactic constituent of some sort, wh + verb seems not to be a constituent. That problem aside, there is a functional problem: If wh + verb ESD means **de**-emphasis of some sort on the wh element, and wh + li represents emphasis on the wh element, then imposing the non-ESD onto such a clause deemphasizes the wh element, contrary to li's purpose in a wh clause. The hypothetical ESD version should be bizarre because question words are always focused and the remainder of the wh question clause is presupposed (i.e., $What \ did \ you \ see$? presupposes that you saw something; Who is he? presupposes that he is someone). This is the focus-presupposition structure characteristic of the XP-li construction in (10); it would be pragmatically weird at best to have a wh question with V-li structure in (11)!

To summarize this subsection, we have shown that a stress domain in Macedonian is enlarged, allowing certain pairs of lexically accented words, along with any clitics between them, to be stressed as one PrWd. The addition of *li* to

 $^{^{24}}$ One informant rejected this example, but appears to reject all $\it wh~+\it li$ questions.

ESD and non-ESD environments disproves one aspect of the now famous Wackernagel's Law (Wackernagel 1892), which describes peninitial clisis as following the first stress **peak**. In fact, as these data show, such clitics follow the first stress **domain**, passing up the chance to follow immediately after the stress. We also discuss the interaction of li in wh questions with ESD. This unique phenomenon in Macedonian also explains one constituent order difference with Bulgarian.

4.3 Optional Auxiliary-Accent in Macedonian

Clitic forms of 'be' in Macedonian appear to be accented, at least in some contexts, while the corresponding forms in Bulgarian are obligatorily stressed after *ne* (cf. §4.5 below) and obligatorily unstressed elsewhere.

Englund (1977: 111) reports example (41a), quoting Živko Čingo's *Paskvelija*, without stress in the standard orthography, but with ti alone not in italics. We have merely transliterated the example:

(Note again, as in (19) above, that ti here is the lexically accented NOM-case form, homonymous with the unaccented DAT clitic.) Responses varied when we elicited stress from informants: Those who did accept (41a) without comment invariably stressed it as shown in (41b). This is predicted if one assumes that li must follow the first stressed word in structures **without** a focused XP in SpecCP, as shown in the tree in (10).²⁵

This example is especially interesting in light of the head-movement account we adopt (in §3) above. It shows what happens when the tensed verb is a clitic auxiliary, which moves to C (in order to check li's focus feature) and there is no other verbal stem to host li. For those speakers who accept (41), the best means of keeping li from being initial seems to be a last-resort strategy of stressing the clitic auxiliary.²⁶

Those who rejected this order suggested (42a-b) instead:

Using *li's* stressed allomorph *dali*, in (41c), is discussed (in §5) below.

 $^{^{25}}$ The stress on ti in (41b) is not very pronounced phonetically. Some authors (cf. Hauge 1976) list NOM-case pronouns as separate PrWds but don't mark stress on them. Lunt (1952) doesn't however indicate word stress on such pronouns.

²⁶ Examples similar to the preceding Macedonian ones are also attested in Bulgarian. Hauge (1976: 2-3) lists examples in which clitic forms of 'be' are stressed, accounting for most of these by either ellipsis or displacement of the would-be host of the clitic, leaving it stranded in clause-final position. Hauge has no explanation, however, for one example with prosodic and syntactic structure and word order seemingly identical to that of (41).

It is not clear, however, what the structure of (41d) is. Two possible S-structures—prior to any prosodic inversion—are shown in (43a-c):

- (43) a. $[CP[SpecCP \underline{ti}][C li][IP[I si]...[PP na mla.di.na.ta]PP]IP]CP$
 - b. [CP[Cli]][IP[SpecIP ti]][Isi]...[PP na mla.di.na.ta][PP][IP][CP]

In (43a) ti is focused (i.e., moved to SpecCP), as diagrammed arboreally in (10) above. This structure does not require prosodic inversion and merely stresses the correct syllable in each PrWd, as shown in (42b). The structure in (43b) shows ti in SpecIP. This structure merely requires li to invert past the first stressed word, ti, in order to keep li prosodically enclitic to some word in this clause. The choice between the structures in (43a-b) might be distinguished phonetically by differing accentuation on ti. We have not conducted such tests and merely present both.

Interestingly, Macedonian has lexically accented 'be' forms, as does Bulgarian. These prosodically heavy forms, formed from the stem /bid-/ (/bâd-/ in Bulgarian). These accented stems, as is apparent from Kramer 1993, are used only as auxiliaries, never as copulas. Korubin (1974: 247–48) also points this distinction, supplying additional examples corresponding to (41) and (42). Thus, these accented stems are not available as ways of making these clitics accented.

Before leaving the issue of copula-stressing, we have one example which appears to involve the interaction of copula-stressing and ESD (cf. in §4.2 above):

```
(44) a. ..., ama ne e li toa otvoren [...]? but NEG is<sub>3.SG</sub> Q it<sub>F.SG</sub>/she<sub>SG</sub> closed '..., but isn't it/she closed [...]?'
b. ama NE e li TO.a OT.vo.ren
```

All our informants accepted (44a), a transliteration of Englund's example (1977: 115, quoting Taško Georgievski's *Zmiski vetar*), each of them supplying the stresses in (44b), which is, at first glance, problematic for the syntactic and prosodic-inversion accounts we adopt here. The problem is that if *ne* is inherently accented, then it should bear stress, with *e* encliticizing to it prosodically. Generally speaking, when *li* prosodically inverts, it appears between the first stressed word and any of its enclitics. This pattern would result in the order *ama ne li e ... We suggest instead that ESD, between ne and the now-accented e, may be involved. If so, then li inverts to the attested place, after ne.

The following structure might also be the S-structure of (42b): (i) [CP [C li] [IP [I si] ... \underline{ti} ... [PP na $\underline{mla.di.na}.ta$] PP]IP]CP

In such a structure li would have to invert prosodically past the first stressed word, ti, resulting in the form *si TI li ČLEN na mla.DI.na.ta; this leaves si in clause initial position. In our discussion of Macedonian type-IV clitics above in §4.1.1—cf. especially exx. (20), (21), (23), and surrounding discussion and footnotes—we show that type-IV clitics can be initial only if proclitic to a (lexically accented) finite verb form. When no such finite-verb host appears in the clause, as in (23), then these clitics appear to be enclitic to the first PrWd. If this is the case, then the structure in (i) is still plausible; si inverts as well.

Hauge (1976: 16, 36–44) makes explicit the auxiliary/copula distinction in Bulgarian. Korubin (1974: 246) also shows that auxiliary clitics lack third-person forms.

To summarize this subsection, we have shown that clitic copulas can be accented under conditions that are still not clear to us. We also show one potential interaction of this phenomenon with ESD. In future work we hope to pursue the conditions under which such forms are possible, as well as how (if at all) their Bulgarian counterparts differ.

4.4 Inversion of Verbal Clitics in Bulgarian

Macedonian type-IV clitics precede the verb only when it finite (and are both prosodically and syntactically proclitic to it). Bulgarian type-IV clitics are also essentially syntactically proclitic to the verb. Unlike those of Macedonian, however, clitic pronominals and auxiliaries (not all type-IV clitics!) are restricted from being clause initial. In addition to li in both languages, type-IV clitics in Bulgarian are prohibited from being initial.³⁰ If no accented word appears in front of these clitics, prosodic inversion takes place (following Halpern 1992/1995). The effect of this process is clearly seen in the difference between the non-li examples in (45) through (48), as well as in the li questions in (1) and (2) above.

(45)	me <u>BO.li</u>	<u>U.vo</u> .to	(Macedonian)
	me _{ACC} hurts _{3.SG}	ear+the	
	'My ear hurts.'		

- (46) a. <u>bo.LI</u> me <u>u.XO</u>.to (Bulgarian) hurts_{3.SG} me_{ACC} ear+the 'My ear hurts.'
 - b. <u>u.XO</u>.to me <u>bo.LI</u> (Bulgarian) ear+the me_{ACC} hurts_{3.SG} 'My **ear** hurts.'
 - c. *me <u>bo.LI u.XO</u>.to (Bulgarian)
- (47) ti ja $\underline{DA.dov}$ (Macedonian) $you_{DAT.SG}$ it_{ACC} gave_{1.SG} 'I gave it to you.'
- (48) a. $\underline{DA.dox}$ ti ja (Bulgarian) gave_{1.SG} you_{DAT.SG} it_{ACC} 'I gave it to you.'
 - b. AZ ti ja DA.dox (Bulgarian) $I_{NOM} you_{DAT.SG} it_{ACC} gave_{1.SG}$ 'I gave it to you.'
 - c. *ti ja <u>DA.dox</u> (Bulgarian)

Unlike li (in both languages), Bulgarian type-IV clitics must be non-initial in a slightly different domain. As we show in §4.1 above, li inverts if no other element dominated by CP appears in front of li. That is, adjuncts to CP do not count. Bulgarian type-IV clitics, however, can make use of adjuncts to CP, coordinating conjunctions and other material not dominated by the CP node. As Hauge (1976: 5) points out, however, type-IV clitics in Bulgarian are also prohibited from following a clause-internal pause. Thus, it seems that the crucial non-initiality domain for these is some sort of phonological phrase.

Clitic inversion is relevant to the syntax of *li* questions because such clitics raise to C along with the verb. We assume (following Rudin 1996) that clitics are functional heads which incorporate V. In Macedonian, *li* is straightforwardly suffixed to the verbal complex (i.e., the complex prosodic word consisting of the verb and its preceding type-IV clitics), as in (49); see also (1), (3) and (5) above.

b. [ste go GLE.da.le] li O.voj FILM (Macedonian) $are_{2.PL} it_{ACC} seen_{PL} Q this_{M.SG} film$ 'Have you seen this film?'

In (49b), for example, the verbal complex is *ste go gledale*. Right adjunction of the verb to li results in the S-structure string li + ste go gledale in C; prosodic inversion then produces the surface order.

In Bulgarian, exactly the same process occurs if a topic phrase (or any other material not belonging to a preceding clause) precedes C. In (50), the verb complex *ste go gledali* can appear pre-verbally

(50) VI.e [
$$ste$$
 go GLE.da.li] TO.zi FILM (Bulgarian) you_{NOM.PL} are_{2.PL} it_{ACC} seen_{PL} this_{M.SG} film 'You seen this film.'

However, when no topic or other material precedes C, rendering the complex verb sentence-initial, then clitic inversion is required to provide a prosodic host for the clitics. In (51b)—the Bulgarian counterpart of (49a)—the verbal complex ste go gledali is adjoined to C, resulting in the string li + ste go gledali. Simple prosodic inversion of li here would produce a sentence with initial clitics, which is not possible in Bulgarian. Clitic inversion is required, resulting in the surface order gledali ste go. In both Bulgarian and Macedonian, when a complex V raises, li cliticizes to the first stressed element of the V; that is, prosodic inversion of li is to the end of the first stressed phonological word to its right. In the examples in (49) the only stressed element is the verb, so li follows it, in both languages; the result for Bulgarian is (51):

'Have you seen it?'

In constructions which involve a stressed auxiliary instead of clitic forms of 'be', *li* follows the auxiliary. These auxiliaries include *ima*- 'have' in Macedonian, the past tense forms of the 'be' auxiliary in Bulgarian, and the negative-future particle *nema*-/*njama*- 'won't':

As we mention at the beginning of this subsection, not all type-IV clitics in Bulgarian are restricted from being clause-initial. In both languages the future particle $ke/\delta te$ is an unstressable clitic, but in Bulgarian it differs from the other type-IV clitics in being able to begin a sentence; this then shields the other (non-li) clitics from being clause-initial:

(54) a.
$$ke$$
 go $ZA.vr.šat$ li (Macedonian) MOD it_{ACC} finish_{3.PL} Q b. te go $SVÂR.šat$ li (Bulgarian) MOD it_{ACC} finish_{3.PL} Q 'Will they finish it?'

On the other hand, the conditional stem $/\underline{bi}$ -/ is lexically stressed in Bulgarian but not in Macedonian, leading to the contrast shown above in (5) and (6). Whereas bi appears clause-initially in both (5a) and (6b), this element is distinct in the two languages. The irrealis element in Bulgarian is conjugated, showing person- and number-agreement and each form is stressed; the form bi in (6b) is the 3.SG (homophonous with the 2.SG) form (cf. Hauge 1976: 36). In Macedonian bi is an invariant clitic, behaving like any other type-IV clitic in that language.

To summarize this subsection, we have shown that Bulgarian, unlike Macedonian, has a non-initiality requirement on its clitic pronominals and auxiliaries (a subgroup of type IV). The modal *ke* is not, however, subject to non-initiality, and the conditional stem /bi-/ is not a clitic at all. These differences cause many clitic sequences to differ between the languages.

4.5 Negative Stress Shift

The second prosodic peculiarity of Bulgarian we will discuss has to do with the unique properties of the negative particle *ne*. In Bulgarian *ne* carries an inherent stress, which is realized on the **following** syntactic constituent, even if that following constituent happens to be a normally-unstressable clitic (see Halpern 1992/1995, Scatton 1984). The relevance of this fact for *li* questions should be clear at this point, given the role of stress in determining the position of *li*. (Izvorski 1994; Izvorski, King, and Rudin 1996; King 1993/1994/1995; Rudin 1990–91, 1992, 1993a, 1993b, 1996; Rudin, King, and Izvorski 1995; and Rudin and Kramer 1994

have recognized this.) Since li cliticizes to the first stressed element of the verb word, li in Bulgarian is automatically placed after the constituent immediately following ne. In Macedonian, however, ne does not induce stress on the following word, and thus doesn't influence the placement of li. So, when li cliticizes to the first stressed element of the verb word (as in Bulgarian), the result is different: It ends up encliticized to the verb (in ESDs) or to ne, since the clitics are not stressed.

We have already seen examples of this in (3) and (4) above; another set is given in (55):

(55) a.
$$\underline{\text{ne}}$$
 mu GO $\underline{\text{da.de}}$ li (Macedonian, with ESD) NEG \lim_{DAT} \inf_{ACC} gave_{3.SG} Q

b. $\underline{\text{ne}}$ MU $\underline{\text{li}}$ go $\underline{\text{da.DE}}$ (Bulgarian) NEG \lim_{DAT} Q \inf_{ACC} gave_{3.SG}

'Didn't (s)he give it to him?'

In (55b) mu is stressed in Bulgarian, as is go in (4b), because of the preceding ne, and li therefore must follow them. This does not happen in Macedonian.³¹ If no type-IV clitic is present, then the Bulgarian order is ne + verb + li, which appears like Macedonian ESD, but is instead the result of ne not having stress and (vacuously) stressing the following accented verb stem.

Thus, constructions with li and ne are complicated by two independent phenomena, causing Bulgarian and Macedonian to diverge markedly. ESD causes Macedonian li to appear after the second inherently accented word, while negative stress shift causes Bulgarian li to follow up to one clitic after ne.

To summarize this section, then:

- The two languages differ significantly in their word prosody. Macedonian has antepenultimate word stress, while Bulgarian has lexically-encoded stress location (§4.1).
- Macedonian allows certain two-word combinations, causing *li* to follow what appears to be two words in that language (§4.2).
- Macedonian also apparently allows certain copulas to be stressed, likewise affecting the position of *li* (§4.3).
- Bulgarian requires clitic auxiliaries and pronominals to be non-initial in the clause. Additionally, the conditional head in Bulgarian, unlike in Macedonian, is not a clitic. This does not affect the placement of *li* as such, but greatly confuses comparisons with Macedonian (§4.4).
- Bulgarian *ne*, the Neg head, has a special property of stressing the following element. Thus, when *ne* follows *li* at the beginning of the S-structure order, *li* must invert prosodically past not just *ne* but the following, stressed element as well (§4.5).

These prosodic differences cause Macedonian and Bulgarian to appear very divergent in their placement of *li*. In fact, the syntactic structure of the two languages, as shown in the preceding section, is quite uniform.

³¹ Englund (1977: 112), quoting Jovan Boškovski's *Izbor*, does list the following example: "Ne ti li se čini deka pticana kako da ni potskažuva za zaludnosta na ova naše metkanje po ulicive?" All of our informants rejected this order, putting *li* instead after *ne* (non-ESD) or after the verb (ESD).

5. Usage

Macedonian yes/no questions differ from those of Bulgarian in another way as well, this time not prosodic, but pragmatic. Unlike in Bulgarian, Macedonian li is optional; it alternates apparently rather freely with \emptyset as well as with the non-clitic yes/no-interrogative complementizer dali. Englund (1977) reports that in her literary corpus 60 percent of all yes/no questions in Bulgarian contained li, the remainder mostly formed with other question words, such as nali 'isn't it'. In Macedonian, only 30 percent of yes/no questions had li, and 44 percent (almost half!) had no question word; cf. Friedman 1993: 286–287 and Kramer 1986: 130–50.

This difference in the usage of li is borne out by a survey of questions in Kramer 1985, a phrase book of Macedonian. Out of 101 questions that would take li in Bulgarian, 52 have li, 29 have dali, and 20 have no overt question word. The three types of questions are apparently synonymous; when visiting the auto mechanic, for example, the tourist is advised to ask the three questions in (56):

(56) a. Imate maslo za avtomobil?
 have_{2.PL} oil for automobile
 'Do you have oil for cars?'
b. Imate li auspuh?
 have_{2.PL} Q muffler
 'Do you have a muffler?'
c. Dali imate svekicki?
 Q have spark-plugs
(Macedonian)
dali (Macedonian)

'Do you have spark plugs?'

Only the second of these corresponds to a normal question in Bulgarian; a question formed with intonation alone, like (56a) is marginal if possible at all in Bulgarian, while *dali* in a main clause signals a rhetorical question in Bulgarian: 'I wonder if you might have spark plugs.'

The optionality of li does not bear directly on the syntax of li questions, but does affect judgments. Speakers may reject an example with li not because the syntax is wrong, but because they prefer dali or \emptyset . The reasons for this may be pragmatic or simply personal preference. One speaker we consulted, an 18-year-old woman, showed an especially strong preference for \emptyset questions, accepting questions with li only reluctantly. However, when forced to use li she had very clear intuitions about where in the sentence it could and could not go. Another speaker frequently commented that an intonation question would be more usual in everyday speech than one with li. This preference certainly leads to a difference in usage and the frequency of li questions in the Bulgarian and Macedonian, but it does not seem to be connected to any difference in the syntax of li itself.

6. Conclusion

Our analysis captures both the essential similarity between the grammars of two closely related languages, and the striking differences between them. As one

 $^{^{\}rm 32}$ Restan 1972 also discusses differences in li usage between the two languages.

might expect, the "deeper" syntax of the two languages is identical; the differences are due to the interaction of the syntax of li questions with a series of relatively "superficial" factors: differing prosodic constraints on clitics, idiosyncratic differences in the stress properties of particular lexical items, and differences in usage. This satisfying result underlines once again the utility of a parametric approach to the grammars of related languages.

References

- Alexander, Ronelle. (1994) "The Balkanization of Wackernagel's law". Indiana Slavic studies 7: 1–8. (G. Fowler, H. R. Cooper, Jr., and J. Z. Ludwig, eds. Proceedings of the 9th Biennial Conference on Balkan and South Slavic Linguistics, Literature, and Folklore. Bloomington, IN: Indiana University, Dept. of Slavic Languages.)
- Anderson, Stephen R. (1996) "How to put your clitics in their place, or why the best account of second-position phenomena may be something like the optimal one." The linguistic review 13(3–4): 165–91. (Special issue: M. Nespor, ed. *Phonology between words and phrases.*)
- Avgustinova, Tania. (1994) "On Bulgarian verb clitics." *Journal of Slavic linguistics* 2(1): 29–47.
- Avgustinova, Tania and Karel Oliva. (1991) *The structure of Bulgarian verb complex*. Saarbrücken: Universität des Saarlandes, Computerlinguistik. (*CLAUS* [Computerlinguistik an der Universität des Saarlandes], 4.)
- Baerman, Matthew. (1996) "The emergence of prosodic constraints in Macedonian." Presented at Workshop on conflicting constraints, Groningen,.
- Billings, Loren A. (1994) "An Optimality approach to the syntax and prosody of discourse particles." Presented at American Association of Teachers of Slavic and East European Languages conference, San Diego.
- Billings, Loren A. (1996a) "Macedonian enlarged stress domain and peninitial clisis". Presented at Cornell University, September.
- Billings, Loren A. (1996b) "Conflicting prosodic and syntactic constraints on special clitics". ZAS papers in linguistics 6: 43–50. (J. Kornfilt, ed. Proceedings of the Workshop on the phonology, morphology and syntax of clitics.)
- Billings, Loren, Christina Kramer, and Catherine Rudin (1996) "The unique prosodic properties of *li* in Macedonian." To be presented at American Association of Teachers of Slavic and East European Languages conference, Washington, December.
- Cavar, Damir. (1996a) "On cliticization in Croatian: Syntax or prosody?" ZAS papers in linguistics 6: 51–65. (J. Kornfilt, ed. Proceedings of the Workshop on the phonology, morphology and syntax of clitics.)
- Ćavar, Damir. (1996b) "Interface phenomena in Croatian and Polish" Presented at Universität Leipzig, November.
- Cavar, Damir and Chris Wilder. (1994) "'Clitic third' in Croatian". Linguistics in Potsdam 1: 25–63.
- Comrie, Bernard. (1976) "Irregular stress in Polish and Macedonian". *International review of Slavic linguistics* 1(2–3): 227–40.

- Deevy, Patricia L. (1995) "An Optimality theoretic analysis of stress in Macedonian." J. N. Beckman et al, eds. *Papers in Optimality Theory*. Amherst: Graduate Linguistic Student Association, 138–66. (*University of Massachusetts occasional papers*, 18.)
- Elson, Mark J. (1976) "The indefinite article in Bulgarian and Macedonian." Slavic and East European journal 20(3): 273–79.
- Elson, Mark J. (1993) "Collocational stress in Contemporary Standard Macedonian." Slavic and East European journal 37(2): 149–61.
- Englund, Birgitta. (1977) Yes/no-questions in Bulgarian and Macedonian: Form. Stockholm: Almqvist & Wiksell International. (Acta Universitatis stockholmensis, Stockholm Slavic studies, 12.)
- Englund, Birgitta. (1979) "The function of yes/no-questions (on basis of Bulgarian and Macedonian material)." In *Papers in Slavonic linguistics*. Stockholm, 16–26. (Meddelanden från institutionen för slaviska och baltiska språk", 18.)
- Ewen, Robert Charles. (1979) A grammar of Bulgarian clitics. Ph.D. dissertation, U. of Washington.
- Franks, Steven L. (1987) "Regular and irregular stress in Macedonian." *International journal of Slavic linguistics and poetics* 35–36: 93–142.
- Franks, Steven (1989) "The monosyllabic head effect." Natural language and linguistic theory 7(4): 551–63.
- Franks, Steven (1991) "Diacritic extrametricality vs. diacritic accent: A reply to Hammond." *Phonology* 8(1): 145–61.
- Friedman, Victor. (1993) "Macedonian". B. Comrie and G. C. Corbett, eds. *The Slavonic languages* (chap. 6). London: Routledge, 249–305.
- Garde, Paul. (1968) "Note sur l'enclise et la proclise en macédonien." Zbornik za filologiju i lingvistiku [Novi Sad: Matica srpska] 11: 31–37.
- Halle, Morris and Michael Kenstowicz. (1991) "The Free Element Condition and cyclic vs noncyclic stress." *Linguistic inquiry* 22(3): 457–501.
- Halpern, Aaron Lars (1992/1995) On the placement and morphology of clitics. Stanford, CA: CSLI Publications. (Published version of his Topics in the placement and morphology of clitics. Ph.D. dissertation, Stanford U.)
- Hammond, Michael (1989) "Lexical stress in Macedonian and Polish." *Phonology* 6(1): 19–38.
- Hauge, Kjetil Rå. (1976) The word order of predicate clitics in Bulgarian. Oslo: Universitetet i Oslo, Slavisk-Baltisk Institutt. (Meddelelser, 10.)
- Ilievski, P. Xr. and K. Peev (1983) "Neobična inovacija vo dolnovardarskite govori (Prokliza pri imperativ)." *Makedonski jazik* 34: 105–13.
- Izvorski, Roumjana. (1993) "On WH-movement and focus movement in Bulgarian". Presented at CONSOLE 2, Universität Tübingen.
- Izvorski, Roumyana. (1994) "Yes-no questions in Bulgarian: Implications for phrase structure". Presented at 9th Biennial Conference on Balkan and South Slavic Linguistics, Literature, and Folklore, Indiana University, Bloomington.
- Izvorski, Roumyana, Tracy Holloway King, and Catherine Rudin. (1996) "Against li lowering in Bulgarian". To appear in *Lingua*.
- Kager, René. (1993) "Stress in windows". ms., Universiteit Utrecht.
- Kenstowicz, Michael. (1991) "Enclitic accent: Macedonian, Italian, Polish." P. M. Bertinetto et al, eds. *Certamen phonologicum II: Papers from the 1990 Cortona Phonology Meeting*. Torino: Rosenberg & Sellier, 173–85.

- Kepeski, Petar and Jože Pogačnik. (1968) Makedonščina. Ljubljana: Državna založba Slovenije.
- King, Tracy Holloway (1993/1994/1995) "Focus in yes-no questions." Configuring topic and focus in Russian. Stanford: CSLI Publications, 137-163 (chap. 6). (Published, somewhat revised version of (chap. 5 [sic.]) of her (1993) Ph.D.: dissertation, Stanford U. This chapter also appeared as her (1994) "Focus in Russian yes-no questions". *Journal of Slavic linguistics* 2(1): 92–120.)
- Koneski, Blaže. (1957/1967) "Za akcentot". Gramatika na makedonskiot literaturen jazik. Skopje: Kultura. (Blaže Koneski: Izbrani dela vo sedum knigi, 6.)
- Korubin, B. (1974) "Mesto glagola sum kak vspomogatel nogo i kak svjazki v makedonskom literaturnom jazyke." N. Ju. Švedova et al, eds. Grammatičeskoe opisanie slavjanskix jazykov: Koncepcii i metody. Moskva: Nauka, 244–50.
- Kramer, Christina. (1985) Anglisko-makedonski razgovornik. Skopje: Univerzitet "Kiril i Metodij".
- Kramer, Christina E. (1986) Analytic modality in Macedonian. München: Verlag Otto Sagner. (Slavistische Beiträge 198.)
- Kramer, Christina. (1993) "The future in Macedonian". Presented at American Association of Teachers of Slavic and East European Languages conference,
- Lüdtke, H. (1979) Lucania. Pisa: Pacini. (Profilo dei dialetti italiani 17.)
- Lunt, Horace G. (1979) A grammar of the Macedonian literary language. Skopje: Državno Knigoizdatelstvo na NR Makedonija.
- Mihaljević, Milan. (1996) "The interaction of li and negation in Croatian Church Slavonic". To appear in G. Zybatow and U. Junghanns, ed. Formale Slavistik. (Presented at the First European Conference on "Formal Description of Slavic Languages", 1995.)
- Mussafia, A. (1898) "Enclisi o proclisi del pronome personale atono qual oggetto". Romania 27: 145-46.
- Penčev, Jordan. (1993). Bâlgarski sintaksis: Upravlenie i svârzvane. Plovdiv: Plovdivsko universitetsko izdatelstvo.
- Peperkamp, Sharon. (1996) "On the prosodic representation of clitics". U. Kleinhenz, ed. Interfaces in phonology. Berlin: Academie Verlag, 102-27. (Studia grammatica, 41.)
- Restan, Per. (1972) Sintaksis voprositel nogo predloženija; obščij vopros (Glavnym obrazom na materiale sovremennogo russkogo jazyka [Syntax of the interrogative sentence; yes-no questions (Mainly on the basis of Russian material). Oslo: Universitetsforlaget.
- Rivero, María-Luisa. (1993) "Bulgarian and Serbo-Croatian yes-no questions: V⁰ raising to -li vs. -li hopping". Linguistic inquiry 24(3): 564–75. Rudin, Catherine. (1990–91) "Topic and focus in Bulgarian". Acta linguistica
- hungarica 40(3-4): 429-47.
- Rudin, Catherine. (1992) "Kakvo li e li: Interrogation and focusing in Bulgarian". To appear in Balkanistika 10. (Special issue: V. Friedman, ed., Gedenkschrift for Zbigniew Gołąb.)
- Rudin, Catherine. (1993a) "On focus position and focus marking in Bulgarian questions". A. Davison et al, eds. Papers from the 4th annual meeting of the Formal Linguistics Society of Midamerica. Iowa City: Univ. of Iowa, 252–65.

- Rudin, Catherine. (1993b) "On the syntax of *li* questions in Bulgarian". Presented at American Association of Teachers of Slavic and East European Languages conference, Toronto.
- Rudin, Catherine. (1996) "Agr-O and Bulgarian pronominal clitics". To appear in S. L. Franks, ed. *Formal approaches to Slavic linguistics*, 5. Ann Arbor: Michigan Slavic Publications.
- Rudin, Catherine, Tracy Holloway King, and Roumyana Izvorski. (1995) "Focus in Bulgarian and Russian yes-no questions". Presented at the Workshop on Focus, Univ. of Massachusetts, Amherst.
- Rudin, Catherine and Christina Kramer. (1994) "Li questions in Macedonian". Presented at American Association of Teachers of Slavic and East European Languages conference, San Diego.
- Sadock, Jerrold M. (1991) Autolexical syntax: A theory of parallel grammatical representations. Chicago: University of Chicago Press.
- Scatton, Ernest. (1984) A Reference Grammar of Modern Bulgarian. Columbus: Slavica.
- [SSKG] (1975) Slovar slovenskega knjižnega jezika. Ljubljana: Slovenska akademija znanosti in umetnosti, Institut za slovenski jezik, v. 2 (*I–Na*).
- Toman, Jindřich. (1996) "A note on clitics and prosody". A. L. Halpern and A. M. Zwicky, eds. *Approaching second: Second position clitics and related phenomena*. Stanford, CA: CSLI Publications, 505–10 (chap. 15). (= CSLI lecture notes 61.)
- Tomić, Olga M. (1996a) "The Balkan Slavic nominal clitics." A. L. Halpern and A. M. Zwicky, eds. *Approaching second: Second position clitics and related phenomena*. Stanford, CA: CSLI Publications, 511–35 (chap. 16). (= CSLI lecture notes 61.)
- Tomić, Olga M. (1996b) "The Balkan Slavic clausal clitics." To appear in *Natural language and linguistic theory*. (Announced in *NLLT* 14(3).)
- Usikova, R. P. (1964) "Ob udarenii v sovremennom makedonskom jazyke". V. D. Koroljuk et al, eds. *Slavjanskaja i baltijskaja akcentologija*. Moskva: Nauka, 27–39. (Kratkie soobščenija Instituta slavjanovedenija, 41.)
- Usikova, R. P. (1985) Makedonskij jazyk: Grammatičeskij očerk, teksty dlja čtenija c kommentarijami i slovarem. Skopje: Makedonska kniga.
- Wackernagel, Jacob (1892) "Über ein Gesetz der indogermanischen Wortstellung." Indogermanische Forschungen 1: 333-436.
- Wilder, Chris and Damir Ćavar. (1994) "Long head movement? Verb movement in cliticization in Croatian". *Lingua* 93: 1–58.
- Zec, Draga and Sharon Inkelas. (1990) "Prosodically constrained syntax." S. Inkelas and D. Zec, eds. *The phonology-syntax interface*. Chicago: University of Chicago Press, 365–78 (chap. 18).

[Billings]
Institut für Slavistik
Universität Leipzig
Augustusplatz 9
D-04109 Leipzig
billings@rz.uni-leipzig.de

[Kramer]
Dept. of Slavic Lang. and Lit.
University of Toronto
21 Sussex Ave.
Toronto, Ont. M5S 1A1 Canada
ce.kramer@utoronto.ca

[Rudin]
Linguistics
Humanities Division
Wayne State College
Wayne, NE 68787 USA
crudin@wscgate.wsc.edu