Jasmina Milićević

PRONOMINAL AND VERBAL CLITICS IN SERBIAN: A MORPHOLOGICAL DESCRIPTION

1 Introduction

The paper proposes a formal description of the morphology of Serbian pronominal and verbal clitics. By 'formal' is meant a description which establishes correspondences between phonemic signifiers of the clitics and their 'reasonable' deeper representations (and vice versa) in such a way that it can be manipulated by a logical device — e.g. a computer program — and (ultimately) be used in automatic text synthesis/analysis. And 'morphology' is to be understood here as the composition of a clitic, which is a non-elementary linguistic sign, in terms of some elementary signs.[1]

I describe clitics in standard Serbian, but most of the data and the proposed description are valid for Croatian as well.

Serbian clitic pronouns and auxiliaries are, strictly speaking, enclitics. However, this distinction is irrelevant for my present purposes, since proclitics (= the negative particle NE, certain prepositions and subordinate conjunctions) will not be considered.

The description of Serbian clitics is based on the Meaning-Text Theory[2] — more specifically, on concepts/terminology of morphological description set forth in Mel'čuk's *Cours de morphologie générale* [= CMG]. Central linguistic concepts, such as word-form, morpheme, morph, inflectional category, etc., are taken for granted and will not be characterized; a few other concepts will be

---

1 On clitics in general, see Zwicky 1977, Klavans 1995 and Halpern 1995. On Serbian clitics, see Browne 1975, Stevanović 1986 and Spencer 1991: 351-358. Morin 1979 provides a morphological description of French clitics, which can serve for an interesting comparison with Serbian clitics.

2 See, for instance, Mel'čuk 1981, 1993: 41-75 and 1997.

introduced along the way. At the same time, the discussion will be made as theory-free as possible.

Pronominal and verbal clitics are dealt with in separate sections. In each case, I provide:

- A list of inflectional categories of clitics, with the corresponding inflectional meanings [= grammemes].

- Morphological description of clitics; it specifies for them the following representations:

  Deep-Morphological Representation. The Deep-Morphological Representation of a word-form w consists of the name of the lexeme L to which w belongs [= L(w)], supplied with all relevant grammemes: $L(w)g^1, g^2, ..., g^n$.

- **Surface-Morphological Representation,** consisting of two levels:

  1. **Morphemic** level, showing the morphemic composition of w, i.e. the way in which a lexical meaning and grammemes (selected at the deep-morphological level) are distributed into corresponding morphemes: $\{M_1\}$, $\{M_1\}$, ..., $\{M_n\}$;

  2. **Morphic** level, showing the morphic composition of w, i.e. the way in which morphemes (selected at the morphemic level) are expressed by corresponding morphs and megamorphs: $m_{1(\Sigma)} + m_{2(\Sigma)} + ... + m_{n(\Sigma)}$. (The symbol (Σ) stands for the syntactics of a morph/megamorph, i.e. a list of its combinatorial features. In this paper, only essential features are indicated, and not for all signs.)

- Discussion of the morphology of clitics; it is intended to highlight special morphological properties thereof.

Clitics are special in that even their morphological status is unclear. The question of whether they are word-forms, affixes or a third type of linguistic entity is controversial. (See, for instance, Spencer 1991, 375-392 and Kaiser (ed.) 1997.) However, within the Meaning-Text framework, morphological items can be either word-forms or parts of word-forms (i.e. radicals or affixes) — no third option is available. I believe that clitics are word-forms — albeit rather bizarre ones — and that this can be proven. This point is discussed in **3.1**, where I analyze the apparently suffix-like behavior of the

Serbian future auxiliary, showing that, in spite of the appearances, it is a word-form and not a suffix.

It will also be shown that Serbian clitics feature a number of phenomena which do not exist or are less prominent elsewhere in the language — zero radicals, unique affixes, pervasive suppletion, and other well-known irregularities observed in pronominal and auxiliary systems, especially in Slavic languages.

Symbols

Ø   : zero morph/megamorph
+   : morphic boundary
#   : word-form boundary
⇔   : correspondence between elements of two adjacent levels of representation
⇒   : correspondence between elements of the same level of representation

Tonal accents

short-falling	short-rising	long-falling	long-rising
["]	[`]	[^]	[´]

Pronunciation

dj /ʒ/	nj /ɲ/	h /x/

2 Pronominal clitics

2. 1 Personal pronouns

2. 1. 1 First and second person pronouns
JA 'I', TI 'thou', MI 'we' and VI 'you'

Inflectional categories

TONICITY   = {full, clitic}
CASE   = {genitive, dative, accusative, ...}

Clitic word-forms

{JA}:

me	JA cl, gen	⇔	{JA}, {CL}, ⏟ m-(pers. pron, 1, sg)	{GEN} ⏟ -e	⇔
mi	JA cl, dat	⇔	{JA}, {CL}, ⏟ m-(pers. pron, 1, sg) +	{DAT} ⏟ -i	⇔
me	JA cl, acc	⇔	{JA}, {CL}, ⏟ m-(pers. pron, 1, sg) +	{ACC} ⏟ -e	⇔

[Full word-forms: mèn+e, mèn+i, mèn+e.]

{TI}:

te	TI cl, gen	⇔	{TI}, {CL}, ⏟ t-(pers. pron, 2, sg)	{GEN} ⏟ -e	⇔
ti	TI cl, dat	⇔	{TI}, {CL}, ⏟ t-(pers. pron, 2, sg) +	{DAT} ⏟ -i	⇔
te	TI cl, acc	⇔	{TI}, {CL}, ⏟ t-(pers. pron, 2, sg) +	{ACC} ⏟ -e	⇔

[Full word-forms: tèb+e, tèb+i, tèb+e.]

{MI}:

nas	MI cl, gen	⇔	{MI}, {CL},	{GEN}	⇔
			n-(pers. pron, 1, pl) +	-a s	
nam	MI cl, dat	⇔	{MI}, {CL},	{DAT}	⇔
			n-(pers. pron, 1, pl) +	-a m	
nas	MI cl, acc	⇔	{MI}, {CL},	{ACC}	⇔
			n-(pers. pron, 1, pl) +	-a s	

[Full word-forms: n+âs, n+ăma, n+âs.]

{VI}:

vas	VI cl, gen	⇔	{VI}, {CL},	{GEN}	⇔
			v-(pers. pron, 1, pl) +	-a s	
vam	VI cl, dat	⇔	{VI}, {CL},	{DAT}	⇔
			v-(pers. pron, 1, pl) +	-a m	
vas	VI cl, acc	⇔	{VI}, {CL},	{ACC}	⇔
			v-(pers. pron, 1, pl) +	-a s	

[Full word-forms: v+âs, v+ăma, v+âs.]

2. 1. 2 Substitute [= 3p] pronoun ON 'he'

Inflectional categories

TONICITY　= {full, clitic}
NUMBER　= {sg, pl}
GENDER　= {masc, fem, neut}
CASE　　= {genitive, dative, accusative, ...}

Clitic word-forms

ga	ON $_{\text{cl, sg, masc, gen}}$ ⇔　　{ON}, {CL},　　{SG.MASC.GEN}　⇔ 　　　　Ø-$_{\text{(pers. pron, 3)}}$ + 　　　**-ga**
mu	ON $_{\text{cl, sg, masc, dat}}$ ⇔　　{ON}, {CL},　　{SG.MASC.DAT}　⇔ 　　　　Ø-$_{\text{(pers. pron, 3)}}$ + 　　　**-mu**
ga	ON $_{\text{cl, sg, masc, acc}}$ ⇔　　{ON}, {CL},　　{SG.MASC.ACC}　⇔ 　　　　Ø-$_{\text{(pers. pron, 3)}}$ + 　　　**-ga**

[Full word-forms: nj+èga, nj+èmu, nj+èga.]

je	ON $_{\text{cl, sg, fem, gen}}$ ⇔　　{ON}, {CL},　　{SG.FEM.GEN}　⇔ 　　　　j-$_{\text{(pers. pron, 3)}}$ + 　　　**-e**
joj	ON $_{\text{cl, sg, fem, dat}}$ ⇔　　{ON}, {CL},　　{SG.FEM.DAT}　⇔ 　　　　j-$_{\text{(pers. pron, 3)}}$ + 　　　**-oj**

je	ON cl, sg, fem, acc ⇔ {ON}, {CL}, {SG.FEM.ACC} ⇔ j-(pers. pron, 3) + **-e** | **not** before BITI, cl, present, 3, sg ON, cl, sg, fem, gen
ju	ON cl, acc, sg, acc ⇔ {ON}, {CL}, {SG.FEM.ACC} ⇔ j-(pers. pron, 3) + **-u** | before BITI, cl, present, 3, sg ON, cl, sg, fem, gen

[Full word-forms: nj+ê, nj+ôj, nj+û.]

ga	ON cl, sg, neut, gen ⇔ {ON}, {CL}, {SG.NEUT.GEN} ⇔ Ø-(pers. pron, 3) + **-g a**
mu	ON cl, sg, neut, dat ⇔ {ON}, {CL}, {SG. NEUT.DAT} ⇔ Ø-(pers. pron, 3) + **-m u**
ga	ON cl, sg, neut, acc ⇔ {ON}, {CL}, {SG. NEUT.ACC} ⇔ Ø-(pers. pron, 3) + **-g a**

[Full word-forms: nj+èga, nj+èmu, nj+èga.]

ih	ON cl, pl, gen ⟺ {ON}, {CL}, {PL.GEN} ⟺ Ø-(pers. pron, 3) + **-i h**
im	ON cl, pl, dat ⟺ {ON}, {CL}, {PL.DAT} ⟺ Ø-(pers. pron, 3) + **-i m**
ih	ON cl, pl, acc ⟺ {ON}, {CL}, {PL.ACC} ⟺ Ø-(pers. pron, 3) + **-i h**

[Full word-forms: nj+îh, nj+ìma, nj+îh.]

Comments

- A question may arise of whether tonicity is an inflectional category (of personal pronouns). On the one hand, word-forms belonging to the same lexeme feature only inflectional differences.[3] Since the clitic and the corresponding full forms of a given personal pronoun clearly belong to the same lexeme, the clitic/full opposition must be considered as inflectional in nature. On the other hand, keeping in mind that regularity of expression (besides obligatory character) is a definitorial feature of an inflectional category, tonicity is a rather problematic one — its grammemes lack their own, independent expression. A grammeme of tonicity is always expressed together with the lexical meaning it characterizes; it is amalgamated to the radical of the pronominal word-form (cf. supra). However, in my view the first consideration overrides the second one, that is, 'distributing' clitic and full pronoun forms into separate lexemes would be more embarrassing than admitting tonicity as an inflectional category. In fact, instances of inflectional categories having no independent expression are known; for example, this is the case with such obvious inflectional categories as (grammatical) case and

---

3 Remember the crucial difference between inflection and derivation: inflectional meanings distinguish word-forms belonging to the same lexeme, while derivational meanings distinguish word-forms belonging to different lexemes.

number in French personal pronouns; cf. the following 3p forms, repre-
senting a fused expression of the corresponding lexical meaning and the
grammemes of tonicity, number and case: [cl, sg, dat] **lui** 'to-him/her', and
[cl, pl, dat] **leur** 'to-them'. (As we shall see, the same considerations hold for
the auxiliaries: for them tonicity is an inflectional category as well.)[4]

- The genitive and the accusative forms of (both clitic and full) pronouns are
  homophonous, the only exception being the feminine singular forms of the 3p
  pronoun ON: genitive **njê** (full) and **je** (clitic) vs. accusative **njû** (full) and **je**
  <ju> (clitic).

- The 1/2p pronouns lack the inflectional category of number — I believe that
  mi 'we' is not the plural of **ja** 'I', no more than **vi** 'you' is the plural of **ti**
  'thou'. It follows that JA and MI, resp. TI and VI, are different lexemes. The
  grammatical number of a 1/2p pronominal lexeme is a feature of its syntactics
  (more precisely, of the syntactics of its radical).

- There are two possible ways of describing substitute (= 3p) pronouns. One
  can say that there are three different lexemes ON 'he', ONA 'she', ONO 'it',
  which do not inflect for gender (as is typical for nouns) and whose oblique-
  case plural forms are homophonous. Alternatively, as I have done here, one
  can postulate a single lexeme ON, with gender neutralization in the oblique
  cases of the plural, and with homophonous oblique forms in the
  masculine/neuter singular. This corresponds to the viewpoint of Serbian
  grammars. Substitute pronouns have the adjectival type of declension, the
  same one that applies to demonstratives OVAJ 'this', TAJ 'that-PROXIMAL', ONAJ
  'that-DISTALL' and to the interrogative/relative pronoun KOJI 'who/which'.

- Clitics **je** 'her-ACC' and **ju** 'her-ACC' are allolexes (of the lexeme ON),
  distributed according to the morphological context — **ju** is used before the
  auxiliary clitic **je** (BITI (aux) pres, ind, 3, sg) and the pronominal clitic **je** (ON
  sg, fem, gen); the form **ju** appears elsewhere:

(1) a. **Vide +o*   *ju <*je>   je.*
    See   PART. MASC.SG   she-CL.ACC   be[aux]-CL.PRES.IND.3SG
    '[He] saw her'.

---

4 Alternatively, as W. Browne has suggested (personal communication), one can say that
   tonicity is expressed by presence/absence of stress (in a pronoun/auxiliary). Although all
   theoretical consequences of such a solution are not clear to me at this point, this path seems
   worth exploring, since under such treatment tonicity would be a 'well-behaved' inflectional
   category.

cf.   *Video je*   *jeste,*   *ali joj se obratio nije.*
       she-CL.ACC   be[aux]-FULL.PRES.IND.3SG
       '[He] did see her, but didn't address her'.

    b. *Liši +o*   *sam*   *ju <*je> je.*
       deprive PART.MASC.SG   be[aux]-CL.PRES.IND.1SG   she-CL.ACC she-CL.GEN
       '[I] deprived her of-her'.

cf.   *Lišio*   *sam*   *je*   *njê.*
                                 she-CL.ACC   she-FULL.GEN
       Id.

• The inventory of pronominal clitics given above includes no reflexive clitics.
  This is at variance with the view of Serbian grammars, according to which
  pronominal clitics do include the following two reflexive clitics:
  1) **se**[refl.accus.], the accusative clitic form of the reflexive pronominal
  lexeme SEBE 'oneself'; which functions as a direct object of genuinely
  reflexive verbs (*brijati se* '[to] shave oneself', *prati se* '[to] wash oneself') and
  2) **se**[refl.invar.], an invariable reflexive clitic (with no corresponding full
  form), used with all other reflexive verbs, such as pseudo-reflexives (*brijati
  se [kod berberina]* '[to] shave [at the barber's]'), reciprocals (*voleti se* '[to]
  love each other'), middle verbs (*otvoriti se* '[to] open'), inherent reflexives
  (*sećati se* '[to] remember'), etc.

However, I believe that in modern Serbian there is just one invariable
reflexive clitic, **se**[refl.invar.], the non-morphological, i.e. analytical, marker
of the reflexive voice, historically related to but synchronically distinct from
SEBE. In other words, there is no **se**[refl.accus.] and the lexeme SEBE has no
clitic forms. (There used to be the opposition **se**[refl.accus.] / **si**[refl.dat.],
but the dative form is extinct in modern Serbian.) Let me quote two linguistic
considerations to support this claim.

1) The relation between **se** and **sebe** is very different from that between the
obvious pronominal clitics and their full forms. Thus, obvious pronominal
clitics alternate freely with their full counterparts; for instance, **te**[2p]cl, acc can
always be replaced by **tebe**[2p]full, acc, just as **ga**[3p]cl, masc, acc, sg can always
be replaced by **njega**[3p]full, masc, acc, sg, etc.: *Volim te / tebe* '[I] love
you/YOU', *Volim ga / njega* '[I] love him/HIM' (if the com–municative
structure licenses this). In contrast to this, the alternation **se** ~ **sebe**, usually
taken as an indication that **se** belongs to the lexeme SEBE, is extremely
restricted. First, **se** is not always replaceable with **sebe**: [?]*Brije sebe svako
jutro* [He] shaves himself every morning' sounds rather bizarre and one would

definitely prefer **se** in this context. Furthermore, to express emphasis, one would rather say *Sâm se brije* '[He] shaves by himself', instead of using **sebe**. Second, **sebe** is not always replaceable with **se**: *Voli sebe* '[He] loves himself' is normal, but **Voli se* Id. is not.

It seems, therefore, that when se 'alternates' with **sebe**, this is actually the alternation between the use of a grammeme — reflexive voice with the meaning of genuine reflexivity, as in *Brije se* —, and the use of a full lexeme SEBE, functioning as a Direct Object of a verb in the active voice, as in *Brije sebe*; cf. the parallelism of the last sentence with *Brije Petra* '[He] shaves Peter', where we see a non-pronominal lexeme in the role of a Direct Object.

The view that there is just one se (which is not a lex of SEBE) allows for a more natural account of the relevant linguistic facts than the traditional one, since, if we posit two se, we have to say, for instance, that *Brije se* features a non-reflexive verb in the active voice with a Direct Object, while *Vole s e* '[They] love one another' contains a verb in the reflexive voice (reciprocal), with no Direct Object involved.

2) The reflexive clitic adjunct se has different combinatorial properties with respect to all obvious pronominal clitics: unlike the former, it combines with the dative possessor and triggers the obligatory deletion of the verbal clitic **je** (see Milićević forthcoming).

2. 2 Morphology of personal pronouns

If we consider only clitic pronouns and their full counterparts (i.e. only the genitive, dative and accusative forms),[5] we can state that they are all word-forms, each made up of two signs:

---

5 For comparison, here are the nominative forms of personal pronouns:

	[1 sg]	[2sg]	[1pl]	[2pl]
[nom]	ja	ti	m i	v i

	[3p]					
	[sg]			[pl]		
	[masc]	[fem]	[neut]	[masc]	[fem]	[neut]
[nom]	on+Ø	on+a	on+o	on+i	on+e	on+a

All nominative forms are special:
• They are always full (in contrast, for instance, to French subject pronouns, which have both full and clitic forms, cf. **moi**-full ~ je-clitic 'I', **toi**-full ~ tu-clitic 'thou', etc.)

- a radical, which is a megamorph[6], representing a 'fused' expression of a given lexical meaning and of a grammeme of tonicity;

- a suffix, which is a morph, expressing cumulatively all other grammemes (see below).

In the case of 1/2sg pronouns, radicals of a clitic and of the corresponding full pronoun are (weakly) suppletive with respect to each other:

Clitic   **Full**

$\{JA\}, \{CLITIC\} \Leftrightarrow$ **m** $\{JA\}, \{FULL\} \Leftrightarrow$ **mèn**
$\{TI\}, \{CLITIC\} \Leftrightarrow$ **t** $\{TI\}, \{FULL\} \Leftrightarrow$ **tèb**

Alternatively, one can say that the radical of a clitic is obtained by truncation of a segment (= stressed V+coda or stressed V+attack) of (the signifier of) the corresponding full form, i.e. by an alternation. It is obvious that such an alternation has a very limited applicability (= it 'works' only for 1/2 sg pronouns). However, the same (or a related) alternation can be said to apply also to the radicals of auxiliary verbs (see 3. 2. 1), and, in the case of 3p pronouns, to the suffixes of full forms in order to get suffixes of clitics (see below). All this enhances the regularity of the alternation. Therefore, we have here two equally plausible descriptions and a principled choice between them seems difficult. Although for the time being I have opted for suppletion, both solutions should probably be admitted (cf. non-uniqueness of morphological solutions, CMG-4, 133).[7]

As for 1/2pl pronouns, they have homophonous radicals in clitic and full forms:

- A 1/2p nominative pronoun is a megamorph expressing a given lexical meaning alongside the grammemes of tonicity and case, for instance $\{JA\}, \{FULL\}, \{NOM\} \Leftrightarrow$ **ja**.
- A 3p nominative pronoun consists of a radical (a megamorph) and a suffix, zero in the masculine: $\{ON\}, \{FULL\}, \{SG.MASC.NOM\} \Leftrightarrow$ **on** + Ø.

6 The other term, more current in literature, is *'portmanteau* morph'. However, the term should be avoided, since a megamorph is not a particular case of morph, but a sign of a different nature. While an 'ordinary' morph always expresses a single morpheme, a megamorph expresses a series of morphemes.

7 Acquisition and/or dialectal data could provide further evidence in favor of one of the two solutions. Thanks to Y-C. Morin for bringing this to my attention.

Clitic			Full		
{MI }, {CLITIC}	⟺	n	{MI }, {FULL}	⟺	n
{VI }, {CLITIC}	⟺	v	{VI }, {FULL}	⟺	v

In 1/2p pronouns, the suffix expresses only the grammeme of case.

Full and clitic forms of the 1/2sg pronouns have identical sets of suffixes. Here are the relevant forms of the 1sg pronoun JA.

	Clitic	Full
[gen]	m + e	mèn + e
[dat]	m + i	mèn + i
[acc]	m + e	mèn + e

Suffixes of the 1/2pl clitic and full pronouns are slightly different; compare the clitic and the corresponding full forms of the 1pl pronoun MI:

	Clitic	Full
[gen]	n + as	n + âs
[dat]	n + am	n + ǎma
[acc]	n + as	n + âs

Here, it is the suffixal part of the full form that bears the stress. (It cannot be otherwise, since the radical consists of a single consonant. The same situation obtains with the 3p pronouns, see below.) This is in contrast with 1/2sg full pronouns, which have stress on the radical.

The suffixes -as and -am(a) are (almost) unique, but with pronouns — as well as with determiners and auxiliaries — this is common enough.

The forms of 3p pronouns consist of 1) a radical megamorph, expressing the lexical meaning (= 'he') plus a grammeme of tonicity, and 2) a suffixal morph, expressing cumulatively the grammemes of number, gender and case. (But in the oblique cases of the plural, only number and case are expressed, because of gender neutralization; cf. supra.)

Radicals of 3p pronoun clitic forms are suppletive with respect to each other and with respect to radicals of corresponding full forms. Masculine and neuter

clitic forms have a zero radical. This is a rather rare phenomenon, not only in Serbian, but cross-linguistically (for obvious semiotic reasons). Zero radicals are found in Kirundi (Bantu) demonstratives, as well as in Awa (Papuan) verbs, see CMG–4, 79-81.

	Clitic	Full
{ON }, {CLITIC} ⇔	Ø I gender ≠ fem	{ON}, {FULL} ⇔ nj
	j I gender = fem	

Full and clitic forms have slightly different sets of suffixes. Compare the relevant paradigms, (partially) repeated here for convenience:

	Clitic	Full
[sg, masc, gen]	Ø + ga	nj + èga
[sg, fem, gen]	j + e	nj + ê
[sg, masc, dat]	Ø + mu	nj + èmu
[sg, fem, dat]	j + oj	nj + ôj
....		
[pl, dat]	Ø + im	nj + ìma
[pl, acc]	Ø + ih	nj + îh

These are typical adjectival suffixes; cf. a fragment of the paradigm of the adjective TUDJ 'foreign': [sg, masc, gen] túdj+eg(a), [sg, masc, dat] túdj+em(u), [sg, fem, gen] túdj+e, [sg, fem, dat] túdj+oj, etc. (neuter forms have been omitted since they are identical to the masculine ones).

3 Verbal clitics
BITI (past/conditional auxiliary) and HTETI (future auxiliary)

3. 1 Auxiliary verbs

Inflectional categories

TONICITY = {full, clitic}
TENSE   = {present, aorist, ...}
MOOD   = {indicative, ...}

PERSON   = {1, 2, 3}
NUMBER   = {sg, pl}

3. 1. 1 BITI

Clitic word-forms

sam	BITI cl, pres, ind, 1, sg ⟺ {BITI},{CL}, S-(past aux) +	{PRES.IND.1.SG} -(a)**m**	⟺
si	BITI cl, pres, ind, 2, sg ⟺ {BITI},{CL}, S-(past aux) +	{PRES.IND.2.SG} **-i**	⟺
je	BITI cl, pres,.ind, 3, sg ⟺ {BITI},{CL}, **j**-(past aux) +	{PRES.IND.3.SG} **-e**	⟺
smo	BITI cl, pres, ind, 1, pl ⟺ {BITI},{CL}, S-(past aux) +	{PRES.IND.1.PL} **-m o**	⟺
ste	BITI cl, pres, ind, 2, pl ⟺ {BITI},{CL}, S-(past aux) +	{PRES.IND.2.PL} **-t e**	⟺
su	BITI cl, pres, ind, 3, pl ⟺ {BITI},{CL}, S-(past aux) +	{PRES.IND.3.PL} **-u**	⟺

[Full word-forms: jès+(a)m, jès+i, jȅst+e, jès+mo, jès+te, jès+u.]

bih	BITI **cl, aorist, ind, 1, sg** ⇔ {BITI},{CL}, {AORIST.IND.1.SG} ⇔ **bi-**(condit. aux) + **-h**
bi	BITI **cl, aorist, ind, 2, sg** ⇔ {BITI},{CL}, {AORIST.IND.2.SG} ⇔ **bi-**(condit. aux) + **-Ø**
bi	BITI **cl, aorist, ind, 3, sg** ⇔ {BITI},{CL}, {AORIST.IND.3.SG} ⇔ **bi-**(condit. aux) + **-Ø**
bismo	BITI **cl, aorist, ind, 1, pl** ⇔ {BITI},{CL}, {AORIST.IND.1.PL} ⇔ **bi-**(condit. aux) + **-smo**
biste	BITI **cl, aorist, ind, 2, pl** ⇔ {BITI},{CL}, {AORIST.IND.2.PL} ⇔ **bi-**(condit. aux) + **-ste**
bi	BITI **cl, aorist, ind, 3, pl** ⇔ {BITI},{CL}, {AORIST.IND.3.PL} ⇔ **bi-**(condit. aux) + **-Ø**

[Full word-forms: bȉ+h, bȋ+Ø, bȋ+Ø, bȉ+smo, bȉ+ste, bȉ++++Ø.]

3. 1. 2 HTETI

Clitic word-forms

ću	HTETI cl, pres, ind, 1, sg ⟺ {HTETI },{CL}, {PRES.IND.1.SG} ⟺ ĉ-(future aux) + **- u**
ćeš	HTETI cl, pres, ind, 2, sg ⟺ {HTETI },{CL}, {PRES.IND.2.SG} ⟺ ĉe-(future aux) + **-š**
će	HTETI cl, pres, ind, 3, sg ⟺ {HTETI },{CL}, {PRES.IND.3.SG} ⟺ ĉe-(future aux) + **-Ø**
ćemo	HTETI cl, pres, ind, 1, pl ⟺ {HTETI },{CL}, {PRES.IND.1.PL} ⟺ ĉe-(future aux) + **-m o**
ćete	HTETI cl, pres, ind, 2, pl ⟺ { HTETI },{CL}, {PRES.IND.2.PL} ⟺ ĉe-(future aux) + **-t e**
će	HTETI cl, pres, ind, 3, pl ⟺ {HTETI },{CL}, {PRES.IND.3.PL} ⟺ ĉe-(future aux) + **-Ø**

[Full word-forms: hòć+u, hŏće+š, hŏće+Ø, hòće+mo, hòće+te, hòće+Ø.]

Comments

- BITI$_{(aux)}$ and HTETI$_{(aux)}$ both have semantically full counterparts. There are at least two semantically full lexemes BITI '[to] be' — the copula and the locative verbs. Both have clitic forms only in the present tense; they are identical to the present tense clitic forms of BITI$_{(aux)}$. Lexical (= semantically full) verb HTETI '[to] want' has clitic forms in the present tense, identical to those of HTETI$_{(aux)}$.

- HTETI$_{(aux)}$ is defective: it has only the present tense forms.

- The present tense forms of BITI$_{(aux)}$ are used as past tense auxiliary and its aorist forms as conditional auxiliary. HTETI$_{(aux)}$ is the future tense auxiliary.

- The clitic and the full forms of the conditional auxiliary differ only in stress, i.e. their segmental signifiers are identical.[8]

  Spoken language has a tendency to use **bi** [3sg] instead of all other forms of the conditional auxiliary. This means that the latter is in the process of becoming an invariable verbal form. In Macedonian and Russian, this process has gone even further: the corresponding forms **bi**, resp. **by**, cognate with the 3sg Serbian form, are not verbs any more, but particles.

- Note the truncation of the stem final vowel before the 1sg marker -u in the corresponding clitic and full forms of the future auxiliary: **ć+u** [$\Leftarrow$ će +u] and **hòć+u** [$\Leftarrow$ hòće+u].

- The future auxiliary features suffix-like behavior in 'synthetic' future tense constructions.

  With -ti infinitives, the future can be expressed either analytically, as in (3a), or 'synthetically,' as in (3b).[9]

  (3) a. On *će* pevati.
        'He will sing'.

---

8 BITI$_{(copula)}$ and BITI$_{(locative)}$ have the same aorist forms as BITI$_{(aux)}$, with the exception of the 3pl form; they have **bi+še** while BITI$_{(aux)}$ has **bi+Ø**. (Note that Serbian has lost the aorist, as well as the imperfect; these tenses are used only by some elder speakers, and only with a small number of verbs.)

9 Interestingly enough, this does not hold for the lexical verb HTETI '[to] want' — it can be used only in an analytic construction. Cf. [1sg] *ću hteti*, [2sg] *ćeš hteti*, etc., and not **hteću, *htećeš*, etc.

(3) b. *Pevaće.*
   Id.

In a 'synthetic' future construction, the auxiliary provokes the truncation of the infinitive ending and fuses with the infinitive stem: *peva +ti će ⇒ pevaće.* The corresponding truncation rules are:

R1~~truncation~~   /i/ ⇒ A ι /t/__/ć/   [cf. Croatian usage *Pevat će.*]

R2~~truncation~~   /t/ ⇒ A ι __ /ć/.

If the verbal stem ends in /s/ or in /st/, the final /s/ is palatalized to become /š/ before *će* :

R3~~assimilation~~   /s/ ⇒ /š/ ι __ /ć/.

For instance:

(4) a. *pas +ti će ⇒ pasće ⇒ pašće* '[he] will graze'

   b. *rast +ti će ⇒ rastće ⇒ rasće ⇒ rašće* '[he] will grow'

Such phonological interaction is typical of parts of a word-form; normally, it doesn't happen between word-forms. Cf. (5), where the same alternations take place in the noun stem before the nominal suffix [instr. sg] -ju:

(5) *mladost +ju ⇒ mladošću* 'youth'.

So, in (3b) and (4), *će* behaves as a suffix; the same is true for all the six forms of the future auxiliary. However, the following properties of the future auxiliary show that it is not a suffix, but a word-form, 'disguised' as a suffix[10]:

- *će* can be factored out of coordinate constructions:
(6) *Pevaće i igraće. ⇒ Pevaćeᵢ i igrati _ᵢ.*
   '[He] will sing and dance'.

---

10 For criteria that can be used to distinguish between inflection and cliticization, see Zwicky and Pullum 1983.

In Serbian, such a factorization is never possible with genuine suffixes:

(7) *Peva +m i igra +m. ≠> *Peva +m$_j$ i igra +_$_j$.<*Peva +_$_j$ i igra +m$_j$.>*
     '[I] sing and dance'.

Note that *će* can be omitted only in the second conjunct, never in the first one: **Igrati _$_j$ i pevaće.* This is typical of word-forms that are syntactic governors, not of suffixes. (In the construction [X→A] *and* [X→B], only the second X, but not the first, can be omitted: X→A *and* B vs. *A *and* X→B. For instance, in the sentence *John reads a book and* [*he*] *reads a newspaper*, only the second occurrence of *read* can undergo deletion: *John reads a book and a newspaper* vs. **John a book and reads a newspaper.*) Cf. the same behavior of the French past auxiliary *avoir* in *Il a$_j$ lu et _$_j$ apprecié....*vs. **Il _$_j$ lu et a$_j$ apprecié...* 'He has read and appreciated...'

With affixes, it is the other way around — in group inflection (see CMG-4, 293), it is always the last 'conjunct' that remains. Cf. also Spanish coordinate tmesis (with a derivational suffix): *rapida + _$_j$ y segura +mente$_j$* 'rapidly and surely', but not **rapida +mente$_j$ y segura + _$_j$*.

- *će* can be coordinated with the corresponding full form *hoće*, as in (8b):

(8) a. *Ne znam   kad   će   doći   i*
        NEG know-PRES.IND.1SG when hteti$_{[aux]}$-CL.PRES.IND.3SG arrive-INF and
        da li   će   nas   posetiti.
        Q-FULL   we-CL.ACC visit-INF
        '[I] don't know when [he] will arrive and whether [he] will call on'.

     b. *Ne znam kad će doći i hoće   li nas posetiti.*
                                hteti$_{[aux]}$-FULL.PRES.IND.3SG Q-CL
        Id.

- *će* alternates with *hoće* in answers:

(9) a. *Pevaće?*
        Sing=hteti$_{[aux]}$-CL.PRES.IND.3SG?   = 'Will [he] sing'?

     b. *Hoće.*
        hteti$_{[aux]}$-FULL.PRES.IND.3SG   = '[He] will'.

This is the standard way of answering general questions, parallel to English usage. DA 'yes' can also be used, but it is more idiomatic to use the full form of the auxiliary with it: *Da, hoće*.

* *će* can take a clause as its surface-syntactic actant:

(10) *On   će* — completive-PROP→ [*da   dodje   sutra*]
    he-NOM.SG   hteti[aux]-CL.PRES.IND.3SG   that-CONJ   come-PRES. IND.3SG   tomorrow
    'He will come tomorrow'.

* *će* can be used alone:

(11) *Kiša   će*.
    rain   hteti[aux]-CL.PRES.IND.3SG
    '[The] rain will'. = 'It is going to rain'./ 'It looks like rain'.

cf. the lexical HTETI in *Šta češ ti ovde?*   what want-CL.PRES.IND.3SG   you-NOM   here
'What [do] you-SG want here'? ≈ 'What are you doing here'? / 'What is the purpose of your being here'?

Another argument in favor of treating the future auxiliary as a word-form comes from regional usage. Serbian has a number of infinitives ending in -*ći*, which cannot be used in 'synthetic' future constructions: only *Doći će* 'He will come' is possible. But, in some regions of Vojvodina (Northern Serbia), forms such as *Dočće*, pronounced /dočće/, are in use. In Serbian, geminate consonants are not found in this position. (They can occur only on a prefix/stem boundary, where there is a strong secondary stress on the prefix, as in *naj+jači* 'the strongest' and *trans+sibirski* 'trans-Siberian'.) So, the sequence *čć* indicates that there are indeed two word-forms in *Dočće*.[11]

Note that if this analysis is correct, i.e. if the future auxiliary is a word-form, then, contrary to the traditional grammar's viewpoint, the so-called synthetic (or simple) future actually doesn't exist in Serbian. For a previous attestation of this, see Browne 1970.

---

11 It should be noted that a more flexible approach to the quesiton of the morphological status of *će* is also possible. According to Y.-C. Morin (personal communication), in a language where a grammaticalization of a clitic is under way, one shouldn't expect all the forms (of the clitic) to be affected (by the grammaticalization) at the same time. Thus, *će* can very well be a word-form in one context (as in *dočće*) and a suffix in another (as in *pevaće*).

3. 2 Morphology of auxiliary verbs

In **3.2.1**, I discuss the affirmative forms of auxiliary verbs (shown above).
Negative auxiliary forms, which have special properties, will be presented and
discussed in **3.2.2**.

3. 2. 1 Affirmative forms

Affirmative auxiliaries feature the same morphological structure as personal
pronouns. They too are composed of a radical and a suffix. The radical is also a
megamorph, expressing a given lexical meaning and a grammeme of tonicity,
while the suffix expresses cumulatively the grammemes of tense, mood, person
and number. This holds for clitic and for full forms alike (cf. supra, p. 244ff.).

As in the case of personal pronouns, clitic and full forms of auxiliaries have
suppletive radicals. (Here, again, the alternation of truncation seems plausible as
an alternative to suppletion; cf. p. 242)

{BITI}, {CL} ⇔ s, j, bï   {BITI}, {FULL}   ⇔ jès, jêst, bï, bï, etc.

{HTETI}, {CL} ⇔ će   {HTETI}, {FULL}   ⇔ hòće, hôće, etc.

Clitic and full forms of auxiliaries have identical sets of suffixes. Compare
again the relevant forms of the past tense auxiliary:

	Clitic	Full
[1 sg]	s+(a)m	jès+(a)m
[2 sg]	s+i	jès+i
[3 sg]	j+e	jêst+e
[1 pl]	s+mo	jès+mo
[2 pl]	s+te	jès+te
[3 pl]	s+u	jès+u

The vowel in -(a)m has been added historically; otherwise the whole form is not
pronounceable. (The combination *sm # is impossible in Serbian. Cf.
socijalizam, nepotizam and not *socijalizm, *nepotizm.*)

The same suffixes, except in 2/3 sg, are used with lexical verbs as well. Cf. for instance, the present tense of *gledati* '[to] watch': gleda+m, gleda+š, gleda+Ø, gleda+mo, gleda+te, gleda(j)+u.

Note that -m is a typical 1sg present suffix in Serbian, while -u (as in hoć+u) is rare in this role. Actually, the only other verb with this suffix is **mog+u** '[I] can'. (In Russian, it is exactly the opposite: only three verbs — DAT´ '[to] give', SOZDAT´ '[to] create' and EST´ '[to] eat' — and their derivatives — have -m as 1sg present marker; all the others have -u.)

3. 2. 2 Negative forms

BITI

present

1sg	2sg	3sg	1pl	2pl	3pl
nísam	nísi	nìje	nísmo	níste	nísu

aorist

1sg	2sg	3sg	1pl	2pl	3pl
nĕ bih	nĕ bi	nĕ bi	nĕ bismo	nĕ biste	nĕ bi

HTETI

1sg	2sg	3sg	1pl	2pl	3pl
néću	nêćeš	nêće	néćemo	néćete	nêće

In the negative forms of auxiliary verbs, the clitic/full opposition is neutralized. (Conversely, one can say that the clitic/full opposition exists only in the present/aorist affirmative for BITI and in the present affirmative for HTETI.) Negative forms are always full (= stressed).

Affirmative		Negative
Clitic	**Full**	**No clitic/full opposition**
Písao je.	*Jȅste písao.*	*Nìje písao.*
write-PART be[aux]-CL.PRES.3SG	be[aux]-FULL.PRES.3SG	be[aux]-PRES.3SG
'[He] wrote'.	'[He] did write'.	'[He] did not write'.

Negation of a verb is expressed in Serbian by adding to it the negative **proclitic** particle ne: *Ne mògu.* '[I] cannot'. The verb and the negative particle form a single accentual word. Lexical (= semantically full) verbs do not have special negative forms, with the following exceptions: BITI [copula/locative] and HTETI '[to] want', IMATI '[to] have' and the special imperative forms [2sg]nemoj/[2pl]nemojte 'do not'.

Negative forms of auxiliary verbs are obtained by adding the proclitic particle **ne** to the full verbal form. This is the same as with lexical verbs, except that with the auxiliaries there is a phonological interaction between the verb radical and the negative particle **ne**. More specifically, the following alternations take place:

* shift of stress from the verb to the particle **ne** in the case of the conditional auxiliary (cf. **ne bïh** ⟹ **nè bih**);

* fusion of the particle **ne** with the first (= stressed) syllable of the radical in the case of the past and the future auxiliaries. Thus, with the past auxiliary, we have ne **jèsam** ⟹ **nísam** (cf. **region. nijèsam**). The same type of fusion is seen in French 'contracted' articles; cf. **du** [⟸ de$_{PREP}$ + le$_{ART}$] and **au** [⟸ à $_{PREP}$+ le$_{ART}$].

On this account, negative auxiliary forms are secondary word-forms (see CMG-1, 205, 238 *ssq*) constructed by morphonological rules. They do not belong to any Serbian lexeme.

Another way of describing negative and affirmative verbal forms would be to postulate for the two auxiliaries the inflectional category of polarity, with grammemes {affirmative, negative}. In that case, negative and affirmative forms of auxiliary verbs should be considered as lexes of the same lexeme. The radical morpheme of BITI resp. HTETI would then be expressed, along with the grammemes of tonicity and polarity, by the following series of megamorphs:

/{BITI}, {FULL}, {AFFIRM}	⟺ **jés, jèst, bï**, etc.
{BITI}, {CL}, {AFFIRM}	⟺ **s, j, bi**
{BITI}, {NEG}	⟺ **nís, nì, nè bi**, etc.
{HTETI}, {FULL}, {AFFIRM}	⟺ **hòće, hòće**
{HTETI}, {CL}, {AFFIRM}	⟺ **će**
{HTETI}, {NEG}	⟺ **néće**

The obvious disadvantage of this solution is that the inflectional category of polarity would apply only to the two auxiliaries (and to the five lexical verbs, mentioned above).

Conclusion

The central aspects of the present paper can be summarized as follows: 1) A complete formal description of Serbian clitics has been proposed; 2) A new inflectional category for personal pronouns/auxiliaries, namely tonicity, has been introduced; 3) It has been demonstrated that the synthetic future construction does not exist in Serbian; 4) It has been pointed out that Serbian clitics feature zero radicals, a rare phenomenon in cross-linguistic terms; 5) A unified treatment of the reflexive clitic as an invariable reflexive voice marker has been proposed.

Acknowledgments

References

Browne, W. 1970. "O oblicima futura tipa čɪTAT ću i čɪTAću" (On the forms of the future tense of the type čɪTAT ću and čɪTAću), *Jezik* 2, 1970/71, Zagreb: Hrvatsko filološko društvo, 43-45.

Browne, W. 1975. "Serbo-Croatian Enclitics for English-Speaking Learners", *Kontrastivna analiza engleskog i hrvatskog ili srpskog jezika*, Zagreb: Institut za lingvistiku Filozofskog fakulteta, 105-134.

Halpern, A. 1995. *On the Placement and Morphology of Clitics*, Stanford, CA: CSLI.

Kaiser, L. (ed.). 1997. *Yale A-Morphous Linguistics Essays. Studies in the Morphosyntax of Clitics*, New Haven: Yale University.

Klavans, J. 1995. *On Clitics and Cliticization: The Interaction of Morphology, Phonology and Syntax*, New York/London: Garden Publishing Inc.

Mel'čuk, I. 1981. "Meaning-Text Models: A Recent Trend in Soviet Linguistics", *Annual Review of Anthropology* 10, 27-62.

Mel'čuk, I. 1993, 1994, 1996, 1997, 1999. *Cours de morphologie générale*, Vol. 1-5, Montréal/Paris: Les Presses de l'Université de Montréal / C.N.R.S. Éditions.

Mel'čuk, I. 1997. *Vers une linguistique Sens-Texte. Leçon inaugurale*, Paris: Collège de France.

Milićević, J. (forthcoming). „Linear Placement of Serbian Clitics: a Description within a Dependency Based Approach".

Morin, Y-C. 1979. "La morphonologie des pronoms clitiques en français populaire", *Cahiers de linguistique* 9, Québec: Les Presses de l'Université du Québec, 67-94.

Spencer, A. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*, Oxford: Blackwell.

Stevanović, M. 1986. *Savremeni srpskohrvatski jezik: gramatički sistemi i književnojezička norma*, Vol. 1-2, Beograd: Naučna knjiga.

Zwicky, A. 1977. *On Clitics*, Bloomington: Indiana University Linguistics Club.

Zwicky, A. and Pullum G. 1983. "Cliticization vs. inflection: English *n't.*", *Language* 59, 502-513.