Linear Placement of Clitics and Syntactic Dependencies (Serbian Second-Position Clitics)[1]

Jasmina Milićević

Department of French, Dalhousie University & Observatory of Meaning-Text Linguistics, Université de Montréal

Описывается линейное размещение клитик второй позиции в литературном сербском языке. Расстановка клитик рассматривается с точки зрения языкового синтеза, т.е. в направлении от смысла к тексту. Теоретической основой работы является модель *Смысл-Текст,* в первую очередь – принятые с этой модели зависимостный подход к синтаксической структуре. Предлагается набор правил для синтеза сербских предложений, содержащих клитики, исходя из их синтаксического представления. Сюда относятся 1) правила введения клитик и 2) правила линеаризации клитик; последние подразделяются на 2.1) правила образования цепочек клитик и 2.2) правила размещения цепочек клитик в морфологическом представлении предложения.

Две основных проблемы – это роль синтаксических зависимостей и роль просодических факторов в размещении клитик второй позиции. Размещение клитик рассматривается как синтаксическая операция, которая существенно, хотя и опосредованно, обусловливается синтаксическими зависимостями. Роль просодических факторов учитывается постольку, поскольку просодические характеристики составляющих влияют на их способности принимать клитики.

Basic facts about Serbian clitics

A CLITIC is a prosodically deficient wordform which does not carry stress/tone and has to depend prosodically on a full-fledged [= stressed/intoned] wordform or a phrase, called the HOST of this clitic. Serbian sentence in (1) contains three clitics—*će **im ga***—leaning on their host *(Jovan)*.

(1)   **Serb.** *Jovan*$_{host}$ *će im ga dati* lit. ⟨Jovan will to-them it give⟩ = ⟨Jovan will give it to them.⟩ Only the clitics with special linear placement properties are considered (clitic prepositions, conjunctions, etc. are left out). These clitics exhibit the following particularities: clustering, rigid ordering within the clause (even in the so-called free word-order languages) and sensitivity to prosody. Morphonologically, clitics form a prosodic unit with the host and may phonologically interact with it (allomorphic variation, sandhis, deletions, etc.), i.e., they may display behavior typical of affixes.

According to the type of the host, clitics fall into two major subclasses: SECOND-POSITION [= 2P] CLITICS, as in (1), typical of Slavic languages, and ADVERBAL CLITICS, as in (2), typical of Romance languages:

(2)   **Fr.**   *Jean va **le leur** donner*_{host} lit.⟨ Jean will it to-them give⟩ = ⟨Jovan will give it to them.⟩

Since adverbal clitics are better known than 2P clitics, we will introduce the latter by contrasting them with the former.

1. <u>Syntactic class</u>: adverbal clitics are pronouns and particles (e.g., the negative **ne** and the reflexive **se** in standard French, the interrogative **tu** in Quebec French), while 2P clitics include pronouns, particles and auxiliaries.

1. <u>Host</u>: adverbal clitics are necessarily hosted by a verb, while 2P clitics attach to an appropriate constituent[2] in the clause (most often the one in the clause-initial position).

2. <u>Syntactic governor</u>: in our framework, the host of an adverbal clitic (in the morphological structure of the sentence) always corresponds to its syntactic governor (in the syntactic structure of that sentence); in contrast, the host of a 2P clitic and its syntactic governor do not necessarily correspond—if they do, this is purely coincidental. Thus, the sequence of clitics ***mu ga je*** in (1) is hosted by the noun *Jovan,* which does not syntactically govern either of the clitics. In point of fact, this noun is governed itself by one of the clitics—the present-tense auxiliary **je** ⟨is⟩.

Basic references on clitics include Zwicky 1977, Spencer 1991: 351-393, Klavans 1995, Halpern 1995. Romance pronominal clitics are discussed, for instance, in Borer (ed.) 1986. For 2P clitics, in particular Serbian/Croatian clitics, see the classic work Browne 1974, Halpern and Zwicky (eds.) 1996 and Kaiser (ed.) 1997.

In Serbian, 2P clitics include:

1. clitic forms of personal pronouns (in the dative, genitive and accusative);

2. clitic forms of auxiliary verbs and copula/locative verbs: HTETI($_{aux}$) ⟨will⟩, BITI($_{aux}$) ⟨be⟩ and BITI (copula/locative) ⟨be⟩;

3. marker of the reflexive voice SE;

4. interrogative particle LI1 [= Q] and emphatic particle LI$_2$ [= EMPH].

All clitics except the reflexive SE and the emphatic LI$_2$ have corresponding full forms. A clitic form of a lexeme is chosen by an operation called CLITICIZATION. In Serbian, cliticization is licensed by communicative and/or syntactic factors, but I won't go into this matter here.

In order to be linearly positioned, all clitics of the same clause must be gathered in a CLITIC CLUSTER; its structure is specified by the following TEMPLATE:

1	2	3	4	5	6	7
LI	**Aux/Cop/Loc Φ j e** ⟨is⟩	**Pron$_{Dat}$**	**Pron$_{Acc}$**	**Pron$_{Gen}$**	**Adjunct$_{Refl}$**	**Aux/Cop/Loc =** **je**

The clitic cluster cannot be interrupted by a non-clitic element.

(3) a.   *Sećate*   ***li#**1 **ga se**?*   lit.⟨ [You$_{PL}$] remember Q it REFL⟩⟨Do you remember him?⟩
             1   4   6

     b.   *Kad*   ***mu ga je*** *dao?* lit. When **to-him it is** [he] having-given?⟩ = When did he give it to him?⟩
             3   4   7

The cluster must be linearly positioned after an appropriate constituent C of the clause, i.e. after a C that can or must host the clitics. The hosting C may be of any syntactic class (cf. (1) and (3)), but must satisfy several requirements concerning its Synt-role/composition, its communicative

---

[2] I will rely on an intuitive understanding of the notion of constituent until p. 0, where it will be characterized. For now, note that constituents in the Maning-Texx approach, although they 'physically' most often correspond to those used in phrase-structure approaches to syntax, are conceptually quite different.

markedness (e.g. contrastiveness), and the number of stresses it carries [= prosodic heaviness]. According to these .requirements, Cs are (POSSIBLE) HOSTS [= can host the clitics] and NON-HOSTS [= cannot host the clitics]. Possible hosts are further subdivided, along two independent axes, into SKIPPABLE HOSTS [= can be overlooked when choosing the landing site for the clitic cluster] ~ NON-SKIPPABLE HOSTS, and INSERTABLE HOSTS [= admit the insertion of the cluster into them; the inserted cluster is placed after the first stressed wordform of the constituent] ~ NON-INSERTABLE HOSTS.

(4) a. [*Juče*]_{host, non-skip} ***me*** ***je*** [*posetio*] [*moj prijatelj Marko*].
          yesterday   me   is_(aux) visited my friend Marko
          (My friend Marko visited me yesterday).
          vs.
          *[*Juče*]_{host, non-skip} [*posetio*] **me je** [*moj prijatelj Marko*].
     b. [*Posetio*]_{host, non-skip} ***me je*** [*juče*] [*moj prijatelj Marko*].
          vs.
          *[*Posetio*]_{host, non-skip} [*juče*] ***me je*** [*mojprijatelj Marko*].

Skippable hosts
(5) a. [*Moj prijatelj Marko*]_{host, skip} ***me je*** [*posetio*] [*juče*].
          or
     b. [*Moj prijatelj Marko*]_{host, skip} [*posetio*] ***me je*** [*juče*].

Non-hosts
(6) a. [*Moj prijatelj, Marko,*]_{non-host} [*posetio*] ***me je*** [*juče*].
          (My friend, Marko, visited me yesterday).
          vs.
     b. *[*Moj prijatelj, Marko,*]_{non-host} **me je** [*posetio*] [*juče*].

A non-host is obligatorily skipped; (6b) is ungrammatical because the initial constituent, which contains a non-restrictive apposition, is a non-host. (The initial constituent in (5a-b) contains a restrictive apposition and is an optional host.)

(7) a. [*Mo* ***me je***   *prijatelj* *Marko*]_{host, skip, insert} *posetio juče*.
          My   me is_(aux)   friend   Marko
          (My friend Marko visited me yesterday).
          vs.
     b. *[*Prijatelj* ***me je***   *Marko*] _{host, skip, non-insert} *posetio juče*.
          Friend   me is_(aux) Marko
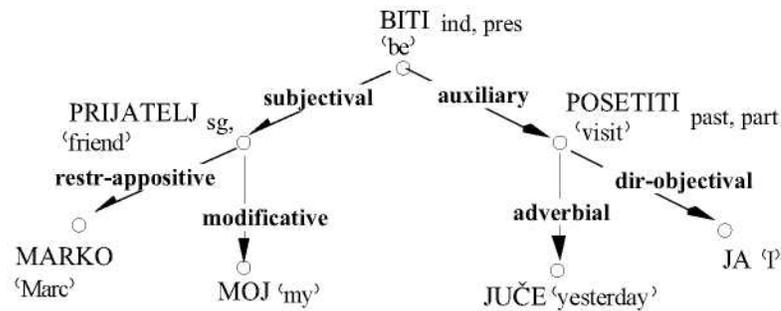
Serbian clitics in the Meaning-Text framework

Meaning-Text Theory uses functional models of languages, called *Meaning-Text models*. Such a model presupposes several levels of representation of sentences and consists of sets of rules [= modules] which establish correspondences between adjacent levels of representation.

Clitic placement rules are part of the Surface-Syntactic module, which maps a Surface-Syntactic Representation [= SSyntR] of a clause to its Deep-Morphological Representation [= DMorphR]. The basic structure of the SSyntR is an unordered labeled dependency tree, and that of the DMorphR is an ordered string of wordforms—or, more precisely, constituents—including the clitics.

A constituent is, roughly speaking, an ordered string of wordforms which is a projection of a complete SSynt-subtree and behaves as a whole from the viewpoint of linearization and prosodization (it is 'moved around' and prosodized as one block). In MTT framework, constituents are used not to represent syntactic structure, as in phrase-structure approaches, but to represent the way in which language expresses it. They are computed from the SSyntR and used as building blocks for the construction of the DMorphR; they appear only in the latter. (On the notion of constituent in the Meaning-Text framework, see Gerdes & Kahane [to appear].)

Here are the SSyntS and the DMorphS of sentence (5a).

BITI ind, pres
'be'

PRIJATELJ sg,
'friend'

subjectival   auxiliary   POSETITI past, part
'visit'

restr-appositive   modificative   adverbial   dir-objectival

MARKO
'Marc'   MOJ 'my'   JUČE 'yesterday'   JA 'I'

This is also the SSyntS of sentences (4a-b), (5b), and (7a). Different word orders in these sentences are the result of communicative differences (oppositions such as Theme/Rheme, Given/New, etc.), which won't be looked into here.

[MOJ$_{masc,sg,acc}$ PRIJATELJ$_{sg, acc}$ MARKO$_{sg, acc}$] **JA$_{cl, acc}$** **BITI$_{cl,pres,ind,3, sg}$** [POSETITI$_{part,masc,sg}$] [JUČE].
The clitic cluster is positioned after the first constituent (constituents are in square brackets). Linear placement of 2P clitics is not determined only by Synt-dependencies and the communicative structure as it is for full-fledged words, where a dependent member of a SSynt-relation is placed with respect to its governor (and perhaps its co-dependents). In clitic placement, other factors are at play—in particular, prosodic properties of sentence elements (~ constituents). Thus, such prosodic properties of a constituent as its heaviness and its capacity to induce a pause after itself are relevant to the identification of clitic-hosting constituents and must be accounted for in the clitic placement rules.

The SSynt-module includes: 1) morphologization rules, further divided into government rules (e.g. case assignment), agreement rules and cliticization rules; 2) linearization rules; 3) prosodization rules.

Only special linearization rules for clitics will be discussed, leaving general linearization, cliticization and prosodization aside. Note, however, that the general linearization rules have to apply first, as the clitic cluster must follow a specific constituent in the partial DMorphS; this means that constituents are built and arranged before the clitic placement rules apply.

Clitic linearization rules

These rules are subdivided into CLITIC-CLUSTER BUILDING RULES and CLITIC-CLUSTER POSITIONING RULES.

Clitic-cluster building rules assign the position in the cluster template (cf. p. 0) to all clitics in the SSyntS. E.g.:

$$\text{BITI}_{clitic, \text{ not [pres, 3p, sg]}} \rightarrow \_2$$

All clitic forms of BITI 'be'_except 3sg [= **je** 'is'] are assigned position 2.

$$\text{BITI}_{clitic, pres, 3p, sg} \rightarrow \_7$$

The 3sg clitic form of BITI ' [to] be'_ is assigned position 7.

Clitic-cluster positioning rules involve identification of potential hosting constituents and the choice of the actual host.

a) Host-identifying rules

apply to a sequence of constituents (~ partial DMorphS of the clause) marked by the features [±**heavy**], [±**detached**], [±**contrastive**], [±**autonomous**] and identify each constituent as a skippable host, a non-skippable host or a non-host. The above features represent (partial) generalizations over prosodic, syntactic and communicative properties of constituents. The feature [+**heavy**] is assigned to a constituent C containing at least two stressed wordforms; in most cases, this makes it into a skippable host, cf. (5). The feature [+**detached**] is assigned to a C in the SSynt-role of an address, a parenthetical, etc., or containing a non-restrictive apposition, etc.; such a C, which is followed by

an obligatory pause, is a non-host, cf. the initial C in (6). The feature [+**contrastive**] is assigned to a C which is focalized, emphasized, etc.; such a C carries a particular contour and is, in most cases, a skippable host, cf. (8a). The feature **[+autonomous]** is assigned to a communicatively prominent C, separated from the rest of the clause by a pause (a sentential adverb, an address, etc.); such a C is a non-host, cf. (8b). A C which is neither of these, is, again, in most cases, a non-skippable host, cf. (4).

(8)   a.   [*NAGRADU*]**[+contrastive]** *dobio je Marko.*

        award got is [aux] Marko
        ⁽It is an award that Marko got⁾.

    b.   **[Draga Marija,]*[+**automomous**] *te volim.*
        ⁽Dear Marija, [I] love you⁾.

In some cases, however, these features are not sufficient: a syntactic property of a C can 'override' them. Thus, a contrastive C is normally a skippable host, but if it is a SSynt-Subject, it becomes a non-skippable host, and so does a contrastive verb; cf., respectively, (9a) vs. (8a), as well as (9b).

(9)   a.   **[MARKO]***Subj[+contrastive]** *nagradu je dobio.*
        ⁽Marko is the one who got the award⁾.

    b.   **[IZNAJMILI]***Verb[+contrastive]** *kuçu su (a ne kupili).*
        ⁽They RENTED the house, rather than bought it⁾.

A subset of (possible) hosts have to be additionally marked for insertability: 1) obligatory insertion, where the SSynt-role of the C is relevant, and 2) optional insertion, where the SSynt-composition of the C is at play.

Obligatory insertion concerns only the clusters containing the copula BITI ⁽be⁾ and the Cs implementing the predicative-attributive construction, cf. (10a-b), while optional insertion concerns all clusters and is possible with the Cs of the form 'Premodifier+X', cf. (7).

(10) a.   Sasvim   mi   je   jasan]**PredAttr**[*njegov stav.*]**Subj**
        completely to-meis[copula]   clear   his position
        (His position is completely clear to me).

    b.   **[Sasvim jasan]***PredAttr** *mi je* [*njegov stav.*]**Subj**
        cf. [*Njegov stav*]**Subj** *mi je* [*sasvim jasan.*]**PredAttr**

The result of the application of host-identifying rules to the partial DMorphS of (5a) follows.

(11)   [MOJ PRIJATELJ MARKO]**host,skip, insert** [POSETITI]**host, non-skip, non-insert** [JUČE]**host, non-skip, non-insert**

In (11), there are four logically possible landing sites for the clitic cluster, one of which—after JUČE— must be discarded by the next set of rules.

b) Clitic cluster placement rules

Basic rule

> *The clitic cluster is placed after the first non-skippable C of its clause or after any possible hosting C which precedes it.*

This rule discards JUSE as a possible host so that (11) yields three sentences—(5a), (5b) and (7a).

No-Clause-Final-Position Rule

> *The cluster should not be placed in the clause-final position if a non-final position is available.*

(12) a.   **[Moj prijatelj Marko]***host,skip,insert** [*posetio*] *me je*. lit. ⁽My friend Marko visited me is⁾
        Cf. a good sentence (5a).

    b.   [*Moj prijatelj, Marko,*]**non-host** [*posetio*] *me je*.

Sentence (12b) is grammatical because the only available position for the clitics is clause-final.

Insertion rules

5. *A cluster containing a copula clitic and positioned after an insertable C in the SSynt-role of the predicative attribute MUST be inserted into C.*

6. *A cluster positioned after an insertable C CAN be inserted into C.*

c) Preference rules for the placement of the clitic cluster

select the actual landing site for the cluster if two or more possible sites have been identified. They take care of optional skipping and optional insertion; only the first operation will be characterized here.

Recall that skippable Cs are those marked as [+**heavy**] or [+**contrastive**].

Heaviness of a C is a gradient: the more stressed wordforms a C contains, the heavier it is. Thus, [MOJ PRIJATELJ] ʿmy friendʾ is much less heavy than [ONIMA KOJI BI ŽELELI DA SE UPOZNAJU SA DOSTIGNUČIMA GENERATIVNE GRAMATIKE]ʿto those who would like to get acquainted with the achievements of generative grammarʾ.

　　The heavier a C is, the more preferable it is to skip it.

For the second C above, the skipping preference is very high, verging on obligatoriness. Optional skipping must not interfere with the **No-Clause-Final-Position Rule** above. There are instances of multiple skipping; cf.:

(13)　a.　[*Ideje Čomskog*]$_{host,skip,insert}$ [*dalju razradu*]$_{host,skip}$
　　　　　ideas of-Chomsky　　　　　　further elaboration

　　　　　[*doživele*]$_{host,non-skip}$　***su***　　***u***
　　　　　having-experienced　are$_{[aux]}$　in
　　　　　ʿChomsky's ideas were further elaborated in ...ʾ

　　　b.　[*U tim okolnostima*]$_{host,skip, insert}$ [*ideje Čomskog*]$_{host,skip, insert}$
　　　　　in these circumstances

　　　　　[*dalju* razradu]$_{host,skip, insert}$ [*doživele*]$_{host,non-skip}$ ***su u***

In (13a), the number of potential landing sites is four, and in (13b) it is six. Here, the preferences seem to be more or less equal, i.e. all these sentences are equally good.

As for contrastive Cs, skipping preferences are less clear for now, as several interacting factors (context, style and global word-order) may impede or favor the skipping; a more advanced study is required.

Conclusion

Serbian clitics are positioned not with respect to their SSynt-governors but after an appropriate C of their clause, regardless of whether there is a Synt-relation between this C and the clitics. This must be so, since the clitics are positioned together, as a cluster, while in the SSynt-structure they have different governors or are governors themselves (an auxiliary is the SSynt-head of its clause). Synt-dependencies are thus not directly relevant to 2P clitic placement—they are not explicitly mentioned in clitic-cluster building/positioning rules. Indirectly, however, they are relevant, since the Cs and their properties, necessary for the placement of the clitic cluster, are computed mostly on the basis of dependencies.

A question hotly debated in literature concerns the role of syntactic vs. prosodic factors in the placement of 2P (in particular Serbian/Croatian) clitics. Syntactic approach is adopted, for instance, in Halpern 1995 and Progovac 1996, while a prosodic one is argued for in Radanovič-Kocič 1996. Here, clitic placement is treated as a syntactic operation, the rationale being that prosody is determined by the Synt- and the Synt-communicative structures, not the other way around: a pause after a parenthetical is a marker of its Synt-role, a particular contour of a constituent is a marker of its communicative role, etc. True, many different syntactic/communicative underlying factors can be

'reduced' to the same prosodic expression, and this allows for a better generalization: e.g., a C which induces a pause after it cannot host the clitics, irrespective of the syntactic/communicative source of the pause; a C with a contrastive contour may but need not host the clitics; etc. However, it is impossible to formulate clitic placement rules in terms of prosody alone, since in a few cases syntactic factors must be mentioned, cf. examples in (9), showing that syntactic role of a C is relevant for determining its hosting capacity. Moreover, the final phonological phrasing must take into account the clitic cluster itself.

References

1. Borer. H (ed.) 1986. The Syntax of Pronominal Clitics. Syntax and Semantics, vol. 19. Orlando/San Diego: Academic Press.

2. Browne, W. (1974). On the Problem of Enclitic Placement in Serbo-Croatian. In: Brecht, R. Et Chvany, C. (eds.). Slavic Transformational Syntax. Michigan Slavic Materials, vol. 10: 36-52. Ann Arbor.

3. Gerdes, K. and Kahane S. (to appear). Phrasing it differently. In: L. Wanner (ed.). Recent works in the Meaning-Text theory in honor of L Mel'čuk.

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

5. Halpern, A. and Zwicky, A. (eds.) 1996. Approaching Second: Second Position Clitics and Related Phenomena. Stanford, CA: CSLI.

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

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

8. Mel'čuk, I. (1974). Opyt teorii lingvističeskix modelej Smysl <r^ Tekst. Moskva: Nauka.

9. Mel'čuk, I. (1988). Dependency Syntax: Theory and Practice. Albany: State University of New York Press.

10. Progovac, Lj. (1996). Clitics in Serbian/Croatian: Comp as the Second Position. In: Halpern & Zwicky 1996:411-428.

11. Radanovič-Kocič, V. (1996). The Placement of Serbo-Croatian Clitics: A Prosodic Approach. In: Halpern & Zwicky 1996: 429-446.

12. Tesniére, L. (1959). Éléments de syntaxe structurale. Paris: Klincksieck.

13. Zwicky, A. (1977). On Clitics. Bloomington: Indiana University Linguistics Club.