On Cliticization in Croatian: Syntax or Prosody?

Damir Ćavar

BBAW & Uni Potsdam

Abstract

In the following paper it will be argued that the phonological approach to clitic place-ment in Serbian/Croatian, as proposed in Zec & Inkelas (1990), not only fails to explain the observed phenomena, but also fails at the level of descriptive adequacy.

Further arguments are presented against accounts which claim that clitic placement is syntactic and which utilize a post-syntactic operation of *Prosodic Inversion* (PI) in order to explain certain cases of apparent split of syntactic constituents (Halpern, 1992; Schütze, 1994).

It will be argued that an alternative analysis which assumes syntactic clitic place-ment as proposed in Wilder & Ćavar (1994) and Ćavar & Wilder (1994) appears to be descriptively adequate.

1 Properties of Clitics in Croatian*

Table (1) gives a brief overview of different enclitic and proclitic categories in Croatian:

(1)

		forms		cliticization
		full	reduced	direction
pronouns	ACC	fem: *njû*	*ju/njú*	$\leftarrow$
		msc: *njega*	*ga/nj*	$\leftarrow$
auxiliaries	pos	*jesam*	*sam*	$\leftarrow$
	neg	*nisam*		
prepositions		*nâ*	*ná*	$\rightarrow$

The enclitic forms of the accusative pronoun for masculine and feminine differ depending on the syntactic context in which they appear. While the forms *njú* and *nj* are only licensed as complements of prepositions, i.e. only appear in prepositional phrases and only cliticize to a non-clitic (e.g. mono-syllabic and bi-moraic) preposition, the other forms *ju* and *ga* may only appear elsewhere.[1]

Another category that appears either as a full or as an enclitic form, is auxiliaries. Auxil-iaries have two full forms, an affirmative and a negative form. The enclitic auxiliary appears in neutral contexts, i.e. neither emphatic nor negative contexts.

Prepositions may be realized as independent words, if they are stressed (when they are bi-moraic, i.e. have a long vowel, for example in *nâ*), or proclitic (when they are mono-moraic, i.e. have a short vowel, for example in *ná*).

With the exception of prepositions and enclitic pronouns that only appear in PPs, all the other clitics, the reflexive pronouns and the question marker *li* form a morphological unit in which (apparently) the individual elements appear in fixed positions:

---

[1]Several informants consider the reduced enclitic pronoun in PPs archaic (e.g. Nedzad Leko (p.c.)), or even unacceptable (e.g. native speakers of Serbian), while other dialects (e.g. in Dalmatia and Hercegovina), or other Slavic languages make more or less extensive use of the two different enclitic forms.

(2)   **The Clitic Cluster** (Spencer, 1991: 356):

      li – AUX – DAT – ACC – Refl. – *je*

There is a tendency among syntacticians to assume that the apparent order in the clitic cluster is the result of syntactic operations or constraints. However, as mentioned in Wilder & Ćavar (1994), and argued in Ćavar (1996), the order of the clitics is not strictly fixed for all native speakers and for all dialects of Serbian/Croatian. Certain clitics may invert with others, depending on their morpho-phonological shape. Therefore it is assumed here that the order in the cluster has to be explained in terms of morphology rather than syntax. The only position that seems to be fixed across dialects and idiolects, is the initial position of the question marker *li*. The auxiliary clitic *je* is preferred in final position, and is usually dropped, if the reflexive pronoun *se* appears in the cluster.[2]

1.1   The 'Tobler-Mussafia-Effect'

The basic property of enclitic elements in Slavic shows up in the so called 'Tobler-Mussafia-Effect' (TM-effect hereafter), namely, that enclitic elements may not appear in string initial position.[3] The examples in (3) and (4) show this for matrix clauses, i.e. the enclitic auxiliary *sam* in (3c), and the enclitic pronoun *me* in (3d) may not appear in absolute string initial position, while their full form counterparts in (3e) and (3f) may:

(3)   a.   *Spavao **sam** čitavi dan.*   b.   *Nije   **me** probudila.*
        sleep$_{ptc}$ be$_{1sg}$ whole day   NEG-be$_{3sg}$ me wake-up$_{ptc}$
        'I slept the whole day.'   'She didn't wake me up.'

    c.   * ***Sam** spavao čitavi dan.*   d.   * ***Me** nije probudila.*

    e.   ***Jesam** spavao čitavi dan.*   f.   ***Mene** nije probudila.*

As argued in Wilder & Ćavar (1994), the same condition holds in embedded contexts. As the examples in (4) show, the TM-effect can be found in embedded contexts as well, i.e. the clitic cluster, which contains the enclitic pronoun *me* and the enclitic auxiliary *je* in (4a) may neither appear in some relative string initial position in embedded finite clauses, cf. (4a), nor in initial position in N-selected infinitive clauses, cf. (4c-d):[4]

(4)   a.   *Senka tvrdi   [$_{CP}$ da   **me je**   probudila]*
        S.   claim$_{3sg}$ that me be$_{3sg}$ wake-up$_{ptc}$
        'Senka claims that she woke me up.'

    b.   * *Senka tvrdi [$_{CP}$ **me je** da probudila]*

---

[2]Note that the full form and the enclitic form of the third person singular auxiliary differ only in vowel length: *jé* vs. *jê*, i.e. the first is a mono-moraic, the second a bi-moraic syllable. On the other hand, the other clitics have a special morphological shape which differs from the full form. One could take *je* to be a simple clitic in terms of Zwicky & Pullum (1983), or Zwicky (1985), while the other enclitic elements are all special clitics, hence the right peripheral position of *je* could be the result of simple cliticization.

[3]There may be some differences with respect to (en)clitic reflexives in Czech, because these apparently appear in string initial position in some contexts (as pointed out in Toman (1993) and by Tobias Scheer (p.c.)). There is no evidence that this is true for enclitic auxiliaries and other pronouns, while there is enough evidence for a special behaviour of reflexive pronouns in other Slavic languages, Polish, as well as Croatian. However, the discussion of these phenomena would go beyond the scope of this paper.

[4]Note that verb-selected infinitival clauses in Croatian are transparent for clitic climbing (see section 1.3), while finite complements are not, i.e. (4b) would be ungrammatical even if the clitics would move to the absolute second position.

c. *Imam mogućnost* [IP *upoznati ga*]
   have_{1sg} possibility get-to-know_{inf} him
   'I have the possibility to get to know him.'

d. * *Imam mogućnost* [IP ***ga** upoznati*]

1.2 The 'Wackernagel Effect'

Another phenomenon observed for enclitic elements in Croatian is the so called 'Wackernagel Effect' (W-effect hereafter) (Wackernagel, 1892): the clitic cluster may not appear in a position deeper than second position in the clause.[5]

The examples (5a-c) show that the enclitic auxiliary *je* may not appear in some absolute third (8b), or absolute forth position (8c) in a clause that contains a fronted *wh*-element, i.e. following a fronted *wh*-word and the subject (8b), or a *wh*-word, the subject, and the participle (8c):

(5) a. *Šta **je** Ivan radio čitavi dan?*
        what be_{3sg} I.   do_{ptc} whole day
        'What did Ivan do the whole day?'

    b. * *Šta Ivan **je** radio čitavi dan?*

    c. * *Šta Ivan radio **je**?*
        what I.   do_{ptc} be_{3sg}
        'What did Ivan do?'

The same holds for sentences with topicalized elements. In (6a) the clitic cluster that contains the enclitic auxiliary *sam* and the enclitic reflexive pronoun *se* may directly follow the topicalized adverb *danas*, but not both the adverb and the participle (6b):[6]

(6) a. *Danas **sam se** naspavao.*   b. * *Danas naspavao **sam se**.*
        today be_{1sg} self have-a-good-sleep
        'Today I had a good sleep.'

One could argue that the ungrammaticality of examples like (6b) results from the final position of the clitic.[7] However the examples in (7) show that a construction in which the clitic appears in third position, following two topicalized phrases, is ungrammatical (7c), even if the clitic is not in absolute final position in the clause:

(7) a. *Stipi **su** Ivan i Marija sinoć dali knjigu.*
        S.   be_{3pl} I. and M.   yesterday give_{ptc} book
        'Ivan and Mary gave a book to Stipe yesterday.'

    b. *Sinoć **su** Ivan i Marija Stipi dali knjigu.*

    c. * *Stipi sinoć **su** Ivan i Marija dali knjigu.*

---

[5]In Wilder & Ćavar (1994) it is argued that one has to differentiate between the TM-effect and the W-effect in an analysis of cliticization in Croatian. While the former has to be explained in phonological terms, the later has to be explained in syntactic terms. This will be discussed in more detail in section 5.

[6]These cases of topicalization differ from the topicalization constructions discussed in Zec & Inkelas (1990). This is discussed in more detail in section 5.1.

[7]See Wilder & Ćavar (1994) for a discussion of such cases.

As argued in Wilder & Ćavar (1994), this placement constraint holds also for relative third position, i.e. enclitic elements may not appear in a position other than second position in embedded clauses. The examples in (8) show that the clitic cluster containing the enclitic pronoun *me* and the enclitic auxiliary *je* may not appear in third position inside a finite sentential complement, neither in absolute final position (8a), nor in some intermediate position (8b):

(8)　　a.　* *Netko tvrdi [da probudila **me je**]*　　　(compare with (4a-b))

　　　　b.　* *Netko*　　*tvrdi [da*　*probudila*　***me je***　*u dva sata]*
　　　　　　somebody　claims that wake-up$_{ptc}$ me　be$_{3sg}$ at two o'clock
　　　　　　'Somebody claims that she woke me up at two o'clock.'

The same condition holds for clitics in noun selected infinitives as the examples in (9) show:

(9)　　a.　*Želja [Mariji dati　ružu] bila　je　velika.*
　　　　　　wish　M.　　give$_{inf}$ rose　be$_{ptc}$ be$_{3sg}$ great
　　　　　　'The wish to give Mary a rose was great.'

　　　　b.　*Želja [Mariji **ju** dati]　bila　je　velika.*
　　　　　　wish　M.　　it give$_{inf}$ be$_{ptc}$ be$_{3sg}$ great

　　　　c.　* *Želja [Mariji dati　**ju**] bila　je　velika.*
　　　　　　wish　M.　　give$_{inf}$ it　be$_{ptc}$ be$_{3sg}$ great

The enclitic pronoun *ju* may not appear in some position deeper than second position in the N-selected infinitive clause, compare (9b) with (9c).

1.3　Clitic Climbing

While clitics cannot raise out of finite sentential complements or noun selected infinitives, the examples in (10) show that they may raise out of infinitive complements into the matrix clause:

(10)　a.　*Ivan **je**　želio [$_{IP}$ čitati　Krležu]*
　　　　　　I.　be$_{3sg}$ wish$_{ptc}$ read$_{inf}$ K.
　　　　　　'Ivan wanted to read Krleža.'

　　　　b.　*Ivan **ga je**　　želio [$_{IP}$ čitati]*
　　　　　　I.　　him be$_{3sg}$ wish$_{ptc}$　read$_{inf}$

　　　　c.　*Želio　**ga je**　Ivan [$_{IP}$ čitati]*
　　　　　　wish$_{ptc}$ him be$_{3sg}$ I.　　　read$_{inf}$

In fact, the clitics have to move out of the infinitive, as the examples in (11) show:

(11)　a.　* *Ivan **je**　želio [$_{IP}$ čitati　**ga**]*
　　　　　　I.　be$_{3sg}$ wish$_{ptc}$ read$_{inf}$ it

　　　　b.　* *Ivan **je** želio [$_{IP}$ **ga** čitati]*

Whether the enclitic pronoun *ga* appears in post-verbal second position in the infinitive clause as in (11a), or in preverbal initial position in the infinitive as in (11b), it doesn't change the fact that the examples are ungrammatical.

2   The Phonological Analysis

Zec & Inkelas (1990) observe that clitics in Serbian/Croatian apparently may split syntactic constituents.

(12) a. *Taj čovjek **joj ga je**   poklonio.*
          that man   her it   be$_{3sg}$ present$_{ptc}$
          'That man presented her with it.'

    b. [*Taj **joj ga je**   čovjek*] *poklonio.*
          that her it   be$_{3sg}$ man   present$_{ptc}$

The auxiliary clitic *joj* may follow an initial constituent, i.e. a complex DP *taj čovjek* in (12a), but it may also appear inside the DP, after the first word, i.e. the demonstrative *taj* in (12b).

    However this splitting of constituents is not possible if the only element that precedes the clitic is a preposition (Zec & Inkelas 1990: 367) as in (13c) below.

(13) a. *Petar **je**   u kući.*
          P.   be$_{3sg}$ in house
          'Petar is in the house'

    b. *U kući **je** Petar.*   c. * *U **je**   kući Petar.*
          in house be$_{3sg}$ P.   in be$_{3sg}$ house P.

While the enclitic auxiliary *je* may appear directly after the fronted PP in (13b), it is not possible for the clitic to split this PP and occupy a position immediately following the initial preposition in (13c).

    Zec & Inkelas (1990) offer an explanation for these facts in terms of phonology. The basic assumption is that there is a fundamental difference between the phonological properties of functional words (closed class elements) and substantives (open class elements). While substantives bear inherent word accent (High tone and pitch accent), functional words do not. Hence, the claim is that open class elements are always phonological words, while functional words can be phonological words only if they are accented.

    As the following examples show, certain conjunctions indeed may host clitics, if they bear High tone and accent (Zec & Inkelas 1990: 368):

(14) a. *Mi smo zvonili, ali nitko   **nam** nije   otvorio.*
          we be$_{1pl}$ ring$_{ptc}$ but nobody us   NEG-be$_{3sg}$ open$_{ptc}$
          'We rang but nobody opened the door for us.'

    b. *Mi smo zvonili, ali **nam** nitko   nije   otvorio.*
          we be$_{1pl}$ ring$_{ptc}$ but us   nobody NEG-be$_{3sg}$ open$_{ptc}$

The basic assumption with respect to clitic placement, as formulated in Zec & Inkelas (1990) is that the distribution of clitics is prosodically restricted, i.e. word order in Serbian/Croatian is subject to prosodic constraints.

    The explanation for the distributional properties of clitics in Serbian/Croatian is given in terms of prosodic properties of the clitics themselves. It is assumed that the prosodic subcategorization frame in (15) is the lexical specification of these enclitic elements (here given for the enclitic auxiliary *je* (3rd sg. 'to be')):

(15) *je*:   [[   ]$_w$ ——]$_w$

    Since Zec & Inkelas (1990) claim that the preposition in Serbian/Croatian is never a phonological word, cliticization to a preposition like in (13c) is ungrammatical, because the subcategorization frame in (15) is not saturated (at a certain level).

2.1 Problems and Consequences

On the one hand, it is not quite clear, what the process of clitic placement is under this analysis. Since placement of the clitic after a complex initial syntactic phrase that contains several phonological words, is possible (cf. (7a) and (12a)), Zec & Inkelas (1990) have to assume that (15) is a (probably syntactic) constraint on clitic placement. On the other hand, since Zec & Inkelas (1990) seem to assume that clitic placement is a syntactic operation, it is not clear, how syntactic operations could access purely phonological information of the categories involved, i.e. what is 'a phonological word in syntax'. One could probably think of (15) as a PF-filter that excludes representations with initial clitics, or with clitics following a constituent which is not a phonological word (cf. Vogel & Kenesi, 1990).

However, while this analysis offers a possible explanation for the TM-effect with respect to clitics in Serbian/Croatian, the major problem for such an approach is the fact that clitic placement underlies for example the Wackernagel-constraint, i.e. clitics may not appear in a position deeper than second position in embedded finite clauses (8) and follow always the first fronted *wh*-element, cf. (5a) vs. (5b).

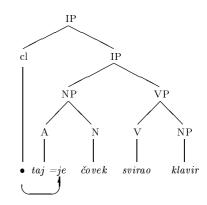# 3 The Syntactic-Phonological Solution

An attempt to analyse the described phenomena and avoid the problems mentioned above with respect to constraints on clitic placement, is presented in Halpern (1992), and Schütze (1994). In addition to adopting the assumption of Zec & Inkelas (1990) in (15), Halpern (1992) assumes that clitics occupy a fixed syntactic position.

The distinction between placement after the first phonological word (1W hereafter), and after the first syntactic constituent (1C hereafter) is explained in the following way: 1C appears after fronted constituents (topicalization, *wh*-movement) and is due to the fact that clitics are adjoined to IP, while the fronted XPs end up in some CP-projection, preceding the clitics.[8] On the other hand, 1W results from the Last Resort operation *Prosodic Inversion*, that inverts two adjacent prosodic entities, i.e. a clitic with the following (or preceding) phonological word, if and only if the subcategorization frame (15) is not fulfilled at some level on the way to PF. It is assumed that PI is operative after syntax, on the way to PF, i.e. clitics may move (after syntax).

The following example (16b) shows the PI-analysis for examples like (16a) in Serbian/-Croatian:

(16) a.   *Taj **je**   čovek svirao klavir.*   b.
           that be$_{3sg}$ man   play$_{ptc}$ piano
           'That man played the piano.'

The enclitic auxiliary *je* which is adjoined to IP (or to C^0 in Schütze (1994)), inverts at PF with the following phonological word.

Halpern (1992) defines PI as follows:

---

(17) **Prosodic Inversion** (Halpern, 1992: 81)

Prosodic adjunction of clitics: For a DCL X, which must attach to a w to its left (respectively right)

   i.   if there is a $w = Y$, comprised of material which is syntactically immediately to the left (right) of X, then adjoin X to the right (left) of Y,

   ii. else attach X to the right (left) edge of the w composed of syntactic material immediately to its right (left).

In other words, after syntax (on the way to PF) any directional clitic (DCL) (enclitic or proclitic), namely, any element that contains (15) as part of its lexical specification, i.e. requires a phonological word w to its left (or right), may invert with a phonological word immediately to its right (or left), only if there is no phonological word preceding (or following) it.

3.1 Problems and Consequences

The PI-account offers a possible explanation, and makes clear predictions with respect to the examples (12) discussed in the previous section.

However, one problem for this analysis arises, when we consider examples with scrambled constituents in finite complement clauses, as in (18). One might assume that in (18b) the direct object *Krležu* is scrambled to VP, and in (18c) to IP.

(18) a.   *Ivan kaže   da   Marija čita   Krležu.*
           I.   say$_{3sg}$ that M.   read$_{3sg}$ K.
           'Ivan says that Mary reads Krleza.'

      b.   *... da Marija Krležu čita*

      c.   *... da Krležu Marija čita*

If the assumption is that clitics are adjoined to IP, and if embedded finite clauses that contain clitics are taken under consideration, it can be observed that the clitics have to be always the highest adjuncts to IP, cf. (19b) vs. (19c):

(19) a.   *Ivan kaže   da   **mu je**   Marija dala   knjigu.*
           I.   say$_{3sg}$ that him be$_{3sg}$ M.   give$_{ptc}$ book
           'Ivan said that Mary gave him the book.'

      b.   *... da* [$_{IP}$ ***mu je*** [$_{IP}$ *knjigu* [$_{IP}$ *Marija dala*]]]

      c. ** ... da* [$_{IP}$ *knjigu* [$_{IP}$ ***mu je*** [$_{IP}$ *Marija dala*]]]

The fact that the W-effect appears in embedded clauses, does not follow from the assumptions in Halpern (1992).[9] On the contrary, this analysis massively under- and over-generates. The relevant data is discussed in the following section.

4 Problems for Phonological Accounts

4.1 Splitting Complex XPs

The PI-account analyses splitting of constituents in examples like (20) as inversion of the enclitic element with the first phonological word immediately to its left. This operation is

---

[9]As already mentioned, Schütze (1994) assumes that clitics occupy the C^0 position, thus solving this inconsistency.

assumed to be a Last Resort operation, because, first, it is only licensed if the clitic appears in string initial position at some post-syntactic level, therefore the subcategorization frame (15) would be violated, and second, the inversion is restricted, namely, only one phonological word may invert with the clitic. Thus the PI-account makes the strong prediction that clitics may only appear after the first phonological word, as in (20b), assuming an underlying representation as in (20a):[10]

(20)   a.   [$_{IP}$ ***su*** [$_{IP}$ *u stara raspala prljava kola* [$_{IP}$ *Marija i Ivan sjeli*]]]
             be$_{3pl}$ in old   rotten   dirty   car   M. and I.   sit$_{ptc}$

       b.   [$_{IP}$ __ [$_{IP}$ *U stara ***su*** raspala prljava kola*] *Marija i Ivan sjeli*]]

However, it is possible for clitics to appear in apparent third or fourth position inside a complex phrase. The following examples show that the enclitic auxiliary *smo* may appear in second position in the clause, when preceded by the complex PP like in (21a), but it may also appear 'inside' the complex PP, following the third phonological word like in (21b):[11]

(21)   a.   [$_{PP}$ *U stara raspala prljava kola*] ***smo*** *sjeli*
             in old   rotten   dirty   car   be$_{1pl}$ sit$_{ptc}$
             'We sat into an old dirty rotten car.'

       b.   [$_{PP}$ *U stara raspala prljava ***smo*** kola*] *sjeli*
             in old   rotten   dirty   be$_{1pl}$ car   sit$_{ptc}$

Since such constructions are neither marked, nor seldom, we may conclude that the PI-account not only fails to offer an explanation for the observed effects, but in fact undergenerates.

It is clear that clitic placement in (21) neither takes place after the first syntactic constituent, nor after the first phonological word. Hence, this data seems to be problematic for both, a phonological and a syntactic clitic placement analysis.

However, the properties of prepositions and split PPs have to be examined in more detail, before an alternative analysis is taken into consideration.

4.2   The Properties of Prepositions

As already mentioned above, in Zec & Inkelas (1990) it is assumed that functional words do not have independent High tone and accent, and therefore are not phonological words, and, therefore, cannot host clitics. Furthermore, it was assumed that certain functional words may bear High tone and accent and function as hosts for clitics, while prepositions may not.

However, as described in traditional grammar books of Croatian, prepositions may be either proclitic or, if accented (stressed), morphologically and phonologically independent (cf. Barić et al, 1990).

While the proclitic version of e.g. the preposition *na* ('on') is mono-syllabic and mono-moraic (short vowel nucleus), the full-form preposition is mono-syllabic, but bi-moraic (long vowel nucleus).

While both the bi-moraic and the mono-moraic (i.e. proclitic) preposition may directly precede substantives or full-form pronouns like *njega* in (22a-b), only the bi-moraic preposition may directly precede enclitic pronouns (22c-d):[12]

---

[10]Note that in (20) the preposition is assumed not to be a phonological word, therefore the clitic *su* inverts with the complex *u stara*.

[11]As confirmed by Nedzad Leko and native speakers of Serbian, every single adjective in the complex PP in (21b) has to/can be stressed, i.e. represent phonological words.

[12]More information on properties of prepositions and the two different enclitic pronouns *nj* and *njú* vs. *ga* and *ju* of the full-form pronouns *njega* and *njû* can be found in Barić et al (1990).

(22) a.   *nâ njega*   b.   *ná-njega*
         on him

      c.   *nâ-nj*   d.   * *ná-nj*

Although the bi-moraic preposition can host clitics, as shown above, it is not possible for the enclitic reflexive pronoun *se* to 'split' the preposition from the pronoun in (23b). As the example (23c) shows, the preposition may function as a host for the enclitic pronoun *nj*, and the whole complex can host the enclitic *se* in string initial position.[13]

(23) a.   *Ná Ivana **se** naslonila.*
         on I.   self lean$_{ptc}$
         'She leaned on Ivan.'

      b.   * *Nâ/*Ná **se** njega naslonila.*
         on   self him   lean$_{ptc}$

      c.   [*Nâ **nj** **se***]$_w$ *naslonila.*
         on him self   lean$_{ptc}$

Therefore, contrary to what was claimed in Zec & Inkelas (1990), we conclude that prepositions may function as hosts for clitics.[14] Furthermore, we can conclude that prepositions, like other functional words may be phonological words, if stressed/accented. However, it is not possible for the preposition to function as a host for clitics in examples like (23b). An explanation for the ungrammaticality of (23b) appears to be straightforward, when this construction is compared with the following phenomena:

(24) a.   * [*Takvoj situaciji*]$_i$ ***smo se*** *našli u, da ...*
         such   situation   be$_{1pl}$ self find in, that ...

      b.   * *U smo se našli* [$_{PP}$ t$_i$ *takvoj situaciji*]*, da ...*

As can be seen in (24a), Croatian does not allow preposition stranding, and it is not possible for the preposition alone to be fronted, leaving the rest behind (24b).

However, what seems to be possible is splitting of complex phrases, DPs as well as PPs. This will be discussed in more detail in the following section.

4.3   Splitting Constituents

As the examples in (25) show, Croatian allows split-topicalization of parts of complex PPs. Apparently the non-constituent part of a PP can be topicalized, while the NP-part remains in situ:

(25) a.   [*U takvoj*] ***smo se*** *našli* [*situaciji*]*, da ...*
         in such   be$_{1pl}$ self find   situation   that

      b.   [*U tako lošoj*] ***smo se*** *našli situaciji, da ...*
         in so   bad   be$_{1pl}$ self find situation that

---

[13]I constructed the examples in (23) together with Nedzad Leko, who gave two comments: 1. the *nj* in (23c) has a default, or the only possible reading as [+human]; 2. the construction itself sounds rather archaic. While I agree with the first comment, I disagree with the second, e.g. native speakers from the Dalmatian coast and from Hercegovina accept constructions like (22c).

[14]Wilder & Ćavar (1994) and Ćavar & Wilder (1994) argue that prepositions may be morphologically and phonologically independent if they appear in contrastive coordinated constructions where the complement of one preposition undergoes ellipsis.

The same is true for complex DPs like in (26), where either the whole DP (26a), if it contains a *wh*-adjective, or the *wh*-adjective alone (26b) is moved to initial position:

(26) a.   *Kakva   kola je   Ivan kupio?*
           what-kind-of car be$_{3sg}$ I.   buy$_{ptc}$
           'What kind of car did Ivan buy?'

      b.   *Kakva   je   Ivan kola kupio?*
           what-kind-of be$_{3sg}$ I.   car buy$_{ptc}$

Since for all split-topicalization constructions, namely, splitting of DPs and PPs, the same constraints seem to hold in Croatian, in the following section we will focus on the split-PP cases.

4.3.1   Constraints on Split Topicalization

The split topicalization constructions in Croatian underly certain constraints. First, as the examples in (27) show, there is no stranding of the left branch of a complex PP with topicalization of the NP-part only:

(27) a.   *Ivan je   bacio   loptu na veliki ravni krov.*
           I.   be$_{3sg}$ throw$_{ptc}$ ball on big   flat   roof
           'Ivan has thrown a ball on a big flat roof.'

      b.   * *Krov$_i$ je   Ivan bacio   loptu* [$_{PP}$ *na veliki ravni* t$_i$]
           roof   be$_{3sg}$ I.   throw$_{ptc}$ ball   on big   flat

It is not possible for the NP *krov* to be extracted out of a complex PP as in (27b). Second, there is no split topicalization of the left branch alone, if the right branch does not move:

(28) a.   * [*Na kakav*]   *je   Ivan bacio   loptu* [*krov*] ?
             on what-kind-of be$_{3sg}$ I.   throw$_{ptc}$ ball   roof

      b.   * [*Na kakav*] *je Ivan bacio* [*krov*] *loptu?*

      c.   [*Na kakav*] *je Ivan* [*krov*] *bacio loptu?*

      d.   [*Na kakav*] *je* [*krov*] *Ivan bacio loptu?*

      e.   [*Na kakav krov*] *je Ivan bacio loptu?*

Whether one assumes the base position of the PP to be right of the direct object (28a), or left of it (28b), the NP-part of the complex PP may not remain in situ, but rather, has to move either to some position preceding VP (28c) or IP (28d). Alternatively, the whole PP may be topicalized, as in (28e).

Further conditions are that the remainder of such split-PP-topicalization constructions has to be a syntactic constituent. This fact and the contrast in (28) suggest an analysis of constructions like (28d) in terms of syntax, rather than, in terms of PI or pure phonology. The observation that the NP-part of a split constituent cannot remain in situ, forces an analysis in which the NP-part moves out of the complex PP first, and in a subsequent step the remnant PP is topicalized.

The examples in (28) show where the phonological and the PI-analysis massively undergenerate. In the following section, the cases of over-generation will be discussed.

4.4 Complex NPs

4.4.1 Nouns and relative clauses

The following examples show that complex NPs containing a head noun plus relative clause may be topicalized as in (29a), or the head-noun may be topicalized while the relative clause may be extraposed to the right peripheral position as in (29c), but the relative clause may not be stranded in some intermediate position in the clause, see (29b):

(29) a. [$_{DP}$ *One knjige* [$_{rel}$ *koje sam želio kupiti*]] *nisam   našao.*
these books   which be$_{1sg}$ wish$_{ptc}$ buy$_{inf}$ NEG-be$_{1sg}$ find$_{ptc}$
'I didn't find the books that I wanted to buy.'

  b. * [$_{DP}$ *One knjige*] *nisam* [$_{rel}$ *koje sam želio kupiti*] *našao.*

  c. [$_{DP}$ *One knjige*] *nisam našao* [$_{rel}$ *koje sam želio kupiti*]

Furthermore, complex DPs which contain a relative clause may be scrambled in embedded finite clauses, where they may occupy a position between the complementizer and the subject:

(30) *Ivan kaže*
I.   say$_{3sg}$

  da su mu [$_{DP}$ *one knjige* [$_{rel}$ *koje je   jučer   kupio*]] *poslali poštom.*
that be$_{3pl}$ him   these books   which be$_{3sg}$ yesterday buy$_{ptc}$ send$_{ptc}$ mail

  'Ivan said that they send him the books, that he bought yesterday, by mail.'

The PI-account predicts that the clitic cluster in (30) inverts with the following phonological word, if the word order in the embedded construction in (30) occurs in a matrix context. Although it seems to be possible to scramble a complex DP to IP in embedded finite clauses as in (30), (31b) shows that it is not possible for PI to operate on an underlying structure (31a):

(31) a. [$_{IP}$ *su mi* [$_{IP}$ [$_{DP}$ one *knjige* [$_{rel}$ *koje sam jučer   kupio*]] *poslali ...*]]
be$_{3pl}$ me   these books   which be$_{1sg}$ yesterday buy$_{ptc}$ send$_{ptc}$

  b. * [$_{DP}$ *One knjige*] *su mi* [$_{rel}$ koje sam jučer kupio] poslali poštom.

Note that this is not only problematic for the PI-account, but also for any account that claims that clitic placement is phonological.

   In order to rescue the PI-approach, one might argue that scrambling of DPs that contain relative clauses is restricted in matrix contexts. However, in examples like (32) with complex subject DPs, we observe the same restriction, namely, PI cannot operate on the underlying representation (32a):[15]

(32) a. [$_{IP}$ *se* [$_{IP}$ [$_{DP}$ *čovjek koji mi je   obećao   pomoći*] *nije   pojavio*]]
self   man   who me be$_{3sg}$ promise$_{ptc}$ help$_{inf}$ NEG-be$_{3sg}$ show-up

  b. * [$_{IP}$ __ [$_{IP}$ [$_{DP}$ *čovjek se koji mi je obećao pomoći*] *nije pojavio*]]

---

[15]Note that *pojaviti* ('show up') requires a reflexive pronoun, and that the reflexive pronoun *se* causes drop of the finite auxiliary *je* ('to be', 3sg).

It is not possible for the enclitic reflexive pronoun *se* to occur in some position between the head noun and the relative clause, although the head noun of the subject DP in (32) is a phonological word.

If it is assumed that the subject DP with relative clause occupies the spec-IP position, it is unclear why PI cannot occur in such cases. While the PI-account would have to stipulate obligatory topicalization of the complex subject DP, the split-topicalization analysis explains the ungrammaticality of (32) in terms of independent syntactic constraints, i.e. splitting of the head noun and the relative clause is only possible if the relative clause is extra-posed.

4.4.2   Noun Selected Infinitives

Other cases where a complex XP cannot be syntactically split in Croatian are complex DPs which contain a noun selected infinitive.[16]

In the examples in (33) the infinitive clause may not be split from the head-noun, i.e. the infinitive may not be extra-posed or stranded in some base-position as in (33c), and, as expected, clitics may not split a complex DP, if it appears in a sentence initial position, see (33d):

(33)   a.   *Ivan **mi je**   dao   mogućnost upoznati   Mariju.*
            I.   me be$_{3sg}$ give$_{ptc}$ possibility get-to-know$_{inf}$ M.
            'Ivan offered me the possibility to get to know Mary.'

       b.   [*Mogućnost* [*upoznati Mariju*]] **mi je** *Ivan dao.*

       c.   * [*Mogućnost*] **mi je** *Ivan dao* [*upoznati Mariju*]

       d.   * [*Mogućnost*] **mi je** [*upoznati Mariju*] *Ivan dao.*

Compared with complex DPs which contain a relative clause, the DPs that contain N-selected infinitives differ only with respect to extraposition. Although DPs that contain N-selected infinitives may be scrambled to IP, or base generated in spec-IP, the Last Resort operation, PI, is not possible.

Again, the PI-approach fails to explain the observed restrictions. In fact, the PI-operation turns out to be completely unnecessary, since all the cases where the operation of PI seems to be motivated, can be explained in purely syntactic terms. Complex DPs cannot be split with clitics, if these DPs cannot be split in syntax. All the cases where the PI-approach over-generates, are cases that are excluded in syntax for independent reasons; all the cases where the PI-approach under-generates, are cases where splitting of complex DPs is independently licensed in syntax.

5   The Syntactic Solution

In the following it will be argued that in order to explain the discussed phenomena, one has to adopt a purely syntactic analysis. The analysis presented in Wilder & Ćavar (1994) and Ćavar & Wilder (1994) is the best candidate for a descriptively adequate theory of clitic placement in Croatian.

In Wilder & Ćavar (1994), it is argued that clitic placement is syntactic, i.e. clitics always occupy the C^0 position. Placement of the clitics in C^0 is responsible for the W-effect. The TM-effect is explained by assuming that special clitics in Croatian include the subcategorization

---

[16]Constructions like (33a-b) don't exist in Serbian, or if some native speakers accept them, they seem to be rather marginal.

frame (15) as part of their lexical specification. This requires a phonological word to precede the clitic cluster in a PF-representation.

Additionally, it is assumed that apparent XP-splits (cf. (12), (28d), etc.) result from scrambling out of a complex XP, with subsequent topicalization of the remnant.

The predictions are that only one XP and/or one X^0 may precede the clitic cluster inside the CP-domain, since CP offers only one landing-site for an XP, and only one head-position. This explains the strong adjacency condition between the complementizer and the clitic cluster in embedded finite complements (cf. (4a-b) vs. (8)), and the strong adjacency between fronted verbs and the clitic cluster in the so called *Long Head Movement* (LHM) constructions.[17]

Furthermore, it is predicted that only one XP may precede the clitic cluster in the CP-domain. Certain types of topicalization that seem to be counterexamples to this prediction will be discussed in the following section.

5.1 Topicalization

As correctly observed in Zec & Inkelas (1990), there seems to be a constraint on topicalized XPs, comparable with the Heavy NP Shift cases: only branching phonological representations are well formed topics (Zec & Inkelas, 1990: 373):

(34) a. *Taj čovek voleo **je** Mariju.*   b. * *Petar voleo **je** Mariju.*
          that man   love$_{ptc}$ be$_{3sg}$ M.   P.   love$_{ptc}$ be$_{3sg}$ M.
          'That man loved Mary.'   'Peter loved Mary.'

Zec & Inkelas (1990) claime that there is a phonological constraint on topicalization, that heavy constituents may be topicalized as in (34a) while light constituents may not, cf. (34b). Heaviness is defined in terms of branching of phonological constituents, i.e. a bi-moraic syllable is heavier than a mono-moraic, a phonological phrase that contains more than one phonological word is heavier than a phonological phrase that contains only one.

However, as argued in Wilder & Ćavar (1994) and Ćavar & Wilder (1994), a distinction between CP external (free) topics and topicalization to spec-CP has to be made.[18] The following examples (see (7) in section 1.2) show that there is a difference between the two types of topics:

(35) a. *Stipi **su** Ivan i Marija sinoć   dali   knjigu.*
          S.   be$_{3pl}$ I. and M.   yesterday give$_{ptc}$ book
          'Ivan and Mary gave a book to Stipe yesterday.'

      b. *Sinoć **su** Ivan i Marija Stipi dali knjigu.*

      c. * *Stipi sinoć **su** Ivan i Marija dali knjigu.*

As in (34b) a non-branching phonological representation, i.e. a subject which is a single phonological word, cannot function as a topic in (34b), an indirect object as in (35a), or an adverb as in (35b), which has a non-branching phonological representation can function as a topic.

Furthermore, a construction like (35c), where two such topics occur, is ungrammatical, while a similar construction (34a) is well-formed with more than one topic:

(36)   *Taj čovjek, prošli tjedan, u nekom parku u Zagrebu, poljubio **je**   Mariju.*
         this man   last   weak   in some   park in Zagreb kiss$_{ptc}$ be$_{3sg}$ M.
         'Last weak this man kissed Mary in some park in Zagreb.'

---

[17]A discussion of the *Long Head Movement* constructions would lead beyond the scope of this paper. For further details see Wilder & Ćavar (1994).

[18]This distinction may also be formulated in terms of adjunction to the clause (free topics as in (34)), and movement to some clause internal specifier position (as in (35a-b)).

Another observation which supports the proposed difference between these two topic-constructions is that [+*wh*]-phrases, i.e. complex *wh*-phrases as in (37)[19] and [+NEG]-phrases,i.e. negative elements that require negative concord as in (38), have to be string adjacent to the head that contains the clitic cluster, if they are moved to a sentence-initial position:

(37) * *Koji   čovjek, prošli tjedan, u Zagrebu poljubio **je**   Mariju.*
       which man   last   weak   in Zagreb kiss$_{ptc}$ be$_{3sg}$ M.

(38) a.   * *Ni u kom slučaju, naša Marija nebi   **ga** poljubila.*
            in-no-case   our M.   NEG-would him kiss$_{ptc}$

      b.   *U nekom parku, naša Marija **ga** ni u kom slučaju nebi   poljubila.*
            in some   park   our M.   him in-no-case   NEG-would kiss$_{ptc}$
            'In no case our Mary would kiss him in some park.'

      c.   *Ni u kom slučaju **ga** naša Marija nebi poljubila.*

These data suggest that, while there might be a phonological restriction on CP external topics (comparable with Heavy NP Shift in English), there is definitely no such constraint with respect to topicalization to spec-CP.

Furthermore, with respect to topicalization, only one XP-position seems to be available preceding the clitic cluster, spec-CP.

5.2   *Wh*-Constructions

Further evidence for analysing clitic placement in Croatian as placement of the clitics in C^0, comes from multiple *wh*-constructions.

In multiple *wh*-questions one *wh*-element has to move to sentence initial position (CP-spec), while the other *wh*-elements may either remain in situ as in (39a), or, all *wh*-elements cluster in sentence initial position as in (39b):

(39) a.   *Šta Ivan daje komu?*
           what I. give who
           'What does Ivan give to who?'

      b.   *Šta komu Ivan daje?*
           what who I. give

However, if the sentence contains clitics, the clitics always have to follow the initial *wh*-constituent. In contrast to topicalization constructions like in (36), the enclitic auxiliary *je* may not appear in third position, following two *wh*-elements (40c-d), but has to follow directly the first *wh*-phrase (40b), if two *wh*-phrases are fronted:

(40) a.   *Šta **je**   Ivan komu dao?*   b.   *Šta **je** komu Ivan dao?*
           what be$_{3sg}$ I.   whom give$_{ptc}$
           'What did Ivan give to whom?'

      c.   * *Šta komu **je** Ivan dao?*   d.   * *Komu šta **je** Ivan dao?*

One can conclude that in multiple *wh*-questions, where apparently all *wh*-elements cluster together in initial position, in fact, only one occupies the spec-CP position, while the others are located lower of C^0, probably scrambled to IP.

---

[19]This is discussed in more detail in section 5.2.

6 Conclusion

It was argued in sections 2 and 4 that the phonological account presented in Zec & Inkelas (1990) does not explain the facts with respect to clitic placement in Croatian. The impression that clitics may appear after the first phonological word of some complex constituent in sentence initial position is not consistent with the syntactic properties of the relevant constructions. Rather, the syntactic properties of constructions where clitics apparently split complex constituents suggest that clitic placement is syntactic in nature. Hence, all analyses that try to explain clitic placement as occuring after the first syntactic constituent (1C), and in addition after the first phonological word (1P) are inadequate (cf. Halpern, 1992; Schütze, 1994; Zec & Inkelas (1990)). Since clitics in Croatian always appear after an XP (*wh*-phrase or topic) and/or a syntactic head X^0, it is only necessary to explain why there is clitic placement after the first syntactic constituent.

Nevertheless, certain phonological conditions with respect to constructions containing clitics seem to hold. The generalization that clitics may never appear in string-initial position seems to require a phonological explanation. The explanation proposed by Zec & Inkelas (1990), in terms of lexical properties of clitics, is the most promising.

Sources

Barić, E. et al, 1990. Gramatika Hrvatskoga Književnog Jezika. Školska Knjiga, Zagreb.

Drilo, S., 1973. Kroatisch–Serbisch. Teil I & II. Julius Groos Verlag, Heidelberg.

Katičić, R., 1986. Sintaksa Hrvastskoga Književnog Jezika. Globus, Zagreb.

References

Ćavar, Damir. 1996 (May). *On Cliticization in Croatian: More Syntax than Prosody!* Paper presented at the Workshop on the Syntax, Morphology and Phonology of Clitics.

Ćavar, Damir, & Wilder, Chris. 1994. "Clitic Third" in Croatian. *Linguistics in Potsdam*, **1**, 25–63.

Halpern, Aaron Lars. 1992. *Topics in the Placement and Morphology of Clitics*. Ph.D. thesis, Stanford University.

Schütze, Carston T. 1994. Serbo-Croatian Second Position Clitic Placement and the Phonology-Syntax Interface. *MIT Working Papers in Linguistics*.

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

Toman, Jindrich. 1993 (June). *A Discussion of the Second Position in Czech*. Talk given at the Wackernagel workshop.

Vogel, Irene, & Kenesi, István. 1990. Syntax and Semantics in Phonology. *Pages 339–363 of:* Inkelas, Sharon, & Zec, Draga (eds), *The Phonology-Syntax Connection*. The University of Chicago Press.

Wackernagel, Jakob. 1892. Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen*, **1**, 333–436.

Wilder, Chris, & Ćavar, Damir. 1994. Long Head Movement? Verb Movement and Cliticization in Croatian. *Lingua*, 1–58.

Zec, Draga, & Inkelas, Sharon. 1990. Prosodically constrained syntax. *Pages 365–378 of:* Inkelas, Sharon, & Zec, Draga (eds), *The Phonology–Syntax Connection*. Chicago, IL: University of Chicago Press.

Zwicky, Arnold M. 1985. Clitics and Particles. *Language*, **61**(2), 283–305.

Zwicky, Arnold M., & Pullum, Geoffrey K. 1983. Cliticization vs. inflection: English *n't*. *Language*, **59**, 502–513.