

**CLITIC PLACEMENT, SYNTACTIC DISCONTINUITY  
AND INFORMATION STRUCTURE**

Damir Cavar                      and              Melanie Seiss  
Eastern Michigan University              Universität Konstanz

Proceedings of the LFG11 Conference

Miriam Butt and Tracy Holloway King (Editors)

2011

CSLI Publications

<http://csli-publications.stanford.edu/>

## Abstract

The phenomenon of so-called second position clitics has received considerable attention in the linguistic literature, and some proposed analyses of the phenomenon have suggested major architectural changes to linguistic theories. In this paper, we look at second position clitics in New-Shtokavian, their syntactic and prosodic properties, and propose a purely syntactic clitic placement analysis. We show that the complex data can be accounted for by an analysis of split constituents and their resulting information structure differences with a simple prosody-syntax interface.

## 1 Introduction

The phenomenon of so-called second position clitics (2PC) in particular in Slavic languages has received considerable attention in the linguistic literature over the last decades. Some proposed analyses of the phenomenon have led to major architectural changes to linguistic theories, and new powerful mechanisms, which for many appeared unnecessary and unmotivated, leading to wrong predictions and over-generation. In this paper, we look at second position clitics in New-Shtokavian (NSh), their syntactic and prosodic properties, and propose a purely syntactic clitic placement analysis.

NSh in the *ije*-kavian variant currently represents the Croatian standard language, while the *i*- and *e*-kavian variants are spoken in e.g. Bosnia–Herzegovina and Serbia respectively. The examples discussed in the following are from the NSh variant spoken in Croatia.<sup>1</sup>

The problem we are concerned with in this paper is illustrated by the examples in (1). In NSh pronouns and auxiliaries can be realized in a (morphologically and/or phonologically) reduced form. We refer to these elements as clausal clitics, i.e. pronominalized verbal arguments or clausal auxiliary verbs. Such clausal clitics seem to be subject to a second position placement constraint, which apparently renders them obligatorily in either a position after a clause-initial syntactic constituent (IP-constructions), or after the initial phonological word (IW-constructions). This is illustrated in (1a,b) for a sentence initial subject NP, and in (1c,d) for an initial object NP.

- (1) a. *Novi auti su stigli u skladište.*  
new cars be.3pl arrive in storage  
'New cars arrived at the storage.'

---

<sup>†</sup>We thank Miriam Butt, Steven Franks, Maribel Romero, and Malgosia Cavar, as well as our colleagues at the University of Zadar, the Institute of Croatian Language and Linguistics, and the University of Konstanz, and the anonymous reviewers and participants of the LFG 2011 in Hong Kong for their comments and help.

<sup>1</sup>We expect judgement difference between our data and data taken from work based on so called BSC (Bosnian, Serbian, Croatian) or Serbo-Croatian, which often focuses on particular variants, or ignores regional differences and variation.

- b. *Novi su auti stigli u skladište.*  
 new be.3pl cars arrive in storage
- c. *Novi auto su naši susjedi kupili.*  
 new car be.3pl our neighbors buy  
 ‘Our neighbors bought a new car.’
- d. *Novi su auto naši susjedi kupili.*  
 new be.3pl car our neighbors buy

The analysis of clitic placement in (1a,c) as placement after the first syntactic phrase (a 1P-construction) appears to be straightforward. The placement of clitics in positions after the initial adjective as in (1b,d) (a 1W-construction) raises questions about the nature of the underlying constraints and mechanisms. The question, whether 1W-constructions are syntactic or phonological in nature, is the matter of many debates. In these 1W-constructions an item contributing information to the clause level is apparently embedded inside a fronted nominal phrase. Would this indeed be the most plausible analysis, it would induce problems for various syntactic theories, e.g. related to level autonomy, syntactic placement constraints, or to the fact that clitics contributing information to the sentence level appear to be inside a subconstituent of arbitrary complexity and embedding depth. An explanation of constructions as in (1b,d) was offered by approaches that suggest a complex prosody-syntax interface, see e.g. Halpern (1995) for a generative approach, and more recently Bögel et al. (2010) within the LFG framework.

In contrast to the prosodic analyses and suggested extensions of the prosody-syntax interface, we argue in this article for a purely syntactic explanation for both construction types, i.e. 1W- and 1P-constructions. We show that the apparent cases of phonological clitic placement in 1W-constructions can be analyzed as instances of split constituent constructions, and that the word order variation is related to information structure, implying scope differences in a hierarchical (syntactic) representation, rather than scope neutral phonological processes. As a consequence of our analysis, the prosody-syntax interface remains rather simple, and does not utilize complex word rearrangement mechanisms outside of syntax, or at the level of phonological representations.

The article is structured as follows. In section 2, we briefly outline previous analyses of the phenomenon and discuss some of their shortcomings. Sections 3 and 4 present the relevant data in more detail. We present evidence that apparent phonological clitic placement in fact involves split constituents in section 3. In section 4 we show that the word order variation is related to information structure. The outline of our analysis is given in section 5, and section 6 discusses how our analysis can account for some of the more complex examples. Section 7 concludes the discussion.

## 2 Previous analyses

There has been extensive work on second position clitics (2PC) in general, and in the recent years in particular, see e.g. Halpern and Zwicky (1996); van Riemsdijk (1999); Franks and King (2000); Anderson (2005). Assumptions and hypotheses related to 2PC in Bosnian, Croatian, and Serbian (or Serbo-Croatian) can be roughly divided into purely phonological, purely syntactic or mixed phonological-syntactic accounts.

Purely phonological accounts for 2PC placement in the respective languages and dialects such as e.g. Radanović-Kocić (1988, 1996) and O'Connor (2002) assume that clitic placement is not subject to syntactic constraints, but rather restricted by purely phonological processes and requirements. Thus, in phonological accounts clitics are either placed after the first phonological word or after the first phonological phrase. In such models the fact that the respective phonological categories usually overlap with syntactic ones is responsible for the specious analysis of purely syntactic clitic placement. Such assumptions require the formulation of a so far missing rich theory of phonological or prosodic representations, where word order and word placement restrictions can be formulated that make the appropriate predictions, and offer plausible explanations.

Diametrically opposed are accounts which suggest purely syntactic mechanisms and constraints for 2PC placement. Representative for a purely syntactic approach to 2PC placement are e.g. Cavar and Wilder (1992) and Progovac (1996). Here we draw on the observations and arguments therein, but we will extend the observations and arguments taking information structure into account. As far as we know, information structure has not been the focus of 2PC placement analyses so far, although some work has hinted at relevant implications (e.g. Diesing et al. 2009).

A third strand of analyses can be called mixed phonological-syntactic accounts. In these analyses it is assumed that 2PC are either placed after the first syntactic constituent, or after the first phonological word. Most of these accounts (e.g. Schütze 1994; Halpern 1995) relate to work presented in Zec and Inkelas (1990). Such proposals have in common that an architecture of a phonology-syntax interface is suggested which involves active movement or placement of elements like clitics at a phonological or prosodic level, or at the interface between these and the syntactic level. In this section, we discuss in more detail in particular Halpern's (1995) account of prosodic inversion, and its recent incorporation into LFG by Bögel et al. (2010).

In the prosodic inversion account proposed by Halpern (1995) and adopted by Bögel et al. (2010), enclitics are assumed to be placed in syntax either after the first syntactic constituent or sentence initially. There is no plausible explanation offered for this particular syntactic placement peculiarity. In constructions where 2PC are placed after a clause initial syntactic constituent, the stipulated prosodic requirement of enclitics requiring a prosodic host to their left is accounted for naturally. However, if 2PCs are placed sentence initially, the enclitics cannot attach

prosodically to a host to their left. In these cases reordering at the prosodic level is suggested as the mechanism that renders the appropriate word order that satisfies the prosodic requirement of enclitics to attach to a preceding prosodic host. This mechanism is called Prosodic Inversion (PI), i.e. a local last resort inversion operation that affects a clause initial enclitic and the prosodic word that follows it, or clause final proclitics and the prosodic word that precedes it. In short, it is suggested that the last resort operation of PI reorders or moves words at the level of prosodic or phonological representation.

Halpern's (1995) concept of PI extends the previously assumed complexity of the phonology-syntax interface significantly. Figure 1 sketches this idea of the mapping between phonology and syntax, as adopted in Bögel et al. (2010). In the c-structure, 2PCs are assumed to be realized sentence-initially. This ensures that the pronominal clitics contribute their information to the sentence level. However, the second line displays the phonological spell out of the c-structure, in which the clitics have been moved to a position following the first phonological word. The interface mapping is assumed to be carried out by a complex rule which ensures that the clitics are only moved minimally and when necessary.

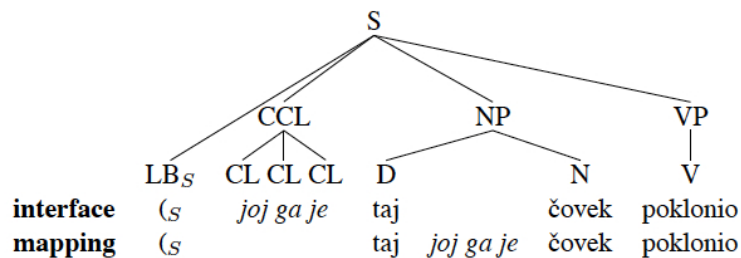


Figure 1: Complex prosody-syntax interface (Bögel et al., 2010)

The PI analysis for 2PC placement faces various problems, e.g. it involves a range of unmotivated conceptual stipulations. For example, it has to stipulate that 2PCs are realized syntactically in sentence initial position. This unmotivated stipulation serves as the main motivation for PI, rendering it conceptually necessary. However, there does not seem to exist any empirical data that supports the initial placement stipulation. Moreover, the motivation for assuming two different 2PC placement strategies in general seems unmotivated and lacks empirical support. In fact, there appear to be clear information structure differences between the two resulting surface constructions in the contrast in (1a,c) and (1b,d). These are not explained by the PI-account.

Further, the PI-account faces empirical problems. It over-generates, as the NSH-examples in (2) show:

- (2) a. *Na nj(ega) je bacila pogled.*  
on him/it be.3sg throw.ptc look  
‘She was looking at it.’

- b. \**Na je njega bacila pogled.*  
 on be.3sg him/it throw.ptc look

(2a) shows that preposition can function as host for an enclitic pronominal complement *nj*. It is possible to lengthen the prepositional nucleus and assign stress to the initial prepositional syllable, rendering them well-formed phonological words.<sup>2</sup> In the case of a stressed preposition, PI would also predict the inversion of an initial enclitic auxiliary and a following adjacent stressed preposition to be possible, contrary to the empirical facts, as (2b) shows.

The PI-account also under-generates. A PI-version that is restricted to inversion with an adjacent phonological word excludes NSh-examples in which 2PCs are placed after the second or third phonological word or syntactic constituent, as shown in (3). Such constructions and their analyses will be discussed in more detail in section 6.

- (3) a. *Taj naš veliki nam prihod neće tako puno pomoći.*  
 this our big us income not-will.3sg this much help.inf  
 ‘This big income of ours will not help us that much.’
- b. *Takav veliki brzi mi auto nije posebno potreban.*  
 such big fast me car not-be.3sg specially necessary  
 ‘I don’t need such a big fast car that much.’

To sum up, besides purely phonological and syntactic accounts, we discussed in particular the shortcomings of the mixed account that introduces the last resort operation of PI. However, all these accounts have in common that they cannot motivate or explain the intra-linguistic variation, i.e. the alternations of the different constructions. In the following we shall argue that clitics are always placed after an initial syntactic constituent, and we shall present an alternative information theoretic approach that explains the intra-linguistic variation between the two core constructions presented in (1).

### 3 Split Constituents

In this section we argue that NSh licenses discontinuous rendering of constituents in syntax, as described in e.g. Progovac (1996), or Fanselow and Cavar (2001, 2002) and citations therein. The possibility of complex syntactic discontinuities leads to the specious analysis of prosodic 2PC placement.

Examples of discontinuous noun phrases are presented in the examples in (4) and (5). The examples (4a) and (5a) show that an enclitic auxiliary *su* or *je* can intervene between the two words of a subject or object noun phrase respectively. However, this is also true for non-clitic sentential adverbs, as the examples (4b) and (5b) show.

<sup>2</sup>See Cavar and Cavar (2011) for a detailed phonetic and phonological analysis of stressed prepositions and their status of being bi-moraic phonological words, capable of hosting clitics.

- (4) a. *Novi su auti stigli u skladište.*  
 new be.3pl cars arrive.ptc in storage  
 ‘New cars arrived at the storage.’
- b. *Novi su danas auti stigli u skladište.*  
 new be.3pl today cars arrive.ptc in storage
- (5) a. *Taj je čovjek nazvao.*  
 this be.3sg man call.ptc  
 ‘This man called.’
- b. *Taj je danas čovjek nazvao.*  
 this be.3sg today man call.ptc

There is in fact ample empirical evidence for the possibility to render complex NPs discontinuously without 2PCs being involved at all. It seems plausible to assume that syntactic discontinuity of complex phrases is responsible for the apparent splitting of these phrases by 2PCs. Well-known examples of discontinuities in interrogative contexts include the examples in (6), as discussed, for example, in Browne (1976).<sup>3</sup>

- (6) a. *Ivan je kupio neki plavi auto.*  
 Ivan be.3sg buy.ptc some blue car  
 ‘Ivan bought some blue car.’
- b. *Kakav je Ivan kupio auto?*  
 what-kind-of be.3sg Ivan buy.ptc car  
 ‘What kind of car did Ivan buy?’
- c. *Kakav je Ivan auto kupio?*  
 what-kind-of be.3sg Ivan car buy.ptc

The possibility of realizing complex phrases discontinuously is not restricted to NPs. For example, PPs can be split in interrogative contexts in the same way, as shown in the examples in (7).

- (7) a. *Ivan je živio u velikom gradu.*  
 Ivan be.3sg live.ptc in big city  
 ‘Ivan lived in a big city.’
- b. *U kakvom je Ivan živio gradu?*  
 in what-kind-of be.3sg Ivan live.ptc city  
 ‘What kind of car did Ivan buy?’
- c. *U kakvom je Ivan gradu živio?*  
 in what-kind-of be.3sg Ivan city live.ptc

<sup>3</sup>Such syntactic discontinuities are common in other Slavic and Non-Slavic languages, see e.g. Fanselow and Cavar (2001, 2002) and the references therein, Obenauer (1976) for French, or Nakanishi (2007) for Japanese.

There are fewer possibilities of rendering discontinuous PPs compared to NPs. While a complex PP can be discontinuously realized in a linearly stretched way, maintaining the underlying canonical word order as in (7), it is not possible to render the elements in a crossing manner, as in example (8). In terms of theories that rely on the concept of movement, example (8) shows that PPs are islands for extraction in NSh, which imposes serious problems for a movement-based explanation of the examples in (7b-c).

- (8) \**Gradu<sub>i</sub> Ivan živi u velikom \_\_\_\_<sub>i</sub>*  
 city Ivan live.3sg in big

The examples in (9) show that 2PCs cannot occur in positions where a complex PP cannot be split. It is not possible for a subject NP to intervene between the preposition and the prepositional complement *nekom grad* in (9a), nor is it possible for an enclitic to intervene in this position (9b). This is true even in cases in which the preposition is clearly an independent phonological word (a bi-moraic and stressed unit) that can host an enclitic complement.

- (9) a. \**U Ivan nekom gradu živi.*  
 in Ivan some city live.3sg  
 b. \**U je nekom gradu Ivan živio.*  
 in be.3sg some city Ivan live.ptc

Furthermore, we observe that in split contexts in NSh one split NP-part, for example the head-noun, can be pronominalized in interrogative (10b) and declarative (10c) contexts, while in non-split contexts (10d) this is impossible. This further supports the assumption that we are dealing with two distinct noun phrases in these constructions, as e.g. proposed in Fanselow and Cavar (2001, 2002).

- (10) a. *Koliko si knjiga pročitao?*  
 how-many be.2sg books read.ptc  
 ‘How many books did you read?’  
 b. *Koliko si ih pročitao?*  
 how-many be.2sg them read.ptc  
 c. *Sve sam ih pročitao.*  
 all be.1sg them read.ptc  
 ‘I read all of them.’  
 d. \**Ivan je pročitao pet ih.*  
 Ivan be.3sg read.ptc five them

As has been discussed in Cavar and Wilder (1999) and O’Connor (2002), 2PCs cannot split the head noun from its relative clause in sentence initial position.



- (11) a. *Čovjek, koji živi u Parizu, sjeća vas se.*  
 man who live.3sg in Paris remember.3sg you.2pl.acc refl  
 ‘The man, who lives in Paris, remembers you.’
- b. \**Čovjek vas se, koji živi u Parizu,*  
 man you.2pl.acc refl who live.3sg in Paris  
*sjeća.*  
 remember.3sg

Given that the head-noun *čovjek* of the complex subject-NP represents a well-formed prosodic word, it would be expected that the 2PC cluster *vas se* can be generated in sentence initial position and invert with the NP-head in the PI approach. This, however, is not possible. There is a simple syntactic explanation for the ungrammaticality of (11b), i.e. the head noun cannot be split from its relative clause by any other element. The only option for a discontinuous realization of an NP with a relative clause would involve right-extraposition of the relative clause, similar to German and English.

In this section we have presented arguments for a syntactic analysis of split NPs and PPs, establishing a parallel between the syntactic splits and splits with the presence of 2PCs. In cases in which syntactic splits are excluded, splits by 2PCs are excluded as well. The fact that e.g. head nouns in split-NP constructions can be pronominalized supports the assumption that each resulting part of a split NP can function as an independent syntactic NP constituent.

## 4 Information structure

In this section, we argue that the different syntactic structures with 2PCs also differ with respect to their specific information theoretic properties. Semantic and pragmatic effects, however, are not expected, if the minimal word order difference involves clitic placement at the prosodic level.

The examples in (12) illustrate that 2PC split constructions are not possible in neutral contexts that form the answer to the question “What happened?”, as shown in (12a) for an oblique argument, and in (12b) for a subject NP.

- (12) a. ?? *U velikom je gradu Petar živio.*  
 on big be.3sg tree Peter climb.ptc  
 ‘Peter climbed on a big tree.’
- b. ?? *Taj nepoznati je čovjek nazvao Mariju.*  
 this unknown be.3sg man call.ptc Maria  
 ‘This unknown man called Maria.’

A further test involves quantifier scope variation in wh-questions with and without splits, as described by Obenauer (1976) for *combien*-split constructions in French,

and by Nakanishi (2007) for split NPs in Japanese.<sup>4</sup>

In example (13), the 2PC unambiguously cliticizes to the fronted direct object phrase. The sentence has two readings, i.e. the collective and the distributive reading. In contrast, in questions involving split NPs as in (14) only the collective reading remains.

- (13) *Koliko članaka su svi ti studenti pročitali?*  
 how-many articles be.3pl all these students read.ptc

?n:  $\exists x$  article(x) &  $\forall y$  [ student(y)  $\rightarrow$  read(y,x) ]

How many articles exist, such that all students read them?

?n:  $\forall y$  [ student(y)  $\rightarrow$   $\exists x$  article(x) & read(y,x) ]

What is the number, such that all students read that number of papers?

- (14) *Koliko su svi ti studenti pročitali članaka?*  
 how-many be.3pl all these students read.ptc articles

\* ?n:  $\exists x$  article(x) &  $\forall y$  [ student(y)  $\rightarrow$  read(y,x) ]

?n:  $\forall y$  [ student(y)  $\rightarrow$   $\exists x$  article(x) & read(y,x) ]

The observation in (14) can be explained as being the result of a syntactic split in a wh-question, as discussed above. Similarly, the same disambiguation effect seems to occur in constructions with an apparent prosodic NP-split, as shown in (15), where only the collective reading is available.

- (15) *Koliko su članaka svi ti studenti pročitali?*  
 how-many be.3pl articles all these students read.ptc

\* ?n:  $\exists x$  article(x) &  $\forall y$  [ student(y)  $\rightarrow$  read(y,x) ]

?n:  $\forall y$  [ student(y)  $\rightarrow$   $\exists x$  article(x) & read(y,x) ]

As shown above, the different word order with 2PCs, after the initial word in split constructions, or after the initial constituent, has consequences with respect to the contribution to the sentence information structure. This is not expected if a purely phonological placement were responsible for the 1W-placement. The observations support a syntactic placement of clitics in the 1W-constructions.

## 5 Basic Analysis

Our analysis has to account for the fact that clitics can be realized in 1P- and 1W-constructions, as discussed above. However, clitics can also be realized in

<sup>4</sup>Thanks to Maribel Romero for the relevant hints and a fruitful discussion of the semantic and pragmatic properties of related split constructions.

third or fourth position (prosodically, and syntactically, see e.g. Cavar and Wilder (1999)). Examples of such a deeper placement are presented in (17) and discussed below. It is important, however, that our analysis is also compatible with these facts. Similarly, our analysis has to explain why clitics cannot be realized after the first prosodic word in embedded contexts, and that there is a strict string adjacency condition between complementizers and clitics.

Any adequate analysis of the relevant construction has to be complex, because it involves the interaction of various linguistic levels, such as syntactic, information, and prosodic structure, as well as semantic properties.

We will not discuss details of relevant prosodic constraints and requirements for the 1W- and 1P-constructions, nor their specific intonation contours. We are confident that our approach is compatible with various recent proposals of prosodic structure within LFG such as Bögel et al. (2010); O'Connor (2004); Mycock (2005); Mycock and Dalrymple (2011).

We also have to skip a detailed discussion of the semantic properties of the relevant constructions. However, we are certain that an appropriate model of the semantic structure can be incorporated in our approach, e.g. along the lines of Dalrymple and Nikolaeva (2011).

The data discussed in sections 3 and 4 showed that the basic levels needed to analyze the data are syntactic structure and information structure. In our analysis, the clitic cluster always follows a syntactic phrase, i.e. 1W-constructions are cases of syntactic structures with discontinuous syntactic phrases. Following work on i(nformation)-structure in LFG (e.g. Bresnan 2001; Choi 1999; King 1997), we assume that certain c-structure positions can be associated with information structure functions. Thus, while the 1W- and 1P-constructions might have the same f-structure, their c- and i-structure might differ.

Figure 2 shows the basic c-structure template assumed for all sentences, i.e. for syntactically discontinuous 1W-constructions, or the continuous 1P-constructions. The 2PCs or the clitic cluster mark the boundaries between TOP(ic) and FOC(us), i.e. the elements following the clitic cluster are in the default focus domain (e.g. VP), associated with the FOC(us) role. The elements before the clitic cluster can be interrogative XPs, TOP(ic) or C(ontrastive) FOC(us) elements.

In the sentential template, the position preceding the clitic cluster is marked as either a topic or a contrastive focus position. This may seem complicated at first sight. However, the similarity between topic and contrastive focus has been noted in various works before (e.g. King 1995; Choi 1999; Mchombo et al. 2005; Cook and Payne 2006). Applying, e.g. Choi's (1999) features "Prom[inent]" and "New" to encode the basic information structure roles (see table 1), topic and contrastive focus share the feature [+Prom]. Thus, the Spec-CP position can be associated with the feature [+Prom], i.e. it would require that all syntactic objects in this position play a prominent role in the information structure.

Specific focus background structures are associated with all split constructions. These are additionally also prosodically marked. These constructions require the initial split subconstituent to be stressed. Thus, without going into details of the

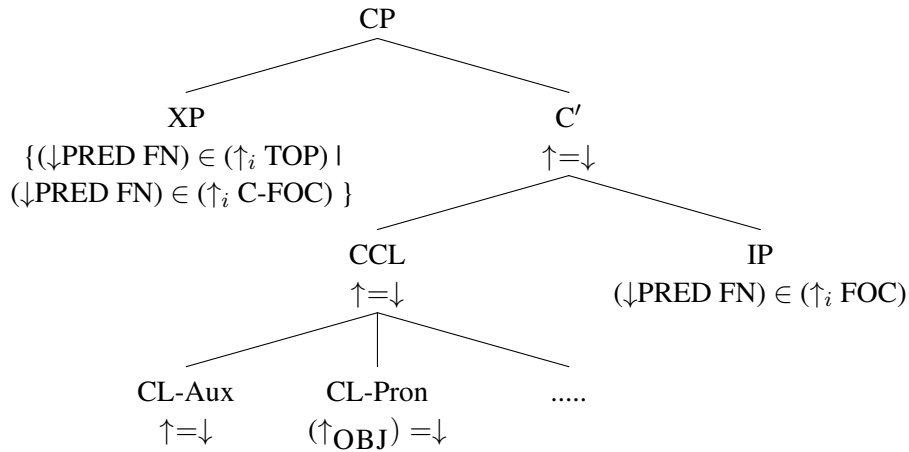


Figure 2: Basic c-structure template including a clitic cluster (CCL)

	- New	+ New
- Prom	Tail	Completive Focus
+ Prom	Topic	Contrastive Focus

Table 1: Information structure features (Choi, 1999)

prosodic structure, stressing the first subconstituent might contribute the information [+New]. This is required to put contrastive focus on the split constituent. Without this stress, the split constituent would be analyzed as the topic, which would be information-structurally odd.

How the template is applied in the different analyses for the two basic constructions is illustrated by using the well-known examples in (16).

- (16) a. *Taj čovjek joj ga je poklonio.*  
 this man her it be.3sg present.ptc  
 ‘This man presented it to her.’  
 b. *Taj joj ga je čovjek poklonio.*  
 this her it be.3sg man present.ptc

Figure 3 shows the analysis for (16a), in which the clitic attaches after the initial syntactic phrase, i.e. the subject NP. The subject does not necessarily receive stress. NSh is a pro-drop language such that any overt subject is actually interpreted as prominent. Thus we analyze this construction with an initial subject NP as an instance in which the initial phrase is representing a topic. The clitic cluster

consists of three clitics in this case, the object<sub>θ</sub>, the object and the auxiliary. The focus is solely projected by the verb.

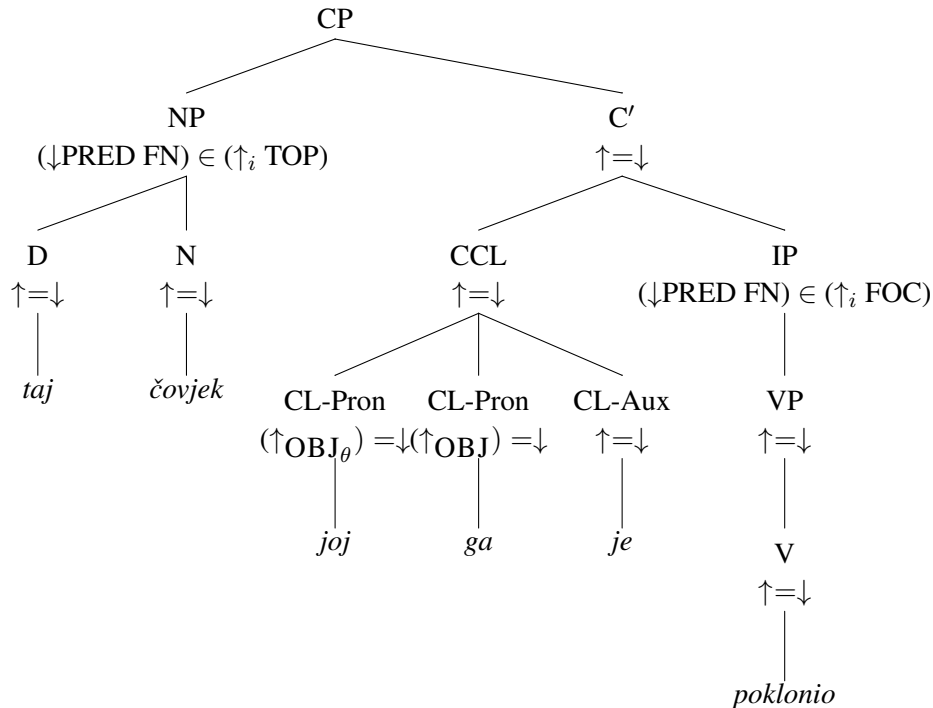


Figure 3: C-structure analysis for (16a): Clitics after the first phrase

In contrast, figure 4 displays the basic analysis for (16b), i.e. the 1W-construction that involves a split constituent in our analysis. The initial constituent in this construction just consists of the sentence initial demonstrative. This construction is only well-formed with a specific intonation contour, i.e. the sentence initial demonstrative receives stress. Thus, in this case the demonstrative is contrastively focused, while the head noun of the subject, *čovjek* ‘man’ is not contributing new information to the sentence, but it might be thematic. The default focus domain is projected by the verb, but it might also include the head noun of the subject.

A full account of the split constructions that includes a discussion of the syntactic structure cannot be provided here because of restrictions of space. Nevertheless, we shall sketch a possible analysis in the following.

The challenges discontinuous constituents pose for syntactic analysis are two-fold. Firstly, an analysis needs to define possible phrase structure rules which define which parts of a constituent can function as phrases on their own. For NSh, for example, a determiner is a possible single constituent of a noun phrase, but a preposition cannot represent a full PP by itself.

Secondly, the analysis has to explain how the single phrases in the c-structure

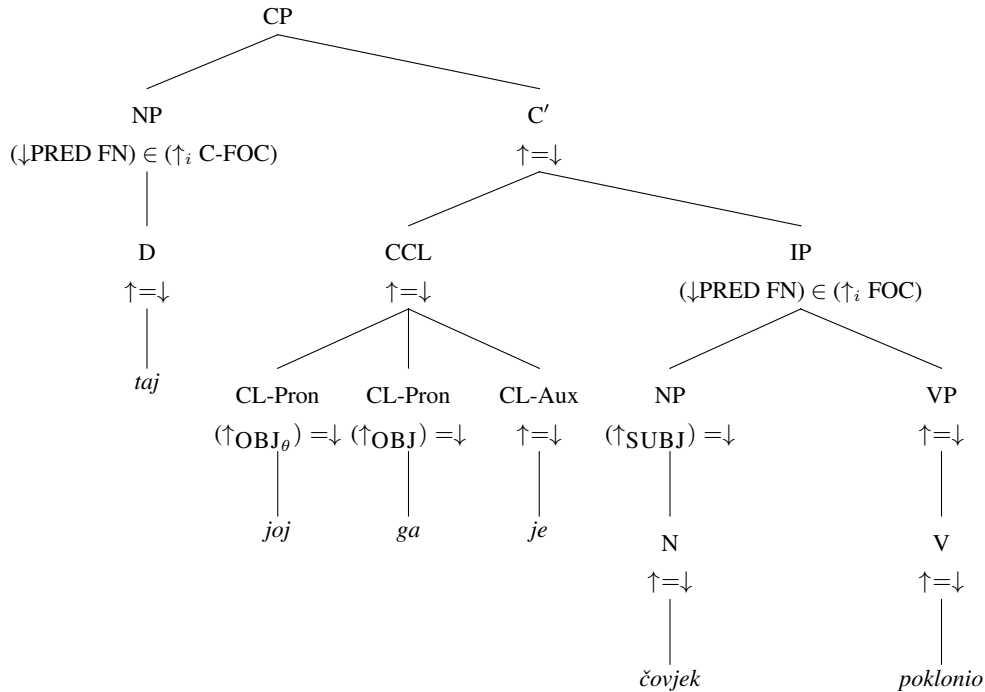


Figure 4: C-structure analysis for (16b): Clitics after the first word

can unify without PRED-value clashes in the f-structure. This question has been addressed within LFG e.g. by Nordlinger (1998) for Australian languages or by Kuhn (1999) for German. Unfortunately, the empirical basis is different in our case.

For the Australian languages discussed by Nordlinger (1998), no distinction exists between nouns and adjectives. Thus, every nominal item can either be used attributively or as a head noun. As a consequence, there are no restrictions on the phrase structure rules of discontinuous constituents.

In German, in contrast, as has been discussed by Kuhn (1999), discontinuous constituents are highly restricted. Only the head noun, or the head noun and its modifier(s) can be preposed in German. This differs from the data in NSh in which the demonstrative or the demonstrative and the nominal modifiers can be preposed. Consequently, Kuhn's (1999) analysis of the German split construction involving elliptical NPs cannot be applied to NSh.

Thus, we propose for the NSh data an alternative analysis in which we treat split-off adjectives and/or demonstratives as headless NP constructions, which in

a way resemble nominalizations, but lack the *pro*-PRED-value they would have when otherwise used alone. This way, we can account for the fact that the individual parts of the split NPs can function as independent NPs while avoiding PRED value clashes in the f-structure.

Reconsidering the examples involving relative clauses in (11), we might add that in our account, (11a) is grammatical because the whole subject NP is in Topic position, whereas in (11b) only the head noun is in topic or contrastive focus position, and the relative clause would have to be assumed to be in the focus domain. This, however, would be information structurally rather odd. Alternatively, if the head noun is in topic position, the relative clause has to be right extra-posed, outside of the topic or focus domain.

Summing up, we propose a purely syntactic analysis for 2PC placement. The clitic cluster always attaches after the first syntactic constituent, which may be a topic or a contrastive focus information-structurally. The first syntactic constituent may either be a complete syntactic phrase, or a split-off part of a syntactic constituent, which then triggers a contrastive focus reading and consequently requires a specific intonational contour. In the next section we turn to even more complex cases and show how the basic analysis laid out here can account for these.

## 6 Further Data and Analyses

In this section we shall discuss more relevant data, which has proven to be difficult for other analyses, in particular the PI-based approaches, and show how our analysis can account for it naturally.

### 6.1 Clitic Third, Fourth ...

The so called sentential or second position clitics do not always have to be realized in second position in the clause (see also (3) for an example of a complex split NP). In (North-)Western NSh variants in Croatia there is a tendency for Clitic-Third (or -Fourth). Syntactic sequences like “XP V CCL ...” are very frequent in these variants. In contrast, (South-)Eastern NSh seems to show a stricter tendency for Clitic-Second, i.e. full phrases seem to be more common in NSh-variants in Dalmatia and West-Herzegovina, while 1W-constructions seem to be more frequently used in the Eastern variants.

The examples in (17) taken from the Croatian Language Corpus (CLC)<sup>5</sup> show constructions in which 2PCs are located in the third or fourth position in the clause. This shows that the topic constituent preceding the clitic cluster can be quite complex.

(17) a. CP V<sub>ptc</sub> CCL ...

---

<sup>5</sup><http://riznica.ihjj.hr>

[ *Da održi koncert* ] [ *pozvao* ] **ga je** *Katolički*  
 that hold concert call-on.ptc him be.3sg catholic  
*pokret za žene*  
 organization of women  
 ‘He was invited by the catholic society for women to arrange a concert.’

b. X NP V.ptc CCL ...

[ *Doduše,* ] [ *hrvatski igrač* ] [ *isprovocirao* ] **ga je**  
 honestly Croatian player needle.ptc him be.3sg  
*startom s leđa*  
 initially from back  
 ‘Honestly, the Croatian player was initially needling him from the back.’

c. C NP NP C NP NP AUX CCL ...

[ *Ali ni jedan ni drugi,* ] [ *dakle ni govoreni ni*  
 but not one not other, that not spoken not  
*pisani jezik* ] [ *nisu* ] **mu mogli biti korisni**  
 written language not-be.3pl him could be helpful  
*neposredno.*  
 directly

‘but not the one nor the other, that is not the spoken nor the written language, could have been directly useful for him’

d. C NP NEG V.imp CCL ...

[ *ali pune glave* ] [ *ne dadoše* ] **mu mira**  
 but full heads not give.imp him piece  
 ‘but the full heads did not give him peace’

As suggested in Cavar and Wilder (1992), constructions like (17a) involve a fronted phrase and a participle head in C preceding the clitic cluster, i.e. so called Long Head Movement constructions. Similarly, the construction in (17b) involves an initial phrase and participle head preceding the clitic cluster, however being preceded by a pragmatic or a discourse element like *Doduše*. The constructions in (17c-d) might involve further positions preceding the clause, which we might assume to be typed as free hanging topics and extra-sentential elements.

As for our arguments and analysis, these examples show that the stipulation of second position clitic placement is problematic. The placement position involves constraints imposed by the possibility of realizing various types of topic and focus elements in the left periphery of the sentence, which again is a syntactic domain, rather than prosodic in nature.



## 6.2 Breakable Clitic Cluster

In this final section we present additional data on clitic clusters which supports the complex syntactic nature view of 2PC placement constraints.

It is often claimed that the clitic cluster is unbreakable in NSh. However, evidence suggests that the clitic cluster is breakable at least in certain syntactic and information theoretic contexts. A detailed analysis of this data, however, is beyond the scope of this paper.

As discussed in the previous sections, clitics usually occur in a syntactic position following contrastively focused or topicalized elements. With respect to various previous examples we mentioned the clitic cluster in NSh, referring to a group of clitics that appear together in the second position in the clause. Numerous clitics can cluster together, and their relative order seems to be subject to placement constraints. The grouping regularities in NSh are described in the slot-model in (18). This seems to be a tendency, not necessarily a strict grammatical constraint, as discussed in Cavar (1999).<sup>6</sup>

(18) *li* – Aux. – Dat.Pron. – Acc.Pron. – Refl.Pron. / *je* (Aux.3sg)

In complex sentences multiple clitic clusters are possible, suggesting that each clause provides a designated structural clitic cluster position, as shown in the example in (19):

(19) *Ivan mu je rekao da mu ga neće dati.*  
Ivan him be.3sg say.ptc that him it not-want.3sg give.inf  
'Ivan told him that he will not give it to him.'

A PI-based approach would have to assume some special handling of multiple-clitic-cluster constructions, and a clause-based application of placement and inversion constraints.

The common assumption is that a clitic cluster cannot be split. The ungrammaticality of examples such as (20b) seems to support this assumption:

(20) a. *Neko dijete mi ga je donjelo.*  
some child me it be.3sg bring.ptc  
'Some child has brought it to me.'  
b. \**Neko mi dijete ga je donjelo.*  
some me child it be.3sg bring.ptc

On the one hand, it seems that a placement stipulation as formulated in the context of PI would have to always uniformly apply to all 2PCs at once in the same way, rather than to individual such clitics. Thus, if one 2PC is placed after the initial syntactic constituent, all of them have to be placed there, i.e. it is not possible to

---

<sup>6</sup>Note that clitic clusters with more than two such clitics are very rare in language use data and corpora, e.g. in the Croatian Language Corpus.

place only one clitic in sentence initial position, and the others in a position after the initial syntactic constituent, otherwise (20b) should be well-formed.

In some contexts like e.g. VP-topicalization constructions, it seems to be possible to split the clitic cluster. In (21b), for example, the two clitics are separated by *auto* ‘car’. This seems to be only possible when a complex ditransitive predicate or verb phrase is topicalized.

- (21) a. *Ivan mu je kupio auto, a ne Stipe.*  
 Ivan him be.3sg buy.ptc car and not S.  
 ‘Ivan has bought him a car, and not Stipe.’
- b. *Kupio mu auto je Ivan, a ne Stipe.*  
 buy.ptc him car be.3sg Ivan and not Stipe.

It seems that the pronominal verbal argument that is realized as a sentential clitic can be located in its complete syntactic and functional complex, i.e. its VP, independent of the other clitic elements. Such an observation is at least problematic for any prosodic 2PC placement model. As mentioned above, on the one hand, in a PI-based approach one would be forced to assume a uniform placement decision for all 2PCs to explain the clustering constraints and the ungrammaticality of examples like (20b). On the other hand, for examples like (21b) one would have to allow for a disjoint initial placement for the 2PCs.

Even more complex problems for PI-based approaches are illustrated by the data on clitic clusters in sentences with infinitival complements. As (22) shows, the clitics that belong to the matrix clause and to the embedded clause can be realized in all possible clitic cluster spots. The direct object clitic of the embedded clause can be realized in the matrix clause as in (22b) or in the embedded clause as in (22c). This could be seen as an instance of optional “clitic raising”. In (22d), the complete embedded clause appears to be topicalized, resulting in two spots for clitic clusters as in the VP-topicalization examples above.

- (22) a. *Ivan je želio čitati knjigu u parku.*  
 Ivan be.3sg wish.ptc read.inf book in park  
 ‘Ivan wanted to read a book in the park.’
- b. *Ivan ju je želio čitati u parku.*  
 Ivan it be.3sg wish.ptc read.inf in park  
 ‘Ivan wanted to read it in the park.’
- c. *Ivan je želio čitati ju u parku.*  
 Ivan be.3sg wish.ptc read.inf it in park
- d. *Čitati ju u parku je Ivan želio.*  
 read.inf it in park be.3sg Ivan wish.ptc

Leaving a detailed explanation and discussion of these constructions to a later time, we might conclude that it appears that, what may find a straightforward syntactic explanation, imposes serious issues and problems for PI-based approaches.

## 7 Conclusion

We have discussed numerous arguments from the literature in favor of a prosodic analysis of 2PC placement in NSh. Given the numerous counterexamples and empirical counter evidence, we can conclude that the concept of prosodic inversion lacks empirical evidence. Thus the proposed extension of theoretical concepts at the prosodic and syntactic level is not motivated.

On the basis of the current empirical evidence we can retain the assumption that clitic placement in the relevant language(s) is syntactic in nature. The fact that some complex phrases can be realized discontinuously, imposes serious theoretical problems in various theories, but also supports the syntactic analysis of 2PC placement.

Due to space restrictions, we were only able to present a brief discussion of the empirical facts, and just sketch a theory of clitic placement and split constructions in NSh. We hope to be able to present an extended and more detailed version in the near future.

## References

- Anderson, Stephen. 2005. *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.
- Bögel, Tina, Butt, Miriam, Kaplan, Ronald M., King, Tracy Holloway and Maxwell III., John T. 2010. Second Position and the Prosody-Syntax Interface. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG 2010 Conference*, pages 106–126, CSLI Publications.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.
- Browne, Wayles. 1976. Two Wh-fronting Rules in Serbo-Croatian. *Južnoslovenski filolog* 32, 196–204.
- Cavar, Damir. 1999. *Aspects of the Syntax-Phonology Interface*. Doctoral dissertation, Universität Potsdam, Potsdam.
- Cavar, Damir and Wilder, Chris. 1992. Long Head Movement? Verb-movement and Cliticization in Croatian. Technical Report, Johann Wolfgang Goethe-Universität, Institut für Deutsche Sprache und Literatur II, Frankfurt a.M.
- Cavar, Damir and Wilder, Chris. 1999. “Clitic Third” in Croatian. In Henk van Riemsdijk (ed.), *Clitics in the Languages of Europe*, volume 20/5 of *Empirical Approaches to Language Typology*, pages 429–467, Mouton de Gruyter.
- Cavar, Malgorzata E. and Cavar, Damir. 2011. Some Cyclicity and Opacity Effects in the Prosody of Two Different Clitic Classes in New-Štokavian Variants. Paper presented at the Ilse Lehiste Memorial Symposium: Melody and Meter.

- Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford: CSLI Publications.
- Cook, Philippa and Payne, John. 2006. Information Structure and Scope in German. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG 2006 Conference*, CSLI Publications.
- Dalrymple, Mary and Nikolaeva, Irina. 2011. *Objects and Information Structure*. Cambridge: Cambridge University Press.
- Diesing, Molly, Filipović-Durđević, Dušica and Zec, Draga. 2009. Clitic Placement in Serbian: Corpus and Experimental Evidence. In Sam Featherston and Susanne Winkler (eds.), *The Fruits of Empirical Linguistics*, pages 61–75, Berlin: Mouton de Gruyter.
- Fanselow, Gisbert and Cavar, Damir. 2001. Remarks on the Economy of Pronunciation. In Gereon Müller and Wolfgang Sternefeld (eds.), *Competition in Syntax*, pages 107–150, Berlin: De Gruyter.
- Fanselow, Gisbert and Cavar, Damir. 2002. Distributed Deletion. In Artemis Alexiadou (ed.), *Theoretical Approaches to Universals*, pages 65–07, Amsterdam: Benjamins.
- Franks, Steven and King, Tracy Holloway. 2000. *A Handbook of Slavic Clitics*. Oxford: Oxford University Press.
- Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. Stanford, CA: CSLI Publications, republished PhD-thesis.
- Halpern, Aaron and Zwicky, Arnold. 1996. *Approaching Second – Second Position Clitics and Related Phenomena*. Stanford: CSLI Publications.
- King, Tracy Holloway. 1995. *Configuring Topic and Focus in Russian*. Stanford: CSLI Publications.
- King, Tracy Holloway. 1997. Focus Domains and Information Structure. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG 1997 Conference*, CSLI Publications.
- Kuhn, Jonas. 1999. The Syntax and Semantics of Split NPs in LFG. In Francis Corblin, Carmen Dobrovie-Sorin and Jean M. Marandin (eds.), *Empirical Issues in Formal Syntax and Semantics 2, Selected Papers from the Colloque de Syntaxe et Sémantique à Paris (CSSP 1997)*, pages 145–166, The Hague: Thesus.
- Mchombo, Sam, Morimoto, Yukiko and Féry, Caroline. 2005. Partitioning Discourse Information: a Case of Chicheŵa Split Constituents. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG 2005 Conference*, CSLI Publications.

- Mycock, Louise. 2005. Wh-in-situ in Constituent Questions. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG 2005 Conference*, CSLI Publications.
- Mycock, Louise and Dalrymple, Mary. 2011. The Prosody-Semantics Interface. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG 2011 Conference*, CSLI Publications.
- Nakanishi, Kimiko. 2007. Measurement in the Nominal and Verbal Domains. *Linguistics and Philosophy* 30, 235–276.
- Nordlinger, Rachel. 1998. *Constructive Case: Evidence From Australian Languages*. Stanford: CSLI Publications.
- Obenauer, Hans Georg. 1976. *Etudes de syntaxe interrogative du français: quoi, combien et le complémenteur*, volume 34 of *Linguistische Arbeiten*. Max Niemeyer Verlag.
- O'Connor, Robert. 2002. The Placement of Enclitics in Bosnian, Croatian and Serbian, unpublished ms.
- O'Connor, Robert. 2004. *Information Structure in Lexical-Functional Grammar: The Discourse-Prosody Correspondence in English and Serbo-Croatian*. Ph.d. thesis, University of Manchester.
- Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: Comp as the Second Position. In Aaron Halpern and Arnold Zwicky (eds.), *Approach Second: Second Position Clitics and Related Phenomena*, pages 411 – 428, Stanford: CSLI Publications.
- Radanović-Kocić, Vesna. 1988. *The Grammar of Serbo-Croatian Clitics: A Synchronic and Diachronic Perspective*. Ph.d. thesis, University of Illinois.
- Radanović-Kocić, Vesna. 1996. The Placement of Serbo-Croatian Clitics: A Prosodic Approach. In Aaron Halpern and Arnold Zwicky (eds.), *Approach Second: Second Position Clitics and Related Phenomena*, pages 429–445, Stanford: CSLI Publications.
- Schütze, Carson T. 1994. Serbo-Croatian Second Position Clitic Placement and the Phonology-Syntax Interface. In Andrew Carnie and Heidi Harley (eds.), *MIT Working Papers in Linguistics 21, Papers on phonology and morphology*, pages 373–473, MIT.
- van Riemsdijk, Henk (ed.). 1999. *Clitics in the Languages of Europe*, volume 20.5 of *Empirical Approaches to Language Typology*. Mouton de Gruyter.
- Zec, Draga and Inkelas, Sharon. 1990. Prosodically Constrained Syntax. In Sharon Inkelas and Draga Zec (eds.), *The phonology-syntax connection*, pages 365–378, Chicago: University of Chicago Press.