

Clitic placement in Serbian: Corpus and experimental evidence

Molly Diesing^a, Dušica Filipović Đurđević^b, Draga Zec^a

^aCornell University, ^bUniversity of Novi Sad

md20@cornell.edu dmfilipo@f.bg.ac.yu dz17@cornell.edu

1. Introduction

We focus on the placement of clitics that occupy the so-called “second” position, characterized cross-linguistically as the sentential position either (i) after the first word, or (ii) after the first constituent. We present a relatively complex case, that of Serbian, in which both (i) and (ii) can serve as sentential positions for clitics (Browne 1975, Zec and Inkelas 1990). Moreover, we found that in Serbian there is more than one type of first position, both in the case of first word, and in the case of first constituent. We found two types of cases depending on whether the sentence initial element is, or belongs to, either an argument or the predicate, which yields a four part classification illustrated in (1)-(2). The native speakers we consulted suggest that, while all four cases are acceptable, (1a) and (2a) are more “common” and less “marked” than (1b) and (2b).

- | | | | |
|-----|----|--|--|
| (1) | a. | <i>Taj zadatak je veoma važan.</i>
that task is-Cl very important
'That task is very important.' | Argument, 1 st constituent |
| | b. | <i>Taj je zadatak veoma važan</i>
that is-Cl task very important
'That task is very important.' | Argument, 1 st word |
| (2) | a. | <i>Veoma je važan taj zadatak.</i>
very is-Cl important that task
'That task is very important.' | Predicate, 1 st word |
| | b. | <i>Veoma važan je taj zadatak</i>
very important is-Cl that task
'That task is very important.' | Predicate, 1 st constituent |

The classification in (1)-(2) provides a different perspective from previous analyses, which account for only two of the four cases we identified: the two argument cases in (1) (as summarized by Anderson 2005 and works cited therein). The predicate cases have not been addressed in the literature. Moreover, the body of judgments reported in the literature is not consistent, due both to differences in speaker judgments, and insufficient control for differences among regional dialects.

We hypothesize that all four cases are attested, yet expect differences in distribution and markedness. We explored this hypothesis by two methods: corpus investigation, and psycholinguistic experiments.

2. First phase: corpus investigation

In the first phase of our project, we collected data from two corpora of the Serbian language, one based on daily press and the other on literary prose (in Laboratory for Experimental Psychology, Department of Psychology, University of Belgrade). We analyzed 2993 sentences with clitics, 1323 from the corpus of daily press and 1670 from the corpus of literary prose. We placed each sentence in one of the four classes in (1)-(2). Our results are shown in the table in (3):

(3)

		Daily press 1323 sentences	Fiction 1670 sentences	Daily press Proportions	Fiction Proportions
a.	Arg- 1Word	7	40	0.53	2.45
b.	Arg- 1Phrase	762	860	57.60	52.60
c.	Pred- 1Word	549	733	41.50	44.83
d.	Pred- 1Phrase	5	2	0.38	0.12

In the argument case, (3)a-b, we see a large proportion of the first phrase sentences, and a small proportion of the first word sentences. The situation is reversed in the predicate case, (3)c-d, where we find a large proportion of the first word sentences and a miniscule proportion of the first phrase sentences. The results we got are striking in several respects. We found support for all types of cases in (1)-(2). More importantly, we found that the two types of cases, arguments and predicates, have different preferred positions for clitics: the preferred position for clitics in the argument case is after the first constituent, and in the predicate case, after the first word. In the second phase of our project, we tested this result by conducting a series of psycholinguistic experiments.

3. Second phase: psycholinguistic experiments

We conducted two psycholinguistic experiments, both based on a set of 120 sentences, specifically designed for this purpose. The sentences include two sets, 60 in each, one for the argument and the other for the predicate case. In experiment 1 we were interested in the differences between the two possible clitic positions in argument and predicate sentences manifested in language production. Sentences were printed with critical clitics omitted, and the two possible positions of clitics replaced with a blank to be filled in, as in (4):

(4) *Njegov ___ auto ___ najbrži u gradu.*
/His ___ car ___ fastest in the city./

The task was to be performed in such a way as to make the sentence sound as natural as possible. Analysis of the responses obtained from the participants revealed a dramatic difference between

clitic positions in the two sentence categories. While 92.98 % of participants placed a clitic after the first phrase in argument sentences, only 2.41% of participants placed a clitic after the first phrase in predicate sentences. The observed difference was significant: $\chi^2(1) = 1874.121$, $p < 0.01$.

In experiment 2 we investigated these differences at the level of language perception, or processing. One-hundred-and-twenty target sentences from experiment 1 were presented as stimuli in a sentence grammaticality judgment task. The participants were given instructions to judge whether the sentence appearing on the screen was acceptable in their language. They were told to base their answers on their intuitions as native speakers, and that there would not necessarily be right or wrong answers. Sentences were presented one-by-one, in a random order, on a computer screen. Prior to each sentence a fixation point was presented for 2000 ms. A sentence would remain on the screen until participant's response, but its duration was limited to 8 seconds. The participant's reaction time was measured. All analyses were conducted on the responses marking the acceptance of the target sentences. A by-participant analysis of variance of reaction time revealed a significant main effect of sentence type: $F(1, 41) = 19.745$, $p < 0.01$ ($F(1, 118) = 3.200$, $p = 0.08$, by item). Predicate sentences elicited shorter processing latencies. There was no main effect of clitic position (although $F(1, 118) = 6.031$, $p < 0.05$, by item), but there was a significant interaction between sentence type, and clitic position: $F(1, 41) = 25.644$, $p < 0.01$ ($F(1, 118) = 94.744$, $p < 0.01$, by item). Argument sentences with a clitic positioned after the first phrase were processed faster than argument sentences with a clitic positioned after the first word, while predicate sentences with a clitic positioned after the first word were processed faster than predicate sentences with a clitic positioned after the first phrase.

4. Conclusion

We interpret these results as clearly establishing preferred clitic placement in the two types of sentences. All four types are represented both in the investigated corpora and in the production and perception patterns albeit in very different proportions. In the psycholinguistic experiments, we observed a striking asymmetry between the two cases. We attribute these differences to different discourse conditions between the first word and first phrase positions within each category. This in turn allows an approach to optionality in grammar based on both structural and discourse factors. Future work will investigate the syntax/prosody interface properties of these cases, with special attention to intonational contours and syntactic structures associated with each of the four types.

References

- Anderson, S. (2005). *Aspects of the Theory of Clitics*. Oxford University Press, Oxford, UK.
- Browne, E.W. (1975) Serbo-Croatian enclitics for English-speaking language learners. *Contrastive Analysis of English and Serbo-Croatian*, 1: 105-134.
- Zec, D. and S. Inkelas (1990) Prosodically constrained syntax. In S. Inkelas and D. Zec, eds., *The Phonology-Syntax Connection*, pp. 365-378.