Clitic doubling in a determinerless language with second position clitics

FRANC MARUŠIČ & ROK ŽAUCER

1 Introduction

It is typically the case that if a language exhibits clitic doubling, it has definite articles (see Bošković 2008a,b,c) and it has verb-adjacent clitics. In this paper, we discuss the case of certain western dialects of Slovenian around Nova Gorica/Gorica/Gorizia, which we subsume under the cover term ‘Gorica Slovenian’. These dialects are interesting in that they appear to defy these generalizations. Whereas they have clitic doubling, their clitics are not verb-adjacent but second-position/Wackernagel clitics. Moreover, these dialects do not have a definite article. In what follows, we discuss how both of these facts are relevant for the claims that have been made in the literature on the basis of the above-mentioned generalizations.

In section 2, we first introduce the phenomenon of clitic doubling, then quickly show (following Marušič & Žaucer 2008) that the phenomenon under investigation is indeed proper clitic doubling rather than some other sort of doubling (e.g. clitic left dislocation, right dislocation, etc.), and then we briefly mention some other characteristics of clitic doubling in Gorica Slovenian. In section 3, we review a major view about clitic doubling in Bulgarian and Macedonian, the only other Slavic languages that are known to have clitic doubling, and explain why it cannot be extended to Gorica Slovenian. In section 4, we discuss the implications of these data on Bošković’s (2008a,b,c) generalization about clitic doubling.

Unless marked otherwise, the Gorica Slovenian data come from spontaneous speech, examples marked as ‘J(udged)G(orica)’ were constructed and tested against several native speakers of Gorica Slovenian, whose help we gratefully acknowledge. We do not mark the ungrammatical examples as judged, even though they obviously are.
2. Clitic doubling in Gorica Slovenian

The doubling of a verb’s argument by a clitic is known from many languages. Among the languages that are geographically close to our Slovenian dialects, these include certain Northern Italian dialects, Romanian, Bulgarian, Macedonian, South Serbian dialects, Greek and Albanian. A typical example of clitic doubling is shown in (1), and a comparable example from Gorica Slovenian is given in (2).

(1) Na mene ne mi e studeno.       Bulgarian to me not I.DAT is cold ‘I am not cold.’ (Franks & Rudin 2004: (3a), p. 106)

(2) Meni mi ni mraz.            I.DAT I.DAT neg-aux3sg cold ‘I am not cold.’

To some extent, this phenomenon appears to be exhibited by most or all western Slovenian dialects, but since there are some differences among the dialects, we limit ourselves to the dialects around Nova Gorica/Gorica/Gorizia.

2.1 Not CLLD, RD or Appositive D

Since clitic doubling can be easily mistaken for clitic left dislocation, right dislocation or appositive dislocation, we have to emphasize that the phenomenon under investigation is not one of these.

As shown in (3), the doubled DP (in (3) the full pronoun mene) does not have to be either first or last in the sentence; this shows that this is neither a case of clitic left dislocation nor right dislocation. The clitic doubled DP also does not have to be marked off with comma intonation, which shows that this is not a case of appositive doubling.

(3) Ma to me mene ne briga.       but this me.GEN me.GEN not cares ‘But I don’t care about this.’

This demonstrates, albeit in a very sketchy and somewhat simplifying manner, that the construction under investigation is indeed a case of clitic doubling rather than one of the other ‘doubling’ phenomena known in the
literature. For a fuller demonstration and some more discussion, we refer the reader to Marušič and Žaucer (2008).1

2.2 The properties of clitic doubling in GoS

Perhaps the most prominent feature of clitic doubling in Gorica Slovenian is that it is limited in a specific sense. It can only happen when the doubled verbal argument is represented by a full pronoun; in other words, a clitic never doubles a non-pronominal argument (whether a common noun or a proper name), as shown in (4).

(4) a. Js se ga njega spomnem še iz šole.
   I refl himACC himACC remember already from school
   ‘I remember him already from school.’

b. *Js se ga Petra spomnem še iz šole.
   I refl himACC Peter ACC remember already from school
   ‘I remember him already from school.’

But whereas clitic doubling is restricted to pronominal arguments, it is available in any person, number and case where the pronouns have a clitic and a non-clitic variant (i.e. genitive, dative, accusative). Here we give examples for all persons and numbers (from Marušič and Žaucer 2008). Examples in (5) are for singular (all taken from spontaneous speech) and examples in (6) for plural persons.

(5) a. Mi lahko daste kar meni?
   Me DAT possible give2.PL ptcl meDAT
   ‘Can you give it to me?’

b. Ma kaj tebe te ne zanima, kako bo šlo končat?
   but Q youGEN youGEN not interest how will go end
   ‘Don’t you want to know how it will end?’

c. Js bi ga njega peljala domov prej.
   I NOM would himACC himACC drive home first
   ‘I would first take him home.’

1 It is hard to establish whether the clitic doubling in Gorica Slovenian could be said to be obligatory, since all speakers are under the influence of non-doubling dialects as well as of non-doubling standard Slovenian (either due to education or everyday usage, and most often both).
(6) a. Peter nam nám ni tou prnest neč za pit. JG
   Peter usDAT usDAT not want bring nothing for drink
   ‘Peter didn’t want to bring us anything to drink.’

   b. Vás si vas ne upam neč prašt. JG
      youGEN refl youGEN not dare nothing ask
      ‘I dare not ask you anything.’

   c. Lahko jih pa njih vpraša.
      possible themACC ptcl themACC ask
      ‘He can ask them.’

It might be just impossible to find spontaneous examples of a double object construction with two full pronouns. But a sentence like (7) was judged possible by our informants

(7) Meni mi ga njega niso teli predstavit. JG
   MeDAT meDAT himACC himACC not want introduce
   ‘They didn’t want to introduce him to me.’

To summarize, clitic doubling in Gorica Slovenian is restricted to pronouns. It is not possible to double a regular NP argument, but it is at the same time completely general, since all strong pronouns with a clitic counterpart, regardless of person/number/case, can be doubled, including the reflexive pronoun.

3. A possible analysis

Franks & King (2000) analyze clitic doubling in Bulgarian and Macedonian as verb agreement (we limit ourselves to testing only the mainstream analysis of clitic doubling in Slavic). Bulgarian and Macedonian clitics are verb-adjacent, so that this analysis makes perfect sense. But this kind of analysis makes sense only if clitics under discussion are verb adjacent. It cannot be used on Gorica Slovenian clitics since clitics in western Slovenian dialects are Wackernagel clitics, as shown in examples (8) (partially repeated from above).

(8) a. Js se ga njega spomnim še iz šole.
    IsOM refl himACC himACC remember still from school
    ‘I remember him already from school.’

b. Js se ga njega dobro spomnim še iz šole.JG
   IsOM refl himACC himACC well remember still from school
   ‘I remember him well already from school.’
c. Lahko jih pa njih vpraša.
   possible them_{ACC} pcl_them_{ACC} ask
   ‘He can ask them.’

d. Zato me mene to moti.
   because me_{ACC} me_{ACC} this bothers
   ‘That is why this bothers me.’

e. Zato me mene to zelo moti.  JG
   because me_{ACC} me_{ACC} this very bothers
   ‘That is why this bothers me a lot.’

f. Kdo me je mene udaru?
   who me_{ACC} aux me_{ACC} hit
   ‘Who hit me?’

None of the clitic clusters in (8) appear next to the verb. In (8a) and (8f)
the clitic and the verb are separated by the pronominal argument – the
object of the clause, in (8b) by the pronominal argument and an adverb, in
(8c) by the pronoun and an additional topic particle, in (8d) by the prono-
nominal argument and the object of the clause and in (8e) by an addi-
tional adverb.

Clitics were in the second position also in all other examples we have
encountered. Sentences where clitics are not in the second position, like
those in (9), are judged ungrammatical by speakers of Gorica Slovenian.

(9)  a.*Js njega dobro se ga spomnim še iz šole.
    I NOM him_{ACC} well refl him_{ACC} remember still from school
    (intended: ‘I remember him well already from School.’)

   b.*Js dobro se ga njega spomnim še iz šole.
    I NOM well refl him_{ACC} him_{ACC} remember still from school
    (intended: ‘I remember him well already from School.’)

   c.*Zato mene to zelo me moti.
      because me_{ACC} this very me_{ACC} bothers
      (intended: ‘That is why this bothers me a lot.’)

Therefore with respect to their position clitics in Gorica Slovenian appear
to be regular Wackernagel clitics, comparable to clitics in Standard
Slovenian, for which see Golden and Sheppard (2000), Golden (2003)
and Marušič (2008). For this reason any analysis of clitic doubling as
verbal agreement cannot be applied.
4. Bošković’s (2008a, 2008b, 2008c) generalization

On the basis of a number of typologically distinct languages, Bošković (2008a,b,c) proposes 10 generalizations, all linking the presence/absence of a definite article and another characteristic. One of these relates definite articles with clitic doubling saying that only languages with definite articles can have clitic doubling. In other words, there should be no language with clitic doubling and no definite article.

Gorica Slovenian seems to be just such a language. It has clitic doubling, but just like other varieties of Slovenian, it does not have a definite article (cf. Toporišič 2000). Just like central Slovenian dialects, Gorica Slovenian has the so-called “adjectival definite article ta”, but this is not a definite determiner, as can be most clearly seen in examples like (10) where ta occurs in an indefinite noun phrase with an overt indefinite determiner (see Marušič & Žaucer 2006, 2007 for discussion).

(10) ene ta velike plošče Central Slovenian/JG
    one PL TA big records
    ’some LPs’ (cf. English ’some (*the) LPs’)

Gorica Slovenian also matches central dialects in having the indefiniteness element en (literally ‘one’), given also in (10) above, but the presence of such indefiniteness elements (which may also turn out to be adjectives) is not relevant for the generalization anyway (Bošković 2008b, but cf. also an alternative speculation in Bošković 2008b: fn. 23)

In short, Gorica Slovenian, which has clitic doubling and no definite determiner, seems to be problematic for Bošković’s otherwise robust generalization.

4.1 Reconciling Gorica Slovenian with Bošković (2008a,b,c)?

In this section we discuss some options that could explain the problematic data. But first let’s go back to Bošković (2008a,b,c). The generalizations he proposes make a clear distinction between languages with and languages without a definite article. He proposes that the difference is not just lexical but rather structural. For him, only languages with a definite article have the DP projection as shown in (11). Difference between NP and DP languages extends also to the way adjectives are merged into the structure since in NP languages adjectives are in Spec.NP, while in DP languages they take NP as their complement.

(11) a. [DP D [AP A [NP N]]] → DP languages (e.g. English)
    b. [NP AP N] → NP languages (e.g. Serbo-Croatian)
The availability of clitic doubling is linked to the presence of the DP phrase also because clitic doubling is typically assumed to bring in specificity. Even though we did not discuss semantics of clitic doubling in Gorica Slovenian, the specificity requirement is trivially satisfied since clitic doubling is limited to personal pronoun, which are always specific (cf. Marušič & Žaucer 2008). Given all these, Gorica Slovenian should have the structure in (11a), but as we said, this dialect lacks a definite article.

One option is to say that Gorica Slovenian has the definite determiner, but that it is null. That would mean that Slovenian would have the structure in (11a) with the DP topping the nominal phrase rather than the one given in (11b), where the topmost projection in a nominal phrase is NP (Bošković presents this as the weaker version of his claim regarding the structural differences between languages with definite articles and languages without definite articles).

This option is problematic for the following reason. If there is a structural difference between Gorica Slovenian and Serbo-Croatian, this structural difference should be observable in all relevant constructions. This would mean that Gorica Slovenian should behave as a DP-language according to all applicable characteristics from Bošković (2008a,b,c). Since not all generalizations are two way generalizations, only a subset of them makes predictions following the presented hypothesis about Gorica Slovenian. We will now go through the relevant generalizations proposed in Bošković (2008a,b,c) and test whether this is indeed the case (this is partly done in Bošković (2008b) for standard Slovenian).

Bošković (2008a,b,c) proposes that languages without articles disallow neg-raising and those with articles allow it. Following this generalization, if Gorica Slovenian has a null article, it should allow neg-raising. But as shown in (12), matrix negation has no effect on the negation sensitive elements in the embedded clause. In (12) genitive of negation is impossible, even though it is generally obligatory in clauses with negation. Matrix negation affects direct objects in an embedded non-finite clause, (12b). As (13) shows, negative concord elements are also impossible even though they are also available inside infinitives under matrix negation. (14) shows that matrix negation does not license an embedded NPI (see Bošković 2008b for the standard Slovenian counterpart). And even semantically, matrix negation, cannot be in a regular sentence, matrix negation cannot be understood inside the embedded clause

(12) a. Vid ne misli, da Peter je fige/ *fig. JG
   Vid neg think that Peter eats figsACC figsGEN
   ‘Vid doesn’t think that Peter eats figs.’
b. Vid neg want eat figs
‘Vid didn’t want to eat figs.’

(13) a.*Vid neg think that aux Peter never nothing eat
(intended: ‘Vid doesn’t think that Peter ate nothing’)

b. Vid neg want never nothing eat
‘Vid didn’t want to eat anything at any time.’

(14) *Ne misli, [da jo je vidla že narmajn dve leti]
‘He doesn’t think that she has seen her in at least two years.’

Secondly, Bošković (2008a,b,c) claims that only languages without articles may allow adjunct extraction out of traditional noun phrases. Again, as shown in (15), Gorica Slovenian does allow adjunct extraction (at least to some degree) and thus patterns with other NP-languages. Following Bošković (2008a,b,c) this should mean that it does not have a definite article.

(15) a. Od kirga autorja je Peter prebral vse knjige?
‘For which author is it true that Peter read all of his books?’

b. Iz kirga kluba praviš, da je Peter srečal vse tipe?
‘For which club do you claim that Peter met all guys.’

Thirdly, Bošković (2008a,b,c) claims that only languages with articles allow transitive nominals with two genitives, as in German example (16), whereas languages without definite articles do not allow transitive nomi-

---

2 As already mentioned in Bošković (2008b), data of this type are tricky. So for example, whereas (15a) and (15b) are fine, minimally different (i) and (ii) are not (neither in Gorica Slovenian nor in Standard Slovenian).

(i) *Od kirga avtorja je Peter včeraj bral najboljšo knjigo?
‘For which author is it the case that Peter read his best book?’

(ii)? Iz kirga kluba je Peter pretepel vse tipe?
‘For which club is it the case that Peter beat all guys?’
nals with two (non-lexical) genitives, Gorica Slovenian, just like standard Slovenian (and Serbo-Croatian) does not allow such constructions, (17).

\[(16) \quad \text{Hannibals Eroberung Roms} \quad \text{German}\]
\[
\text{Hannibal’s conquest of Roma} \quad \text{(from Bošković 2008b)}
\]

\[(17) \quad \text{a.}^\ast \text{zavzetje Rima Hanibala} / {^\ast \text{zavzetje Hanibala Rima}}
\quad \text{conquest Roma}_\text{GEN} \text{ Hannibal}_\text{GEN}
\]
\[
\text{b.}^\ast \text{petje Zdravljice Toneta}
\quad \text{singing Zdravljica}_\text{GEN} \text{ Toneta}
\]
\[
\text{(intended: ‘Hanibal’s conquest of Rome.’)}
\]

Fourthly, independently of Bošković (2008a,b,c), Živanović (2008) explicitly links the presence/absence of the majority reading to the presence/absence of a definite article (based on a substantial sample of languages). Gorica Slovenian, (18a), patterns with standard Slovenian, given in (18b), in not allowing the majority reading known from e.g. English, (19).

\[(18) \quad \text{a. Narveč tipov pije vino.} \quad \text{JG}\]
\[
\text{most guys drink wine}
\]
\[
\text{‘More guys drink wine than any other drink.’}
\]
\[
\text{b. Največ ljudi pije pivo.} \quad \text{Standard Slovenian}\]
\[
\text{most people drink beer}
\]
\[
\text{‘More people drink beer than any other alcoholic drink.’}
\]

\[(19) \quad \text{Most people drink beer.} = \text{More than half of people drink beer.}
\]

If this correlation indeed holds both ways, the absence of the majority reading in Gorica Slovenian means that Gorica Slovenian cannot have a definite article, not even a null one.

Fifthly, Bošković (2008a,b,c) proposes that only languages without articles may allow scrambling. Here again, Gorica Slovenian seems just as flexible as Standard Slovenia, as should be obvious also from many of the examples given above and also from (20) below, where the direct object of the non-finite complement is scrambled to the front of the sentence.

\[(20) \quad \text{a. Mene me ne morte zjebat.}
\]
\[
\text{I} \text{ACC neg} \text{ can} \text{ fuck-up}
\]
\[
\text{‘You cannot fuck me up.’(a line by the local rapper Valterap)}
\]
Bošković also gives a generalization linking the availability of left branch extraction to the presence/absence of the definite determiner, but here the data both in Gorica Slovenian and standard Slovenian are really not clear, as pointed out also by Bošković (2008b), so we will not discuss this any further.

To summarize, Gorica Slovenian patterns with Standard Slovenian with all relevant generalizations, except that it has clitic doubling, and both Standard Slovenian and Gorica Slovenian mostly pattern with Serbo-Croatian. This should mean that both Gorica and Standard Slovenian are NP-languages. In turn this should mean that the possibility of analyzing Gorica Slovenian as having a null definite article is not feasible. Therefore the fact that Gorica Slovenian has clitic doubling remains problematic for Bošković’s generalization.

4.2 A Change in Progress?

In a footnote Bošković (2008b, fn. 23) also mentions the option that the situation in Slovenian, where some of the characteristics diagnose the language as an NP language less clearly than is the case in Serbo-Croatian, could reflect a change in progress, i.e. a change from a typical NP language to a DP language. For example, it might be possible that the structure of Slovenian and Gorica Slovenian nominal phrases has begun to grow, with the lowest parts of the extended projection of the noun (but not all of it) having already been added on top of the NP.

Potential support for this could be seen in the fact that Slovenian (as mentioned under 4. above) has an indefinite article, which is also true for Gorica Slovenian, as shown in (21) (repeated from (10) above).

(21) ene ta velike plošče Central Slovenian/JG onePL.TA big records

‘some LPs’ (cf. English ‘some (*the) LPs’)

If we assume that the indefinite article sits in some projection lower than the DefP/DP (in the spirit of Julien 2005), one could claim that the fact that Gorica and standard Slovenian have an indefinite article but not a definite article suggests that the structure of the nominal phrase in these systems indeed does contain a part of the extended projection of the noun, but only its lower part. That is, the structure of the Gorica and standard
Slovenian noun phrase would not be like that of Serbo-Croatian (i.e. as in (11b) above). Rather, its NP would be dominated by the functional projection which the indefinite article is associated with, but that would be where the functional structure of Slovenian nominal phrases stops, so it is not topped off by a DefP/DP.

Furthermore the presence of this additional functional structure in Gorica and standard Slovenian nominal phrases could mean that adjectives are merged into the nominal phrase as in DP languages rather than as in NP languages. This could also explain the relative unavailability of Left Branch extraction in Slovenian. Though at the same time, if the position of the adjective in Slovenian and Gorica Slovenian is comparable to English, then left branch extraction should be just as bad in Slovenian as it is in English, but this is actually not the case (as discussed at some length in Bošković 2008b).

However, if the existing but deficient functional structure above NPs in Gorica and standard Slovenian is supposed to explain the non-perfectly-clear situation with respect to the NP/DP characteristics, it simply cannot explain the presence of clitic doubling in Gorica Slovenian, assuming, with Bošković, that clitic doubling can only occur in the presence of DP.

Clitic doubling in Gorica Slovenian is limited to personal pronouns, which are standardly analyzed as Ds. So at least for pronouns, doubling could still be analyzed as linked to the presence of the DP projection. But if we accept that personal pronouns are DPs in Slovenian, they should probably be DPs universally, also in Serbian, but that only means that the existence of clitic doubling limited to personal pronouns might not affect the generalization, which should probably be restated as (22).

(22) Only languages with articles may allow clitic doubling of non-pronominal arguments.

5. Conclusion

Given the fact that Gorica Slovenian has clitic doubling of strong pronouns in Accusative, Genitive and Dative case in all numbers and persons, we can conclude that clitic doubling is not restricted to languages

---

5 If we accept that personal pronouns are DPs, we should also analyze as DPs demonstratives when used as pronominal elements, e.g. when used to refer to persons as in (i). But such demonstrative pronouns cannot be doubled in Gorica Slovenian.

(i) Včera sem (*ga) tega to videl.
     yesterday aux him this here see
     'I have seen this guy yesterday.'
with verb-adjacent clitics and overt definite articles. Gorica Slovenian thus presents a bit of a problem for the generalization made in Bošković (2008a,b,c), which says that clitic doubling only exists in languages with overt definite determiners. This generalization can only be saved if we modify it to exclude clitic doubling of personal pronouns. The generalization is not as strong as before, but it is still factual.

We leave for future research the detailed investigation of the other clitic-doubling western Slovenian dialects, including the well documented dialect of Rezija, and the variation between these dialects that is briefly described in Marušič and Žaucer (2008). Also interesting is the question how this syntactic phenomena is related to clitic doubling in the neighboring Romance languages: e.g. Friulian (see Erat 2006) and Triestino (Pinguentini 1984).

References
Bošković, Željko. 2008a. What will you have, DP or NP?. In E. Elfner & M. Walkow (eds.) NELS 37: Proceedings of the 37th Annual Meeting of the North East LS. Amherst, MA: GLSA.


Ramovš, Fran. 1957. Karta slovenskih narodov v priročni izdaji. Ljubljana: Cankarjeva založba
