UNAGREEMENT BETWEEN ITALIAN
AND SOUTHERN ITALIAN DIALECTS[1]

Georg F.K. Höhn
University of Cambridge

Giuseppina Silvestri
University of Cambridge

Maria Olimpia Squillaci
University of Cambridge

Abstract: *This paper discusses new data from two southern Italian dialects, northern Calabrese and southern Calabrese, which contrast with standard Italian in exhibiting the flexible agreement patterns with definite plural subjects characteristic of the unagreement phenomenon of Spanish and Modern Greek. We highlight a problem that this raises for recent proposals connecting the availability of unagreement to the obligatory presence of a definite article in adnominal pronoun constructions like "we children". Adopting the distributed morphology framework, we suggest that the contrast between standard Italian and the southern varieties may be due to variation in the feature specification of the vocabulary items realising the D head.*

Keywords: *unagreement, pronominal determiners, demonstratives*

1. Introduction

This paper presents new data concerning a contrast between standard Italian, where definite plural subjects cannot co-occur with non-third person agreement as illustrated in (1), and the availability of such constructions in two southern Italian dialects (SIDs), northern Calabrese of Verbicaro (Cosenza) and southern Calabrese of Bova Marina (Reggio Calabria), illustrated in (2) and (3).

(1)   *I bambini giochiamo.   [Standard Italian]
          the children play.1PL
          intended: 'We children are playing.'

---

(2)   I quatrarə iucamə.   [Northern Calabrese]
        the children play.1PL
        'We children are playing.'

(3)   I figghioli iocamu.   [Southern Calabrese]
        the children play.1PL
         'We children are playing.'

Taking the data from the SIDs to be instances of unagreement (Hurtado 1985), we argue that they pose a problem for recent analyses of this phenomenon by Choi (2013, 2014) and Höhn (in press). Both accounts postulate a correlation between the availability of unagreement and the presence of a definite article in adnominal pronoun constructions (APCs), e.g. Spanish *nosotros los estudiantes* 'we students'. The SID data discussed here appear to defy that correlation.

Adopting the framework of distributed morphology (Halle & Marantz 1993, Embick 2010), we argue that the contrast in (un)agreement patterns between standard Italian and non-standard southern varieties is related to a difference in whether personal pronouns carry a demonstrative feature. The lack of unagreement in standard Italian arises because personal pronouns are underspecified for demonstrativity and compete for insertion with the vocabulary items for the definite article as proposed in Höhn (2016). Due to the difference in the vocabulary, no such competition arises in the SIDs.

The paper is articulated as follows. Section 2 introduces the unagreement phenomenon and sketches the previous analyses. Relevant data from northern and southern Calabrese will be described in section 3 and, in the light of this evidence, our proposed analysis will be put forward in section 4. Section 5 concludes the paper.

2. Unagreement

This section introduces the unagreement phenomenon (Hurtado 1985) and outlines the analyses proposed for it by Choi (2013, 2014) and Höhn (2016). Definite plural subjects in languages like Standard Modern Greek (SMG), Spanish, Catalan and Bulgarian are compatible with all three plural forms of the verb, as illustrated in (4) for SMG.

(4)   Oi foitites diavaz(-oume/-ete/-oun) kathe mera.   [Standard Modern Greek]
        the students study-1PL/-2PL/-3PL every day
        'We/you/the students study every day.'

This conflicts with the common assumption that definite noun phrases such as *oi foitites* 'the students' are inherently third person, which would predict only third person verbal agreement to be allowed. The availability of first and second person plural verb forms in this context (*diavazoume, diavazete* in (4)) implies an apparent feature mismatch, or *unagreement*, between subject and verb.

It seems that languages allowing unagreement generally also have null subjects (Choi 2013, Höhn 2016). However, not all null subject languages allow unagreement

configurations.[2] This was illustrated for Italian in example (1) of section 1, and (5) provides a further example from European Portuguese:

(5) *Os portugueses bebemos bom café.   [European Portuguese]
     the Portuguese drink-1PL good coffee
   *intended*: 'We Portuguese drink good coffee.'   (Höhn 2016:547)

Simply put, it seems like definite noun phrases obligatorily behave as third person in languages like Italian and European Portuguese, but are underspecified for person in unagreement constructions in languages like SMG (4) or similar constructions in Spanish.

Both Choi (2013, 2014) and Höhn (2016) suggest that this correlates with crosslinguistic variation in the presence of a definite article in APCs, as illustrated in (6) based on Höhn (2016:559f). In null subject languages lacking unagreement, adnominal pronouns are in complementary distribution with definite articles, see (6a). In languages with unagreement, on the other hand, an adnominal pronoun matching the verbal agreement marking may optionally *precede* the mandatory definite article as illustrated in (6b).

(6) a. *No unagreement*   b. *Unagreement*
        noi (*gli) studenti   [Italian]   emeis *(oi) foitites   [Greek]
        nós (*os) estudantes   [Eur. Port]   nosotros *(los) estudiantes   [Spanish]
        mi (*a) diakók   [Hungarian]   nie studenti-*(te)   [Bulgarian]
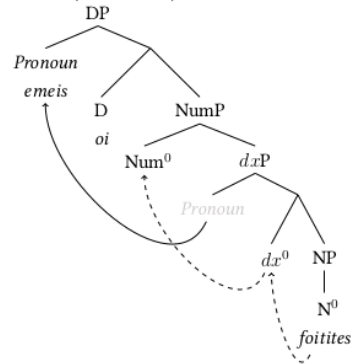   we (*DEF) students   we *(DEF) students/students-*(DEF)

Choi (2013, 2014) rejects the classical pronominal determiner analysis (Postal 1969, Abney 1987), illustrated in (7) for Italian *noi studenti* "we students", as it cannot accommodate the co-occurrence of definite articles and adnominal pronouns in the languages in (6b). To provide a unified analysis, he argues that rather than realising the head D, adnominal pronouns are phrases that move to SpecDP as sketched in (8a) for Greek with an overt definite article in D, and in (8b) for Italian with a covert one.

(7) Pronominal determiner structure (e.g. Abney 1987)

```
        DP
       /  \
      D    NP
     noi  studenti
```

---
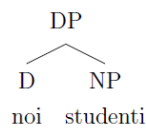
2 This marks a problem for Longobardi's (2008: 204) brief discussion of unagreement, where he suggests that languages with strong Person allow unagreement (for the strong vs. weak Person distinction see fn. 3). Having strong Person may be a necessary condition for unagreement, although, insofar as strong Person seems to correlate with the availability of null subjects, one may wonder which of them is the crucial factor. However, neither of them can be a sufficient condition, since the difference between SMG and standard Italian with respect to unagreement would remain unaccounted for, as both are strong Person languages with null subjects.

(8)   a. Choi (2014:141) for Greek   b. Corresponding structure for Italian
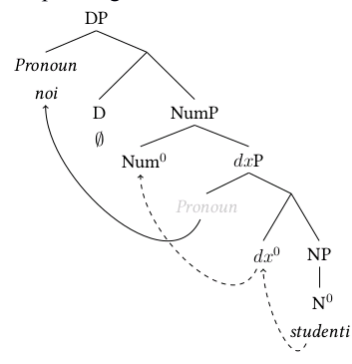

Choi treats unagreement as a form of pro-drop which requires licensing and identification by a fully specified T in the clausal domain, following Rizzi's (1982, 1986) classical analysis, but pro-drop has to be additionally licensed by the D head as per the Condition on D^0 in (9).

(9) **Condition on D^0**: D^0 must be overtly realized by a definite article (but, being a
     mediating pro-drop licenser, may not be as fully specified with its phi-features as T^0).

                                                                                    (Choi 2013, (20))

On this view, there is no structural difference between languages with and without unagreement. Instead, their different behaviour results from the phonological properties of D in interaction with the condition in (9).

Since we adopt the alternative account of Höhn (2016) below, we provide a slightly more detailed overview of that approach here. Höhn retains the pronominal determiner analysis (7) for languages without definite article in adnominal pronoun constructions, like standard Italian or European Portuguese, but proposes that there is crosslinguistic variation in the structural position of person to the effect that grammatical person occupies a position independent from the article in unagreement languages as illustrated in (10).

(10)
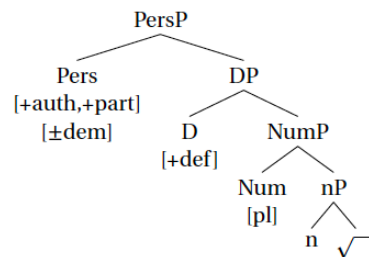
Adopting the general framework of distributed morphology (Halle & Marantz 1993, Embick 2010), Höhn proposes that the minimal difference between APCs and unagreeing DPs is the spell-out of the Pers head in (10).[3] Unagreement results if Pers receives null spell-out, the same process which is arguably responsible for pro-drop. The realisation of Pers is argued to depend on the specification of a [±dem(onstrative)] feature on Pers as illustrated by the Greek vocabulary items (VIs) in (11). The practice of decomposing person into two binary features [±part(icipant)] and [±auth(or)] follows Nevins (2007, 2011).

(11)   Pers [-dem] ↔ Ø   [Standard Modern Greek]
         Pers [+dem, +auth, +part, pl] ↔ *emeis*

In languages like standard Italian with the pronominal determiner structure in (7), simple pro-drop is essentially the result of the same process of null spell-out applying to the head encoding person. Due to the structural difference between (7) and (10), however, the relevant head in this type of language is D. Consequently, the mechanism responsible for pro-drop, namely null spell-out of the head carrying person features, cannot derive an output comparable to unagreement in this language. The application of this process to a D head with a lexical NP complement would predict the existence of definite bare nouns with first or second person interpretation in Italian, contrary to fact. Indeed, both Italian and Greek require overt definite articles even in generic DPs in subject position.[4] Höhn (2016) suggests that the only context in which a definite D can circumvent this overtness requirement is if DP does not contain any other overt material, e.g. an adjective or a noun, so effectively definite D can only be silent in pro-drop. For illustration, consider the sample list of VIs in (12):

(12)   D [+def, -dem] ↔ Ø / _]$_\varphi$ [Standard Italian]
         D [+def, +auth, +part, pl] ↔ *noi*
         D [+def, -auth, -part, pl, masc] ↔ *(gl)i*[5]

---

3 This framework distinguishes between the featural content of a syntactic node (e.g. person features) and its phonological realisation. The significance of this distinction for the present analysis lies in the way it allows a simple mechanism of accounting for syncretisms. Concerning the null-realisation of Person features in (11) and (12), it overcomes the need to postulate individual silent lexemes for each person-number combination.

4 Longobardi (1994, 2008) analyses this as a requirement for a connection between N and D in argumental DPs, terming languages that need an overt reflex of this connection *strong D* or *strong person languages* (as opposed to *weak person languages* like English where this connection can be established covertly, i.e. at LF). Possible ways of establishing this connection are N-to-D movement, notably in the case of proper names, or overt definite articles.

5 In standard Italian the definite plural article has a phonologically conditioned allomorph *gli* that is used before words beginning with vowels, glides, certain consonants (/ts/, /dz/, /ɲ//) and some consonantal clusters (/s/+C, /p/+/n/, /p/+/s/, /k/+/s/) (Marotta 1993). Since *i* is the elsewhere form, we will refer to this non-restricted vocabulary item throughout. The existence of the contextually restricted, competing form *gli* is silently assumed and changes nothing in the current discussion. The same reasoning applies for *quei*, the masculine plural of the demonstrative *quello*, *quegli* being its allomorph.

The first VI in (12) can realise a [-dem] D head as null if it is located at the right edge of the spell-out domain φ (in the sense of Embick 2010), that is only if none of the material in its complement domain is phonologically realised. The relevant configuration can be sketched as [$_{DP}$ D [$_{NumP}$ Ø]]. Notice that while the syntactic structures serving as input for spell-out are featurally fully specified, this VI with null realisation is underspecified for person, number and gender. This means that one VI is sufficient to regulate the null realisation of D, obliterating the need for a dedicated silent lexical item *pro* for each person number combination, as assumed by Choi (2013, 2014).

If the above structural and phonological conditions are not met, the null VI for D cannot be inserted. Consequently, another VI wins the competition for insertion and definite D is spelled out overtly according to its feature specification, e.g. by one of the other two VIs in (12). The fact that in the realisation of a pronominal determiner structure like (7) the definite plural article *i* competes for insertion with the pronouns accounts for the lack of unagreement. When D is specified as non-third person in typical APCs, the definite article loses out, irrespective of whether we assume that the vocabulary items realising it are underspecified for person or specified for third person ([-auth], [-part]) as in the third VI in (12). Remember that the input for vocabulary insertion consists of featurally fully specified syntactic structures, so that the D head in a typical APC would carry some non-third person feature specification. Consequently, the definite article would lose the competition against the VI of a non-third person pronoun either because the latter represents a better, more specific fit to the (person) features of the input (if the definite article lacks person features) or, if the VI(s) for the definite article are specified as [-auth, -part], because the third person definite article VI is incompatible with the non-third person syntactic input. In either case, this yields a plain adnominal pronoun construction with a pronominal determiner. A plain definite plural subject DP, the hallmark signature for unagreement, would not be derivable from a syntactic context that gives rise to non-third person verbal agreement on the assumption that this implies a structure with non-third person features on D.

3. Problems raised by northern and southern Calabrese

The data observed at the outset from northern and southern Calabrese, (2) and (3) respectively, clearly suggest that these languages allow unagreement in contrast with standard Italian. In spite of this difference, however, the SIDs under discussion do not allow the definite article in APCs, just like standard Italian, see (13).

(13)   a.   Noi (*i) bambini giochiamo.   [Standard Italian]
                we the children play.1PL

         b.   Nua (*i) quatrarə iucamə.   [Northern Calabrese]
                we the children play.1PL

         c.   Nui (*i) figghioli iocamu.   [Southern Calabrese]
                we the children play.1PL

This is problematic for the correlation between definiteness marking and unagreement proposed by Choi (2013, 2014) and Höhn (2016). In particular, the second part (14b) of

Höhn's generalisation in (14) as well as Choi's Condition on D^0 in (9) above are contradicted by the SID data:

(14)   Null subject languages with definite articles
        a. show unagreement if they have a definite article in APCs, and
        b. do not show unagreement if they have no definite article in APCs.

In the next section, we propose a modification to Höhn (2016) to account for the variation between standard Italian and the SIDs based on the hypothesis that the difference is morphological in nature.

4. A morphologically based analysis

In section 2 we pointed out that the specification of a [±dem] feature accounts for the difference between APCs and unagreement in Höhn's (2016) proposal. Following Sommerstein (1972) and Rauh (2003), we observe a similar distinction between regular and 'demonstrative' uses in languages with classical pronominal determiners.
A 'demonstrative' pronominal determiner is illustrated in the examples in (15) adapted from Sommerstein (1972).

(15)   a. YOU troops will embark but the other troops will remain.
        b. He said that (those/*the) troops would embark but the other troops would
           remain.

The pronominal determiner *you* in an utterance like (15a) cannot be reported by an observer using the definite article but only by a demonstrative. In Höhn's terms, this is suggestive of the pronominal determiner in (15a) carrying a [+dem] feature.

Against this background, we can observe a contrast between standard Italian and the SIDs under discussion with respect to how demonstrative and non-demonstrative pronominal determiners show up in reported speech. Similar to what we saw for English in (15), standard Italian adnominal pronouns can have a demonstrative reading in addition to their simple pronominal determiner use.

(16)   a. NOI bambini giochiamo.
           we children play.1PL
           'We children are playing.'

        b. Ha detto che quei bambini giocano.
           has.3SG said that those children play.3PL
           'S/he said that those children are playing.'

(17)   a. Noi bambini giochiamo.
           we children play.1PL
           'We children are playing.'

    b. Ha detto che i bambini giocano.
       has said that the children play.3PL
       'S/he said that the children play.'

The difference between (16) and (17) is very subtle and can be captured only if the context behind the examples is explained. Suppose that in (16) there are two groups of children, one is playing and the other is not. The group of children who plays produces the sentence in (16a) using a contrastive focus on the adnominal pronoun *noi* 'we' so as to differentiate themselves from the other group of children. As a consequence and for the same reason, the demonstrative *quei* 'those' must be used to report this utterance, as illustrated in (16b). A simple definite expression like in (17b) would not represent a proper report of (16a) in the sketched context. By contrast, in the second example, (17a), there is only one group of children (probably opposing to a group of adults). Since there is no need to differentiate between groups, in this case the demonstrative *quei* is not required.

Therefore, in Italian, on a par with the English examples, the contrastive adnominal pronoun in (16a) has to be reported using the demonstrative *quei* as in (16b), whereas a plain adnominal pronoun can be reported using the definite article *i* as shown in (17). Although there are intonational differences between the adnominal pronouns in (16a) and (17a), they are segmentally identical. Insofar it is fair to say that pronominal determiners in standard Italian can but need not be used as demonstratives.

Adnominal pronouns in the Calabrese SIDs pattern differently insofar as they have to be reported using demonstratives. Consider the examples below:

(18)   a. NUA quatrarə iucamə.   [Northern Calabrese]
        we children play.1PL
        'We children are playing.'

        b. Ha dittə ca quiddə quatrarə iocənə.
          has said that those children play.3PL
          'S/he said that those children play.'

(19)   a. I quatrarə iucamə.
        the children play.1PL
        'We children are playing.'

        b. Ha dittə ca i quatrarə iocənə.
          has said that the children play.3PL
          'S/he said that children kids play.'

(20)   a. NUI figghioli iocamu.   [Southern Calabrese]
        we children play.1PL
         'We children are playing.'

        b. Dissi chi ddi figghioli iocanu.
          said.3SG that those children play.3PL
          'S/he said that those children play.'

(21)   a. I figghioli iocamu.

the children play.1PL
'We children are playing.'

b. Dissi chi i figghioli iocanu.
said.3SG that the children play.3PL
'S/he said that the children play.'

The example pairs in (18) and (20) from northern and southern Calabrese respectively illustrate that in both varieties an observer reporting the (a) utterance needs to replace the adnominal pronoun by a demonstrative, whereas the definite article can be retained in cases of reported unagreement constructions as in (19b) and (21b).[6]

We suggest that this contrasting behaviour of adnominal pronouns in standard Italian and the SIDs is due to specific differences in the vocabulary of the different varieties, and that this also accounts for the exceptional availability of unagreement in the SIDs. More specifically, the crucial difference between Standard Italian and the southern varieties lies in the feature specification of personal pronoun. Also, while for standard Italian the question of whether VIs for the definite article are specified for third person is tangential to the availability of unagreement as discussed in section 2, in the southern varieties they are crucially underspecified for person. The crucial differences in the vocabulary between standard Italian and the SIDs are illustrated in (22) and (23) below.

(22)   D[+def, +auth, +part, pl] ↔ *noi*   [Standard Italian]
        D[+def, (-auth, -part,) pl, masc] ↔ *i*
(23)   D[+def, +dem, +auth, +part, pl] ↔ *nua/nui*   [Northern/southern Calabrese]
        D[+def, pl, masc] ↔ *i*

Crucially, the VIs corresponding to personal pronouns are underspecified for [±dem] in standard Italian, while they are specified for [+dem] in the southern varieties. Moreover, the VIs realising the definite article are crucially underspecified for person in the southern varieties, while they may or may not be specified for third person ([-auth,-part]) in standard Italian. This absence of person features in the definite article allows it to appear even if D is specified for non-third person and is therefore a precondition for the appearance of unagreement. However, it is possible that standard Italian definite articles are similarly underspecified. We propose that it is the difference in demonstrative specifications on personal pronouns that accounts for the variation in the availability of unagreement.

Consider the SIDs first. If D has a [+dem] feature, the pronoun with the appropriate person feature specification is inserted, yielding an adnominal pronoun construction. If D is specified for [-dem], the pronominal VIs are incompatible due to their [+dem] specification and the article is inserted instead as the elsewhere case. This yields a plain definite subject DP that is, however, syntactically specified for non-third person. Therefore, we get an unagreement configuration with non-third person verbal agreement.

In standard Italian, on the other hand, the VIs corresponding to personal pronouns are underspecified for [±dem]. This crucially accounts for the lack of unagreement in standard

6 Notice that a demonstrative would be possible in the latter case as well depending on the communicative intentions of the observer reporting the (a) utterance. The crucial point is that the overt adnominal pronouns in (18a) and (20a) have to be reported by a demonstrative.

Italian, since it puts them in competition with the definite article VIs for insertion into non-third person D nodes irrespectively of whether those are specified as [+dem] or [-dem]. The definite article VIs lose in all non-third person contexts even if it were underspecified for person because there is always a "pronominal" VI that is specified for the person features present in the syntactic structure. Hence, the definite article cannot be inserted in any non-third person contexts. In other words, unagreement simply cannot arise.

5. Conclusion

We have presented new data from two southern Italo-Romance varieties, northern and southern Calabrese, showing that in contrast to standard Italian they display the unagreement phenomenon, that is, definite plural subjects can agree with first, second and third person. This is theoretically relevant because it presents a problem for previous cross-linguistically based generalisations stating that in order for a language to allow unagreement its adnominal pronoun constructions like *we linguists* require an overt definite article (Choi 2013, 2014; Höhn 2016). Although both varieties differ from standard Italian with respect to unagreement, their adnominal pronoun constructions are like those of the standard language in not allowing a definite article.

We propose an extension of Höhn's (2016) analysis to account for the exceptional behaviour of the southern varieties. Building on observations by Sommerstein (1972) and Rauh (2003) for English pronominal determiners, we argue that while the SIDs and standard Italian share the same nominal structure, encoding person and definiteness on the same head (Postal 1969; Abney 1987), the crucial difference is a morphological one.

Standard Italian pronominal determiners can behave like demonstratives, while in the SIDs they *have to* behave like demonstratives. We propose that this is due to a difference in the specification of Höhn's (2016) [±dem] feature on the vocabulary items realised as personal pronouns. While the pronominal VIs are specified for [+dem] in the Calabrese varieties, their standard Italian counterparts are underspecified for that feature. As a consequence, the pronominal VIs do not compete with the VIs corresponding to the definite article for vocabulary insertion in a D node that is specified as [-dem]. Since the definite article VIs are underspecified for person, they can be inserted in non-third person D nodes, leading to the unagreement configuration. On the other hand, because pronominal VIs are underspecified for [±dem] in standard Italian they compete with the definite article VIs in [-dem] contexts as well and win over them in all non-third person contexts.

If this account is on the right track, it underlines the importance of morphology, in particular of the vocabulary, in the derivation of unagreement structures. An alternative analysis is conceivable that would directly extend Höhn's (2016) analysis of the variation in unagreement as a result of a structural difference. On such a view, the SIDs would share the structure proposed for Spanish and Greek in (10) above, with definiteness and person encoded on separate heads.

The advantage of such an approach would be a unified structure underlying the unagreement phenomenon. However, it would raise questions as to why D would receive zero spell-out in adnominal pronoun constructions in northern and southern Calabrese, in contrast to other unagreement languages. We leave further elaboration of such an approach and its comparison to the one presented here to future research.

References

Abney, Steven. 1987. *The English noun phrase in its sentential aspect*. PhD dissertation: MIT.

Choi, Jaehoon. 2013. "Pro-drop in pronoun-noun constructions". In *North East Linguistic Society* (NELS) 42, ed. by Stefan Keine and Shayne Sloggett, 119–128. Amherst: GLSA

Choi, Jaehoon. 2014. *Pronoun-noun constructions and the Syntax of DP*. PhD dissertation, University of Arizona.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge (MA): MIT Press.

Halle, Morris, and Alec Marantz. 1993. "Distributed morphology and the pieces of Inflection". In *The view from building* 20, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge: MIT Press.

Höhn, Georg F.K. 2016. "Unagreement is an illusion. Apparent person mismatches and nominal structure". *Natural Language and Linguistic Theory* 34 (2): 543-592. Available at: <http://ling.auf.net/lingbuzz/001790>

Hurtado, Alfredo. 1985. "The unagreement hypothesis". In *Selected papers from the Thirteenth Linguistic Symposium on Romance Languages*, ed. by Larry King and Catherine Maley, 187-211. Amsterdam: John Benjamins.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Longobardi, Giuseppe, 1994. "Reference and proper names: a theory of N-movement in syntax and logical form". *Linguistic Inquiry* 25, 4:609-665.

Longobardi, Giuseppe. 2008. "Reference to individuals, person, and the variety of mapping Parameters". In *Essays on Nominal Determination: from morphology to discourse management* ed. by A. Klinge & H. Müller, 189-211. Amsterdam: John Benjamins.

Marotta, Giovanna. 1993. "Selezione dell'articolo e sillaba in italiano: un'interazione totale?". *Studi di Grammatica Italiana* 15: 255-93.

Nevins, Andrew. 2007. "The representation of third person and its consequences for person-case effects". *Natural Language and Linguistic Theory* 25 (2): 273–313.

Nevins, Andrew. 2011. "Multiple agree with clitics: Person complementarity vs. omnivorous number". *Natural Language and Linguistic Theory* 29: 939–971.

Postal, Paul. 1969. "On so-called "pronouns" in English". In *Modern studies in English: Readings in Transformational Grammar*, ed. by David A. Reibel and Sanford A. Schane, 201–226. Englewood Cliffs: Prentice Hall.

Rauh, Gisa. 2003. "Warum wir Linguisten "euch Linguisten", aber nicht "sie Linguisten" akzeptieren können. Eine personendeiktische Erklärung". *Linguistische Berichte* 196: 390-424.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Forris.

Rizzi, Luigi. 1986. "Null Objects in Italian and the Theory of *pro*". *Linguistic Inquiry* 17.3 501-557.

Sommerstein, Alan H. 1972. "On the so-called definite article in English". *Linguistic Inquiry* 3: 197-209.

REDUPLICATION IN ROMAN:
A CASE OF MICROVARIATION

Dara Jokilehto
Université de Genève

Abstract: *This paper discusses a case of reduplication found in Roman Italian, a case of remnant movement. The data is compared to similar phenomena in River Plate Spanish (Saab 2011) and Italian (Gullì 2003). It is argued that Saab (2011) cannot account for the Roman case of reduplication, which involves the reduplication of phrases, since his approach can only prevent the deletion of heads. Gullì's (2003) account fails to predict that Focus-Fronted elements appear above the reduplicated phrase in Roman. Furthermore, it is argued that the projection between the reduplicated phrases must be that of Mirative Focus (Cruschina 2012), due to the different interpretation of Roman with respect to Gullì's examples. For these reasons, it is proposed that the projection triggering remnant movement, RedupP, is actually situated below FocP, and that what Gullì uses to host the material sandwiched between the reduplicated strings is a Mirative Focus projection, MFocP.*

Keywords: *syntactic reduplication, cartography, information structure*

1. Introduction

The phenomenon to be examined in this paper is that of VP reduplication in the Roman variety of Italian. Previous accounts of similar phenomena, Saab (2011) for River Plate Spanish (RPS), and Gullì (2003) for Italian and Calabrese, are shown to differ from Roman in discrete ways. This paper argues for a Cartographic analysis (cf. Rizzi 1997) of the Left Periphery as shown below, in (1):

(1)   ForceP
```
        ForceP
        /\
          TopP
          /\
            FocP
            /\
              RedupP
              /\
                MFocP
                /\
                  RP
                  /\
                    FinP
                    /\
                      TP
                      /\
                      ...
```
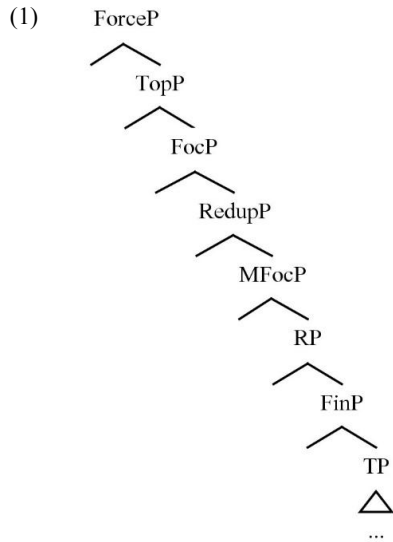
Roman Reduplication (RR) involves the extraction of an element, followed by remnant movement, creating the conditions for preventing the deletion of a copy, but the RR data cannot be accounted for by the above approaches: it features phrasal movement, incompatible with Saab (2011), since phrasal movement of the verb (and reduplicated string) hampers the mechanism preventing deletion on Saab's approach. Also, RR expresses Mirative Focus (MFoc; cf. Cruschina 2012), unlike Gullì's data for Italian, which expresses Contrastive Focus (ConFoc). Given the position of (Corrective) Focus-Fronted elements, it is proposed that the landing site for the remnant movement is below FocP, contra Gullì (2003). Following Cruschina's (2012) account of Mirative Focus, it is further proposed that, given the different interpretation of RR, the Focus projection to host the material between the reduplicated strings is MFocP, due to the particular interpretation of RR utterances.

This paper is divided as follows: section 2 will describe the Roman data.

Section 3 will compare Saab's (2011) account of RPS to RR. Section 4 will outline Gullì's (2003) account of a similar reduplication in Standard Italian and argue for some changes to be made in order to fit the Roman data. Section 5 concludes the paper.

2. Roman Reduplication

This section will provide a description of Roman Reduplication (RR), arguing that RR is a root phenomenon expressing Mirative Focus (Cruschina 2012). (2) shows how RR consists in the reduplication of (at least) part of the VP in what might be called the right periphery:

Roman (Vernacular)[1]

(2)　　M'　　　　hanno　　fregato 'r motorino　m'　　　hanno fregato!

　　　　CL.1SG.DAT　have.3PL　steal.PP the scooter CL.1SG.DAT　have.3PL stolen

　　　　'They stole my scooter!'

RR is used to express surprise, excitement or agitation: the interpretation is therefore that encountered with Mirative Focus (Cruschina 2012), which is not usually realised with this structure in Italian varieties elsewhere in Italy, but with Focus Fronting (see also Bianchi and Bocci 2012). The focus of the clause is found sandwiched between the verbs, auxiliaries and relevant clitics: this is often the complement of the verb (2), though it may also be the subject (3a), an adverbial (3b), or a participle (3c). Adjacency of the two reduplicated strings is bad (3).

(3)　　a.　　Te　　　　　scerca　　Gianni, te　　　　scerca!

　　　　　　CL.2SG.ACC　seek.3SG　John CL.2SG.ACC　seek.3SG

　　　　　　'John is looking for you.'

　　　　b.　　Hai　　　fatto bene, hai　　fatto!

　　　　　　have.2SG　done well　have.2SG　done

　　　　　　'You did well.'

　　　　c.　　M'　　　　hanno　　fregato, m'　　　hanno!

　　　　　　CL.1SG.DAT　have.3PL　steal.PP　CL.1SG.DAT have.3PL

　　　　　　'They've cheated me!'

　　　　d.　　Tô[2]　　　　　　　　　dico　　*(domani)

　　　　　　CL.2SG.DAT-CL.3SG.ACC　say.1SG　tomorrow

　　　　　　tô　　　　　　　　　　dico.

　　　　　　CL.2SG.DAT-CL.3SG.ACC　say.1SG

　　　　　　'I'll tell you tomorrow.'

There are restrictions to what may be reduplicated, however. The reduplicated string must start with the highest verb (or auxiliary) (4a), and be contiguous (4b). Any clitics present must also be reduplicated, including optional clitics (4c). There is no restriction on the type of verb (argument structure, control, restructuring, etc.).

(4)　　a.　　S' è　　　　voluto　　compra'　　tutto　　'r negozio…

　　　　　　si be.3SG want.PP　buy.INF　　all　　　the shop

　　　　　　'He wanted to buy the whole shop for himself!'

---

[1] Abbreviations key: CL = clitic; SG = singular; PL = plural; F = feminine; ACC = accusative; DAT = dative; PP = past participle; IMPERF = imperfect; INF = infinitive; LOC = locative. Indicative mood, present tense and masculine gender are unmarked.

[2] The circumflex accent marks a long vowel.

    b.   *se ~~è~~   voluto compra'!

    c.   *s' è   ~~voluto~~ compra'!

    d.   *~~se~~   è/ha   voluto compra'![3]

            si be.3SG/have.3SG want.PP buy.INF

Strings inside embedded clauses may not be reduplicated (5c, d, g, h), which indicates that RR is a root phenomenon (regardless of whether the complementiser is reduplicated). But material inside embedded clauses can be reduplicated, as long as the string begins in the matrix clause and the embedded clause is untensed (5b, f). (6) shows that *far fare* causative constructions pattern like untensed embedded clauses with respect to the availability of RR:

        Tensed embedded clauses

(5) a.   Vedo   [ch' è   arivato Gianni]…

          see.1SG that be.3SG arrive.PP John

          'I see that John's arrived!'

    b.   *vedo [ch'è arivato]

    c.   *~~vedo~~ [ch'è arivato]

    d.   *~~vedo [ch'~~è arivato]

        Untensed embedded clause

    e.   Ciò   paura [d' ariva'   troppo tardi]…

          CL.LOC-have.1SG fear   of arrive.INF   too   late…

          'I fear I'll arrive too late!'

    f.   Ciò paura d'ariva'!

    g.   *~~Ciò paura~~ d'ariva'!

    h.   *~~Ciò paura d'~~ ariva'!

        Causative construction

(6)   Tâ   faccio   [fa'   tutta a   piedi,

          CL.2SG.DAT-CL.3SG.ACC.F make.1SG make.INF all.F at   foot

          tâ   faccio   [fa']]

          CL.2SG.DAT-CL.3SG.ACC.F make.1SG make.INF

          'I'll make you walk the whole way!'

Echo/reprise questions are compatible with RR (7c), but not normal Wh-questions with Wh-fronting (7d). In Yes/No questions, RR is allowed, with the effect of increasing their "forcefulness" (7e).

---

[3] The auxiliary *ave'*, 'to have' is included here, since the auxiliary *esse* 'to be' is only used here because of the reflexive *si*.

(7)   A: a.   John said [he's leaving].
        B: b.   John said *what*?!

            c.   Gianni   ha   detto   (**che**) **cosa**   ha   detto?
                John   have.3SG say.PP what thing   have.3SG say.PP
                'John said what!?'

            d.   (Che)   cosa   ha   detto   Gianni,   (*ha   detto)?
                  what thing   have.3SG say.PP   John   have.3SG say.PP

            e.   Vòi   'na machina tua,   vòi?
                want.2SG a   car   your.F want.2SG
                'You want a car of your own?'

       The Left Periphery is available for fronting operations, allowing both corrective foci (8a) and contrastive topics (8c). (8a) also shows that when the contrastive element is fronted, some material (the past participle here) must remain to prevent the reduplicated strings from being adjacent: unpronounced copies or traces cannot fulfil this role. (8c) shows that in-situ Focus is awkward (8c), especially with negative tags (*no 'r portafojjo*, 'not [my] wallet'). Instead, focus fronting is preferred, interpreted as a Corrective Focus. Based on Bianchi and Bocci (2012), I take this as evidence that Focus in-situ is not allowed with RR, but Focus-Fronting is and is interpreted as Corrective Focus. Clitic Right Dislocation is also possible (8c):

            Focus fronting (Corrective)
(8)   Context: 'I heard they stole your scooter.'
      a.   **'A machina**   m'   hanno   fregato,   m'
          the car   CL.1SG.DAT   have.3PL   steal.PP   CL.1SG.DAT
          hanno (*fregato)!
          have.3PL steal.PP
          'The car, they stole from me!'

          Context: 'I heard they stole your wallet.'
      b.   *Focus in-situ (Contrastive)*
          #M'   hanno   fregato tutto   m'   hanno
          CL.1SG.DAT   have.3PL   steal.PP   everything   CL.1SG.DAT have.3PL
          fregato,   (??no/mica   'r portafojjo).
          steal.PP   not   the wallet
          'They stole everything, not my wallet!'

Contrastive Topic

c. Quer giubbotto lì   me   l'   aveva
    that coat there   CL.1SG.DAT   CL.3SG.ACC   have.IMPERF.3SG
    comprato   mi nonno,   me   l'
    bought   my grandfather   CL.1SG.DAT   CL.3SG.ACC
    aveva   comprato!
    have.IMPERF.3SG   bought
    'That coat there, my grandfather bought me.'

Clitic Right Dislocation

d. Me   l'   hanno   fregato in due, me
    CL.1SG.DAT   CL.3SG.ACC   have.3PL steal.PP in two CL.1SG.DAT
    l'   hanno   fregato, er   motorino!
    CL.3SG.ACC.m have.3PL steal.PP   the   scooter
    'It was two people who stole my scooter!'

3. *Roman vs. River Plate Spanish*

This section will outline Saab's (2011) account of *capicúa*, a case of reduplication found in River Plate Spanish (RPS). It will be argued that Roman Reduplication (RR) cannot be accounted for by Saab (2011), since said account only allows the reduplication of heads, which is insufficient to account for the phrasal movement encountered in RR.

Essentially, all the approaches mentioned in this paper involve remnant movement, differing mainly in two respects: how deletion is averted and the first movement out of the remnant, constituting what is, in fact, reduplicated. This is the approach taken by Saab (2011) for River Plate Spanish (RPS), following Martins (2007) on European Portuguese. A similar approach was also taken by Gullì (2003).

Saab's (2011) approach to reduplication in RPS, known in the region as *speaking capicúa*, also extracts the string to be reduplicated before using remnant movement, but, unlike RR, in RPS only the finite verb can be reduplicated (9)[4]. Essentially, the verb moves through T, Pol and Foc. After the verb has moved out of PolP, PolP raises, in turn, to Spec,FocP. Saab makes no mention of compound tenses, but it is assumed only the auxiliary would be reduplicated (if this is at all possible in RPS):

River Plate Spanish

(9)   a. [$_{FocP}$ [$_{PolP}$ vino Juan] [$_{Foc'}$ [$_{Foc}$ vino+Foc^o] [$_{PolP}$ ~~vino Juan~~]]]
                 came John   came
     'John came!'

<div align="right">Saab (2011:317)</div>

---

[4] Also the case in Martins (2007) for European Portuguese.

b.

```
        FocP
         |
        Foc'
       /    \
     Foc    PolP
      |      /  \
    vino   Pol'
          /    \
        Pol    TP
         |     / \
       vino       T'
                 /  \
                T   VP
              [+I]  / \
                   V'
                  / \
                 V
               [+I]
```

c.

```
            FocP
           /    \
        PolP    Foc'
       /  \     /   \
   vino Juan  Foc   PolP [+I]
              |      / \
            vino   vino Juan
```
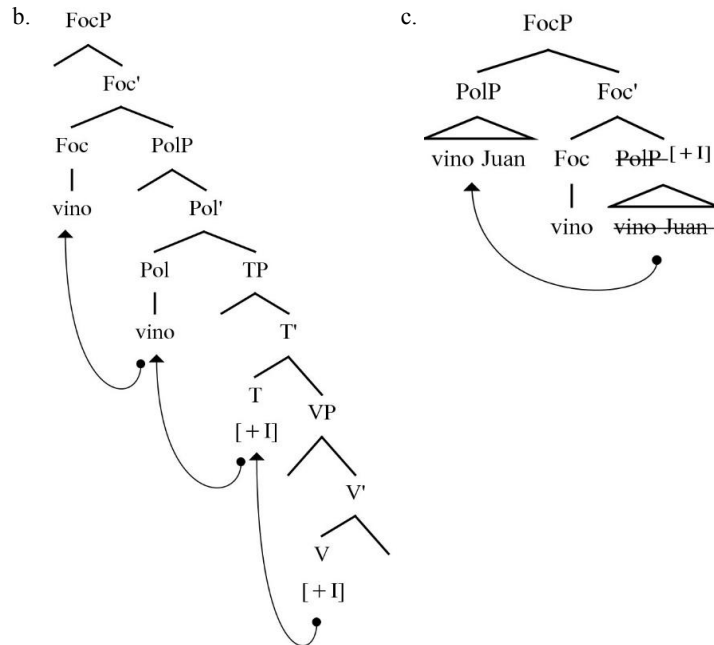
In Saab's system, deletion[5] targets elements which have been assigned a [+I], or *identity*, feature, following the rule in (10). The formulation of the rule prevents the deletion of *subwords* (11), when these are marked for deletion but have been incorporated into a morphosyntactic word not marked for deletion: this is relevant for heads though, not for phrases, and is therefore not useful for the RR data. This [+I] assignment is not identical for phrases and heads: while phrases are marked for deletion in the syntax, heads are only assigned [+I] after linearisation. The lower copy of PolP, being a phrase, is marked for deletion under c-command (9c). The mechanism for head deletion (and the prevention thereof), which was crucial for RPS, we can leave aside, since it will not aid us with the RR examples, where what moves is not a head. Because of this and because of the c-command relation, which is retained in RR, it is not possible to prevent its deletion on Saab's approach.

---

[5] Or rather, Non-Insertion, as Saab follows Halle and Marantz (1993) in adopting Late (Lexical) Insertion, where Spell-Out is a post-syntactic operation.

(10) **Non-Insertion:** No Lexical Insertion Rule, IR, applies in the domain of X^0, X^0 a MWd, if X^0, or some projection of X^0, is specified with a [+I] feature.
Associated definitions:
The domain of X^0, X^0 a MWd, is the set of terminal nodes reflexively contained in X^0.
Morphosyntactic word: At the input to Morphology, a node X^0 is (by definition) a *morphosyntactic word* (MWd) if X^0 is the highest segment of X^0 not contained in another X^0.
Subword: A node X^0 is a *subword* (SWd) if X^0 is a terminal node and not a MWd.
((ii) and (iii) from Embick & Noyer 2001:574).

<div align="right">Saab (2011:314-5)</div>

(11) **Sub-Word Deletion Corollary:** No SWd can be subject to non-insertion if the MWd that contains it is not *I*-assigned.

<div align="right">Saab (2011:315)</div>

The approach in Saab (2011) relies on head movement to save a copy of the verbal head from ellipsis. The reduplication of non-heads in RR makes such a solution ineffective for Roman, since head movement would fail to extract the required string. If we wish to retain the spirit of Saab's (2011) approach, it may be necessary to create more functional projections to provide landing sites for the phrase to be moved. The [+I] assignment rule would also fail to allow lexical insertion, since the morphological ellipsis mechanism only applies to heads.

4. Microvariation in the Left Periphery

In this section, Gullì's (2003) account of the reduplication of emphatic declaratives is outlined and then compared to the Roman data. It is shown that Roman remnant movement targets a lower position than that proposed by Gullì. Furthermore, based on semantic-pragmatic differences, it is argued that the projection hosting Gullì's [A*A] operator depends on the interpretation of the construction.

Gullì (2003) discusses a number of different types of reduplication in Calabrese and Italian. Due to space constraints, only the most relevant form of reduplication will be examined: plain reduplication of emphatic declaratives (12). Gullì's examples for this type of reduplication are from Italian.

(12) a. È   andato   a   Parigi   è   andato
        be.3SG   go.PP   to Paris   be.3SG   go.PP
        'He really did go to Paris.'

    b. È   andato   a   Parigi   è   andato   (non a Londra).
        be.3SG   go.PP   to Paris   be.3SG   go.PP   (not to London).
        'He really did go to Paris. (Not to London.)'

<div align="right">based on Gullì 2003:15</div>

The sentence in (12a) is the emphatic declarative type of plain reduplication. Gullì argues that it expresses Contrastive Focus, since a negative tag is possible to contrast the sandwich element with something else (12b), in this case *a Parigi* 'to Paris' with *a Londra* 'to London'. Based on the English gloss, it may be that Gullì's emphatic declaratives also express a form of Verum Focus, in the sense of Höhle (1992), although this is not mentioned in the paper.

Gullì's account for this involves extraction and remnant movement. What is responsible for the non-deletion of elements is a feature [+r], which causes a lower XP to move, while indicating that the copy of a moved element must not be deleted. This feature is closely linked to a conjunction/disjunction operator [A*A], based on the [A‾A] operator proposed by Huang (1982), which is situated in a projection RP (Relator Phrase). RP establishes a relation between the element in Spec,RP and the element in its complement's specifier position: for emphatic declaratives, this is DeclP. Gullì claims the [+r] feature is in a higher C projection RedupP and also in the lower RP, though the latter can bestow this feature to any Phrase in its specifier position.

(13) a. RedupP ... b. RedupP ...

Gullì (2003:35)
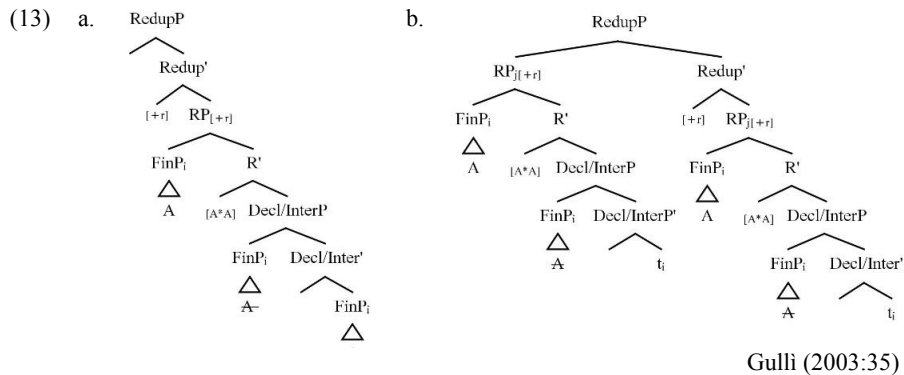
In (13a), the [+r] feature prevents the deletion of the FinP in Spec,RP, while the lower copy is deleted. This is not enough, however; the c-command relation must also be disrupted, adopting Kayne's (1994) Chain-Link Deletion Rule, given in (14). Therefore, RP must move to Spec,RedupP, whereupon FinP no longer c-commands the lower copy, which is no longer affected by the Chain-Link Deletion Rule. Thus, Gullì makes use of two principles in order to prevent reduplication from being over-generated: remnant movement is too common to account for reduplication on its own and without a [+r] feature, Gullì argues, reduplication could never occur. Note that, unlike Saab (2011), in this case, the element not being reduplicated is extracted before remnant movement, while in RPS, it remained in the remnant.

(14) **Chain-Link Deletion:** A given chain c_k can license PF deletion of another link c_i of the same chain only if c_i does not c-command c_k.
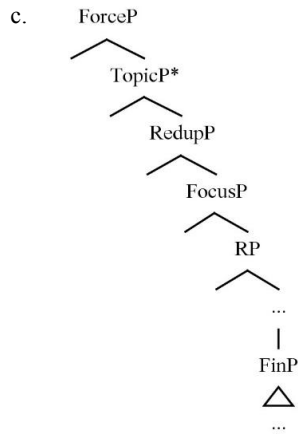
Kayne (1994:96)

Gullì's (2003) reduplication of emphatic declaratives is very similar to RR, but there are some crucial differences. Gullì claims the reduplication structure for emphatic declaratives is underlyingly the same as that of another structure: predicate clefts, an

example of which is (15a). Focus-fronted DPs, however, appear after the cleft (this is also the case with fronted Wh-elements in questions). Hence, Gullì assumes the structure in (15c), where RedupP is merged between the higher TopP and the lower FocP:

(15) a. Mangiare   ho   mangiato.
          eat.INF   have.1SG   eat.PP
          'As for eating, I have eaten.'

      b. Mangiare (PASTA) ho   mangiato.
          eat.INF   pasta   have.1SG   eat.PP
          'As for eating, I have eaten pasta.'

                                                        Gullì (2003:92)

      c. ForceP
                TopicP*
                     RedupP
                          FocusP
                               RP
                                   ...
                                   |
                                   FinP
                                   △
                                   ...

                                                        Gullì (2003:91)
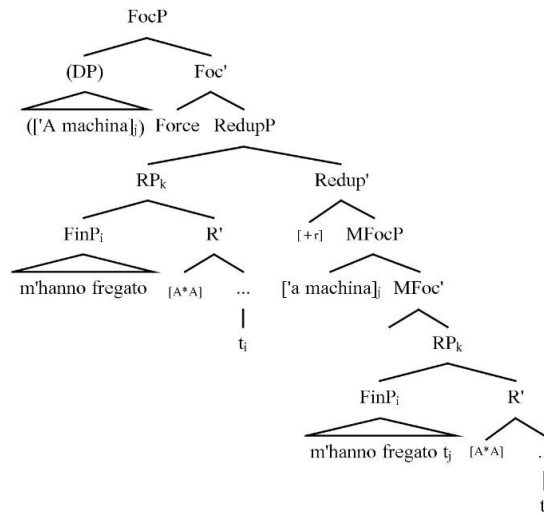
        As shown in (8a), repeated below, RR conflicts with Gullì's proposed structure: Focus-Fronted elements appear above the reduplicated phrase, hence above RedupP, and cannot appear anywhere below it (8aʹ). The landing site for the remnant movement must therefore be Spec position lower than the landing site for Focus-Fronting (assuming the landing site for Focus Fronted elements does not vary from Roman to Standard Italian).
        The proposed structure is shown in (16), with the position of the Focus-Fronted element in parentheses.

(8)   a. **'A machina**   m'   hanno   fregato,   m'
          the car   CL.1SG.DAT have.3PL   steal.PP   CL.1SG.DAT
          hanno!
          have.3PL
          'The car, they stole from me.'

      aʹ. M'hanno (*'a machina) fregato (*'a machina) m'hanno!

(16)

```
                          FocP
                    ┌───────┴───────┐
                  (DP)             Foc'
                   ┌┴┐          ┌────┴────┐
          (['A machina]ⱼ)    Force      RedupP
                                    ┌──────┴──────┐
                                  RPₖ           Redup'
                            ┌──────┴──────┐   ┌────┴────┐
                         FinPᵢ          R'  [+r]     MFocP
                        ┌──┴──┐      ┌───┴──┐     ┌────┴────┐
               m'hanno fregato [A*A] ...  ['a machina]ⱼ   MFoc'
                                     │               ┌─────┴─────┐
                                     tᵢ            RPₖ
                                              ┌──────┴──────┐
                                           FinPᵢ           R'
                                          ┌──┴──┐       ┌──┴──┐
                                 m'hanno fregato tⱼ [A*A]   ...
                                                             │
                                                             tᵢ
```

Keep in mind that the interpretation of RR utterances is not quite the same as that of Gullì's examples. RR expresses Mirative Focus, which communicates surprise, excitement or agitation. Furthermore, these structures are infelicitous as answers to questions, do not allow negative tags and are not contrastive. They are ungrammatical in unmarked questions, but allow (Corrective) Focus-Fronting. Hence, it is proposed that, at least in Roman, RedupP is merged below FocP (and probably the lower Top projections used for clitic left dislocation). The projection between RedupP and RP is not FocP, but MFocP. If predicate clefts do indeed involve RedupP (rather than being a Contrastive Topic), this must either be a separate, higher projection than that used for RR, or the position of RedupP must vary cross-linguistically.

5. Conclusion

This paper has argued that the reduplication found in Roman differs from those found in River Plate Spanish (Saab 2011), and Standard Italian (Gullì 2003). Saab (2011) makes no provision to prevent the reduplication of non-heads, inadequate for Roman Reduplication (RR), given that the reduplicated string consists of a phrase. Gullì's (2003) unified approach to reduplication in Calabrese and Italian predicts that Focus-Fronted elements should appear lower than the reduplicated string, which is not the case for RR: consequently, the landing site for remnant movement (Spec,RedupP) must be situated below FocP. Given the different interpretation of RR with respect to Gullì's Italian data, it is suggested that the operator [A*A] linked to RedupP is in a projection below MFocP (cf. Cruschina 2012).

References

Bianchi, Valentina, and Giuliano Bocci. 2012. "Should I stay or should I go? Optional focus movement in Italian". In *Empirical issues in syntax and semantics, vol. 9*, ed. by Christopher Piñón, 1-18. Available at: <http://www.cssp.cnrs.fr/eiss9/eiss9.pdf>

Cinque, Guglielmo. 2002. *Functional structure in DP and IP. The Cartography of Syntactic Structures. Volume 1*. New York: Oxford University Press.

Cruschina, Silvio. 2012. *Discourse-Related Features and Functional Projections*. New York/Oxford: Oxford University Press.

Embick, David, and Ralf Noyer. 2001. "Movement operations after syntax". *Linguistic Inquiry* 32:555-595.

Gullì, Antonino. 2003. "Phrasal Reduplication Syntax". PhD dissertation, the City University, New York.

Halle, Morris, and Alec Marantz. 1993. "Distributed Morphology and the Pieces of Inflection". In *The View from Building 20*, ed. by Ken Hale and Samuel Jay Keyser, 111-176. Cambridge: MIT Press.

Höhle, Tilman. 1992. "Über Verum-Fokus im Deutschen". In *Informationsstruktur und Grammatik*, ed. by Joachim Jacobs, 112-141. Opladen: Westdeutscher Verlag.

Huang, Cheng-Teh James. 1982. "Logical Relations in Chinese and the Theory of Grammar". PhD Dissertation, MIT.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, M.A.: MIT Press.

Martins, Ana Maria. 2007. "Double realization of verbal copies in European Portuguese emphatic affirmation". In *The Copy Theory of Movement*, ed. by Norbert Corver and Jairo Nunes, 77-118. Amsterdam: John Benjamins Publishing Co.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". In *Elements of Grammar*, ed. by Liliane Haegeman. Dordrecht: Kluwer.

Rizzi, Luigi. 2004. *The Structure of IP and CP. The Cartography of Syntactic Structures. Volume 2*. New York: Oxford University Press.

Saab, Andrés. 2011. "On verbal duplication in River Plate Spanish: Anti-adjacency and head copy deletion". In *Romance Languages and Linguistic Theory 2009: Selected papers from 'Going Romance' Nice 2009* ed. by Janine Berns, Haike Jacobs and Tobias Scheer, 305-322. Amsterdam: John Benjamins Publishing Co.

CLITIC DOUBLING AS DIFFERENTIAL
OBJECT MARKING[*]

Dalina Kallulli
Universität Wien

Abstract: *Taking my cue from clitic doubling constructions primarily in Balkan languages where this phenomenon is prevalent, my goals in this paper are to show that: (i) clitic doubling is a form of differential object marking; (ii) clitic doubling is the spell-out of agreement with a [+given] XP; (iii) clitic doubled indefinites are "non-novel" indefinites; (iv) the restrictor of 'all'-quantifiers is always 'given'; (v) generalized quantifiers are formed in two steps, crucially involving an intermediate DP-layer as the complement of the quantifier head; (vi) the XP in (ii) is invariably a DP; (vii) the person-case constraint is in effect differential object marking.*

Keywords: *clitic doubling, differential object marking, person case constraint*

1. Introduction

Taking my cue from clitic doubling constructions primarily in Balkan languages where this phenomenon is prevalent (see Kallulli and Tasmowski 2008), I put forward the following claims: (i) clitic doubling is a form of Differential Object Marking (DOM) (Bossong 1991); (ii) clitic doubling is the spell-out of agreement with a [+given] XP; (iii) clitic doubled indefinites are "non-novel" indefinites (Krifka 2001); (iv) the restrictor of 'all'-quantifiers is always 'given'; (v) generalized quantifiers are formed in two steps, crucially involving an intermediate DP-layer as the complement of the quantifier head (Matthewson 2001); (vi) the XP in (ii) is invariably a DP; (vii) the Person-Case Constraint (PCC) is in effect DOM.

In particular, inspired by Kiparsky (2008), I focus on the relation between generalizations drawn in previous work and a true universal grammar principle, namely the "D-hierarchy". While clitic doubling is always the spell-out of agreement with a topic XP, a concern of the present paper is to show how this phenomenon can be brought in line with Givón's (1975) idea that (verbal) agreement is always topic agreement, and with the vast typological literature on (other well-known cases of) differential object marking.

The rest of this paper is organized as follows. Section 2 provides some preliminaries and sets the stage for what is to follow. Then, in section 3 I revisit the Albanian and Greek clitic doubling patterns, which on top of violating Kayne's Generalization (Jaeggli 1982), serve as a rather good illustration of syntactic micro-variation in this domain. Section 4 juxtaposes these to the patterns of the definite objective conjugation in Hungarian, which

---

for all intents and purposes has the same function as clitic doubling, and which I contend reflects the workings of the same underlying universal, namely the "D-hierarchy" (Kiparsky 2008). Section 5 wraps the discussion up.

2. Setting the stage

Many clitic doubling languages (e.g. Albanian, Aromanian, Greek, Megleno-Romanian and varieties of Spanish) violate the so-called "Kayne's Generalization" (Jaeggli 1982), which informally stated, says that clitic doubling is possible whenever the (doubled) noun phrase can get case by means of some non-verbal device that has case assigning properties, namely prepositions. Simplifying somewhat, the idea was that the doubling clitic absorbs Case, so unless a preposition (or some other case-assigning device) could be inserted, the DP-argument would remain caseless, and the Case Filter would cause the derivation to crash.[1] In contrast, Suñer (1988) argues that *a* in Spanish is an animacy marker, which is why in spite of clitic doubling it is missing in the examples in (1) that she provides.

(1) a. Yo **lo** voy a comprar el diario justo antes de subir. (Porteño Spanish)
         'I am going to buy it-the newspaper just before coming up.'
     b. Yo **la** tenía prevista esta muerte.
         'I had foreseen it-this death'
     c. Ahora tiene que seguir usándo**lo** el apellido.
         'Now she has to go on using it-the surname.'

Across languages, clitic doubling affects interpretation in ways subject to various idiosyncratic constraints that make it hard to define its function in a unitary manner. Early generative studies described clitic doubling as sensitive to the feature *humanness* in Rumanian and *animacy* in Spanish, a view that was already untenable for particular varieties of Romance (see (1)). With the Balkan patterns coming into the focus of research on the topic, other semantic properties such as *prominence*, *specificity*, *presuppositionality*, *familiarity*, *definiteness* and *topicality* have increasingly been scrutinized as to their relevance for the phenomenon of clitic doubling (see Kallulli and Tasmowski 2008). Be it as it may, the mention of (each of) these notions enables one to see how a rather direct connection of this phenomenon to what Kiparsky (2008) refers to as the "D-hierarchy", given in (2), can be established. Importantly, drawing on Wierzbicka (1981), Kiparsky remarks that the hierarchy involves neither "animacy" nor "agentivity", which makes a

---

[1] In an effort to account for the violation of Kayne's Generalization with respect to indirect object clitic doubling in (Standard) Romanian, Dobrovie-Sorin (1994) appeals to the fact that indirect objects in this language bear morphological dative case, which in pre-minimalist Case theory was considered to be inherent, and as such, different from structural case that had to be assigned by a governing category. However, further research soon revealed that this phenomenon is quite extent in the Balkan languages: Albanian and Greek exhibit clitic doubling not only of inherently case marked indirect objects, but also of structurally governed direct objects bearing morphological accusative case but that nonetheless do not co-occur with a prepositional element. Moreover, Macedonian, Bulgarian, (and among the Romance languages) Aromanian and Megleno-Romanian can double a purely structurally governed direct object that bears no morphological case (see Kallulli and Tasmowski 2008).

direct functional explanation implausible, and that a category related to definiteness, such as individuation or "topic-worthiness" is a more likely candidate.[2]

(2) The D-hierarchy:

| 1Pro | 2Pro | 3Pro | Proper Noun / Kin term | Human | Animate | Inanimate |
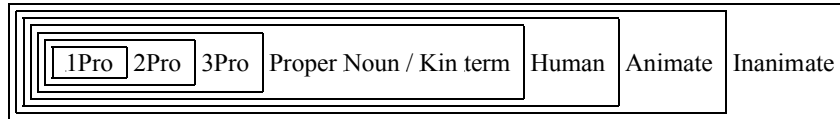
With these prerequisites in mind, I turn to clitic doubling in Albanian and Greek, which unlike standard Spanish and Romanian, violates Kayne's Generalization.

3. Clitic doubling as DOM: the case of Albanian (and Greek)

Both Albanian and Greek only have object pronominal clitics. Only in Albanian but not in Greek dative/genitive objects and direct objects instantiated by local (i.e., 1st and 2nd person) full pronouns are invariably clitic doubled.[3] These facts alone point to clitic doubling as a DOM strategy, at least in Albanian, an issue that gains in significance when taking into account the nominative-accusative case syncretism for indefinites, illustrated in (4), versus the lack of such syncretism for definites, see (3).

(3) a. Lib**ri**   u botua.
        book.**the**$_{NOM}$ was published
        'The book was published.'
    b. Botoi   lib**rin**.
        published.3S   book.**the**$_{ACC}$
        'S/he published the book.'

(4) a. Një   libër   u botua.
        a   book   was published
        'A book was published.'
    b. Botoi   një   libër.
        published.3S   a   book
        'S/he published a book.'

A (pan-Balkan) aspect of clitic doubling as DOM is the fact that, as illustrated in (5) for Albanian, while both definites and indefinites with articles can be clitic doubled, bare indefinites cannot, a fact which directly relates to the D-hierarchy in (2).[4]

---

[2] As an anonymous reviewer justly notes, the animacy hierarchy has been discussed extensively in the functionalist literature, which is why the Wierzbicka/Kiparsky conjecture on the re-interpretation of the D-hierarchy in terms of individuation, or topic-worthiness is particularly important for the purposes of the present paper.

[3] Albanian and Greek have identical case systems except for the fact that the Greek counterpart of the Albanian dative is the genitive, the Greek dative having been supplanted by the genitive.

[4] It also entails that dative objects in Albanian cannot be instantiated by bare nouns, as these are invariably doubled. This is especially clear in the case of bare singulars, which cannot occur as dative objects (and subjects). Matters are however slightly more complex with plural bare nouns; crucially, however, unlike direct objects, dative objects instantiated by plural bare nouns are always interchangeable (i.e. semantically equivalent) with their

(5) a. (E)   botoi   **librin**   më në fund.
       CL,ACC,3S   published.3S   book.the_ACC   at long last
       'S/he published the book.'
    b. (E)   botoi   **një libër**   më në fund.
       CL,ACC,3S   published.3S   a book   at long last
       'S/he published a book (at long last).'
    c. (*E)   botoi   **libër**   më në fund.
       CL,ACC,3S   published.3S   book   at long last
       'S/he published a book (at long last).'

As I have argued in Kallulli (1999 et seq.), a clitic doubled object functions as a familiarity topic, i.e. it is [+given], as illustrated by the complementarity of felicity conditions between the 'minimal pairs' in (6B)/(7B), where a doubling clitic is impossible in the given contexts (i.e. when the VP, or the object DP is focused), versus (8B)/(9B), where a doubling clitic must be present in Albanian, and is strongly preferred in Greek.

(6) A: What did Ana do?   B: Ana   (*e)   lexoi   libr**in**. (Albanian)
                                       I Ana (*to)   dhiavase **to** vivlio. (Greek)
                                       the Anna CL   read   **the** book
(7) A: What did Ana read?   B: Ana   (*e)   lexoi   librin.
                                       I Ana (*to)   dhiavase to vivlio.
(8) A: Who read the book?   B: Ana   *(e)   lexoi   librin.
                                       I Ana ?(to)   dhiavase to vivlio.
(9) A: What did Ana do to/with B: Ana   *(e)   lexoi   librin.
          the book?   I Ana ?(to)   dhiavase to vivlio.
                                       'Anna read the book'

Thus, it stands to reason that focused objects, among them *wh*-objects, cannot be doubled:

(10) a. Al: Së/çfarë   (*e)   pe?   (Kallulli 2000:220)
             [who_ACC/what]   CL,ACC,3S   saw-you
     b. Gr: Pjon/ti   (*ton/*to)   idhes?
             [who/what]_ACC   him/it_CL,ACC   saw-you
             'Who/what did you see?'

In contrast, the object of a subject question, forming part of the presupposition, must be doubled in Albanian and is strongly preferred doubled in Greek, too:

(11) a. Al: Kush   *(e)   pa   fëmijën?   (Kallulli 2000:220)
     b. Gr: Pios   ?(to)   ídhe   to pedhí?
             who   CL,ACC,3S saw   the child
             'Who has seen the child?'

---

definite plural analogues, which suggests that dative bare plurals are only seemingly bare, in the sense that there is morphological determiner drop which is however structurally present, which would account for the semantics. The fact that definite plural dative objects have an archaic flavour (see Kallulli 1999) corroborates this idea.

The link to presupposition is further highlighted through (12)/(13), which show that even for a "non-factive" verb such as 'believe' (Kiparsky and Kiparsky 1970), factivity can in fact be triggered by clitic pronouns 'doubling' the clausal complement, though this 'doubled' CP, in turn, can be said to be the complement of an empty D-head corresponding to the pleonastic *it* in English or the so-called 'correlate' *es* in German, which likewise trigger factivity in these languages (for details, see Kallulli 2006). That is, factivity is the correlate of givenness, or topichood, in the propositional domain.

(12) a. Besova   se Beni shkoi   (por në fakt ai nuk shkoi). (Kallulli 2006: 212)
     believed.I   that Ben left   (but in fact he not left)
     'I believed that Ben left (but in fact he didn't).'

     b. **E**   besova   se Beni shkoi   (*por në fakt ai nuk shkoi).
     CL,ACC,3S   believed.I   that Ben left   (but in fact he not left)
     'I believed the fact that Ben left (*but in fact he didn't).'

(13) a. Pistepsa   oti o Janis efije (ala stin pragmatikotita den ejine kati tetio).
     believed.I   that the Janis left (but in.the reality neg happened a thing such)
     'I believed that John left (but in fact he didn't).'

     b. **To** pistepsa oti o Janis efije (*ala stin pragmatikotita den ejine kati tetio).
     it$_{cl}$ believed.I that the Janis left (but in.the reality neg happened sth such)
     'I believed the fact that John left (*but in fact he didn't).'

In sum, (at least direct object) clitic doubling is a topic-licensing operation. Clearly, however, as (8B)/(9B) highlight, the grammaticalization of this phenomenon across Albanian and Greek is at different stages. Greek clitic doubling has been claimed to be dependent on definiteness (Anagnostopoulou 1994), though this is a contentious issue given the existence of doubled indefinites, as in (14b), from Kazazis and Pentheroudakis (1976).[5]

(14) a. Do t**a**   pija   me kënaqësi   **një**   uiski.   (Albanian)
     FUT SUBJ.**CL**$_{3S,ACC}$ drink.I with pleasure   **a**   whisky

     b. **To**   pino   eukharistos   **ena**   ouiskaki.   (Greek)
     it   I-drink with-pleasure   one   whisky
     'I would gladly drink a whisky.'

Be it as it may, even if in Greek clitic doubling of indefinites is more restricted than in Albanian, it is certainly not the case that definite expressions can always be clitic doubled; recall the (Greek) examples in (6B) and (7B). Crucially, just like clitic doubled definites, clitic doubled indefinites are necessarily interpreted as [+given]/[+topic], i.e. they are 'non-novel' in the sense of Krifka (2001). That the doubled indefinites in (14) are non-novel is evidenced by several diagnostics. First, just like the 'doubled' CPs in (12b) and (13b), they are deaccented; that is, the nuclear pitch accent cannot be borne by the clitic doubled expression. Secondly, the indefinite picks up a discourse referent whose existence in the input context is obviously presupposed, as can be seen by the fact that the sentences in (14) can be uttered felicitously in either of the contexts in (15); while the clitic doubled indefinite in (14a,b) function as a kind of quotation in the context of (15a), it stands in a

---

[5] Anagnostopoulou (1994) does in fact acknowledge this example as a counterexample to her claim that Greek clitic doubling is contingent on definiteness.

part-whole relationship with the indefinite 'a drink' in (15b), and its referent is presupposed through accommodation in the context of (15c).

(15) a. What about a whisky? / Would you like a whisky?
    b. What about a drink? / Would you like a drink?
    c. I have just stepped out of work.

Turning to the (other) differences between Albanian and Greek clitic doubling, given that datives in Albanian are invariably clitic doubled, as are direct objects instantiated by local pronouns, it seems sensible to describe the function of doubling clitics as mere object agreement markers in such configurations. But if the nature of agreement and topic markers is indeed substantially different, why are doubling clitics employed as means for fulfilling both functions? I suggest that these two seemingly different functions are not that different after all, and that crucially, clitic doubling is always agreement with a topic (object) DP, which is fully in line with Givón's (1975) claim that (object) agreement is topic-verb agreement. Furthermore, I hypothesize that PCC effects (Perlmutter 1971, Bonet 1991) in Albanian and other languages arise due to the competition for topic-prominence. Note in this context that Albanian is a so-called "strong" PCC language, in which local (i.e. 1st and 2nd) person direct objects in double object constructions are banned. In other words, I hypothesize that PCC effects within and across languages are straightforwardly derived from the D-hierarchy given in (2), i.e. they arise due to competition for the first slot (occupied by local pronouns) in the D-hierarchy.

The corollary that I want to add to Givón's (1975) claim that (object) agreement is topic-verb agreement is the one implicitly contained in the D-hierarchy. I contend that datives and (local) pronouns, just like subjects, are always DPs (i.e., they always contain a D-projection), which is however not invariably the case for (non-pronominal) direct objects. Indeed as already mentioned and illustrated in (5c) for Albanian, bare singulars can only occur as direct objects, not as indirect objects. And as also noted earlier, bare singulars cannot be clitic doubled either in Albanian or in Greek. The implication here is that bare singulars are truly bare, in the sense that there is no D-layer projected in their structure. Interestingly, as I have argued in Kallulli (2005), bare singulars just like their plural counterparts (i.e. existential bare plurals) cannot serve as topics, which in Kallulli (2005) I relate to their property-denoting (i.e. semantic <e,t> type) status.[6] Consequently, datives (and subjects) are presuppositional (or, in Kiparsky's 2008 terminology "topic-worthy"/"individuated") in a way that direct objects are not,[7] and this is precisely what their marking (via clitic doubling) relates to.[8] It is therefore not surprising to witness the effects that the phenomenon of clitic doubling produces also in languages with no canonical clitic doubling, such as Hungarian, which uses a particular conjugation, namely the so-called

---

[6] An anonymous reviewer notes that "in a language like Italian bare plurals and even bare singulars can appear in left topicalization", doubled with *ne* 'them/some', as in: *Orsi, ne ho visti* ('Bears, I have seen them/some'), or: *(Di) orso, ne aveva visto uno* ('As for a bear, I have seen one'), and raises the question of how the semantics of Albanian bare nouns relates to these. Obviously the very fact that such differences exist between the two languages (bare nouns cannot be clitic left dislocated in Albanian) suggests that the relevant structures (and consequently their semantics) are different.

[7] It follows that datives (and subjects), when focused, can only be contrastive topics. The fact that quirky subjects across Balkan languages must be clitic doubled (see Krapova and Cinque 2008) is further evidence of the topic-worthiness of such quirky (i.e. dative and/or accusative) subjects.

[8] While subjects are not clitic doubled in any of the Balkan languages, they invariably agree with the finite verb. That is, subject-verb agreement can be viewed as a further stage in this grammaticalization process.

"definite" (or "objective") conjugation, which I turn to after discussing the following fact that further corroborates the analysis outlined so far and in particular the role of D in the structure. As discussed in Kallulli and Tasmowski (2008), across several clitic doubling languages, 'all'-quantifiers are invariably clitic doubled:

(16) a. *(I)   pashë   të gjithë.   (Albanian)
     b. *(Los)   vi   todos.   (Argentinian Spanish)
     c. *(Tus)   idha   olus.   (Greek)
     d. *(I-)am   văzut   pe toți.   (Romanian)
        them$_{CL,ACC}$ saw.1S   all$_{ACC}$
        'I saw them all'

This fact can be straightforwardly derived by combining the view that the restrictor of 'all'-quantifiers is always 'given' (i.e. non-novel) with the idea that doubling clitics trigger givenness of their associates (Kallulli 2006). In line with Matthewson's (2001) analysis of quantification, (17b), which differs from the traditional one in Barwise and Cooper (1981), (17a), in that the generalized quantifier is formed in two steps, crucially involving an intermediate DP-layer as the complement of the quantifier head, I assume that the restrictor of 'all' is invariably a DP, both when it is phonetically overt or null. As such, the clitic does not double the quantifier (QP) but only its DP-complement, which being 'given', may be silent (Merchant 2001).[9]

(17)

a.
```
                    DP
                ⟨⟨e, t⟩, t⟩
               /          \
          D                    NP
  ⟨⟨e, t⟩, ⟨⟨e, t⟩, t⟩⟩        ⟨e, t⟩
         most                 chiefs
```

b.
```
                    DP
                ⟨⟨e, t⟩, t⟩
               /          \
          Q                    DP
  ⟨e, ⟨⟨e, t⟩, t⟩⟩             ⟨e⟩
                             /      \
                         D              NP
                    ⟨⟨e, t⟩, e⟩       ⟨e, t⟩
```

(Matthewson 2001: 146)   (Matthewson 2001: 153)

---

[9] As both reviewers note, undoubtedly there is more to say on the interaction of clitic doubling with (other) quantifiers. Luckily, (part of) this work has already been done in Dočekal and Kallulli (2012), whose conclusions that: (i) the DP associated with the clitic (i.e. the 'doubled' DP) must be interpreted as generating admissible minimal witnesses, which in turn makes the DP topical; and that (ii) as a consequence of (i), clitic doubling systematically produces information structure effects in that the doubled DP is unequivocally interpreted as topical, are fully in line with my analysis here. Crucially, Dočekal and Kallulli (2012) also provide conclusive evidence from Albanian against an analysis of clitic doubling along the lines of Gutiérrez-Rexach (1999) for Spanish.

4. *Other guises of the D-hierarchy: the Hungarian 'objective' conjugation*

Hungarian verbs have two subject agreement inflectional paradigms, the so-called "objective" (or "definite") and "subjective" (or "indefinite") conjugations, reflecting the presence or absence of a definite object, as in (18a,b,c), from Coppock (2013).[10]

(18) a. Lát-om   **a** madar-at.
          see-1S,DEF   the bird-acc
          'I see the bird.'
     b. Lát-ok   **egy** madar-at.
          see-1S,INDEF   a   bird-acc
          'I see a bird'
     c. Vár-ok.
          wait-1S,INDEF
          'I'm waiting'

Person also affects the choice of conjugation: the subjective conjugation is used with 1st and 2nd person objects, despite their definiteness:[11]

(19) Lát-nak   engem/téged/minket/…
     see-3PL,INDEF   me_ACC/you_ACC/us_ACC/…
     'They see me/you/us'

While the distribution of the definite conjugation is rather complex (see É. Kiss 2002, 2005, 2013, Coppock and Wechsler 2012, Coppock 2013 and references therein), (19) identifies an obvious gap, in that, as Coppock (2013) points out, first and second person non-reflexive, non-reciprocal pronouns are certainly definite, and under the hypothesis that the objective conjugation is governed by definiteness, they should trigger the objective conjugation. In spite of this, a dominant view on what conditions the use of the objective conjugation is what Coppock and Wechsler (2012) refer to as the "DP-hood hypothesis", and which merely states that DP-hood is a necessary but not sufficient condition for the objective conjugation. In fact, Coppock (2013) relates the objective conjugation to *familiarity* (i.e. non-novelty), which she however connects to a morphological [+definiteness] feature as follows: if the referential argument of a phrase is lexically specified as *familiar*, then the phrase bears the feature [+def], and this feature governs the objective conjugation. (If, on the other hand, the referential argument of a phrase is specified as *new*, then the phrase bears the feature [-def], which governs the subjective conjugation.) Coppock defines the notion 'referential argument' as follows: "The referential argument of a phrase is the discourse referent *u* such that: when the phrase combines an expression denoting property P, P is predicated of u." While this comes quite close to the

---

[10] According to É. Kiss (2013), the types of objects eliciting the definite conjugation include nouns with a definite article, possessive constructions, proper names, 3rd person personal pronouns, reflexive pronouns (which have the morphological make-up of possessive constructions of the type 'my body', 'your body'), and demonstratives. Object clauses also trigger the definite conjugation, which É. Kiss ascribes to a presumable overt or covert pronominal head in their structure. On the other hand, the types of objects eliciting the indefinite conjugation include bare nouns, nouns with an indefinite determiner, and indefinite and universal pronouns.

[11] There are further qualifications, or exceptions, to this generalization, which I will however not go into as they are not important in the present context. For details on these exceptions, see Coppock (2013) and references therein.

trigger of clitic doubling in Albanian and Greek, it doesn't account for the gap in (19), since the parallelism with Albanian (though not with Greek) breaks down here. Coppock claims that the person effect (i.e. (19)) is explained on the grounds that local non-reflexive pronouns are not anaphoric but rather purely indexical, unlike third person pronouns and local reflexive and reciprocal pronouns. This is also the most problematic part of her account, since, even though she takes familiarity to be broader than anaphoricity, indexicals are obviously not given in the associated context for her, which is puzzling.

Interestingly, É. Kiss (2013) reports that although object noun phrases supplied with indefinite determiners require the indefinite conjugation according to all grammars of Modern Hungarian, examples are ample where speakers hesitate whether the indefinite or the definite conjugation is more appropriate, often accepting both, or preferring the definite conjugation. Furthermore, the examples in the case of which the unexpected definite conjugation is accepted, and even preferred by the majority of speakers (up to 85% of them) all involve a topicalized [+specific] indefinite object.

Taken together, these facts suggest that an account of the impossibility of the use of the objective conjugation in (19) along the lines of É. Kiss (2005, 2013), according to which this gap is a manifestation of the Inverse Agreement Constraint (Comrie 1980) which blocks object verb agreement if the object is higher ranked in the animacy hierarchy than the subject, seems more feasible, as it is more in tune both with phenomena known from other languages, including languages genetically related to Hungarian (e.g. Ostyak), and as it can be easily derived from the universal D-hierarchy in (2). In particular, relying on Nikolaeva's (2001) work on Ostyak, a sister language of Hungarian, which shows that agreeing objects function as secondary topics, É. Kiss (2013) suggests that the objective conjugation in Hungarian is fossilized topic-verb agreement. The object thus competes with the subject for the higher slot in (2), and the objective conjugation is ruled out in this case because of the Inverse Agreement Constraint, in a manner that is analogous to the PCC effects observed in Albanian and other languages (section 3). That is, 1st and 2nd person pronoun objects which trigger the subjective conjugation in Hungarian (i.e. the data in (19)) are no less DPs than e.g. their Albanian (direct object) counterparts, which are invariably clitic doubled, and cannot partake in a clitic cluster given that dative objects, which asymmetrically c-command direct objects (Massey 1992), are invariably doubled, too, hence giving rise to PCC effects.[12]

5. Conclusions

Undoubtedly, clitic doubling is a form of DOM, which our faculty of language includes as a possible ingredient of a natural language, and which I have related to the D-hierarchy, a "true intrinsic universal" in the sense of Universal Grammar (Kiparsky 2008). Furthermore, I have contended that PCC effects arise as a result of competition for the same slot within the D-hierarchy, a conclusion that is corroborated by the distribution of the Hungarian objective conjugation. An outcome of this comparison is the idea that both PCC and the Inverse Agreement Constraint are direct consequences of the D-hierarchy.

In other words, the D-hierarchy that governs split case assignment, number marking, and agreement (for details, see Kiparsky 2008) also governs PCC and the Inverse Agreement

---

[12] The order within a clitic cluster is invariably dative > accusative in Albanian (and genitive > accusative in Greek).