Testing linguistic theory and variation to their limits: The case of Romance

Adam Ledgeway

Electronic version
URL: http://journals.openedition.org/corpus/2408
ISSN: 1765-3126

Publisher
Bases ; corpus et langage - UMR 6039

Printed version
Date of publication: 1 January 2013
Number of pages: 271-327
ISSN: 1638-9808

Electronic reference
Adam Ledgeway, « Testing linguistic theory and variation to their limits: The case of Romance », Corpus [Online], 12 | 2013, Online since 14 May 2014, connection on 01 May 2019. URL : http://journals.openedition.org/corpus/2408

© Tous droits réservés
Testing linguistic theory and variation to their limits:  
The case of Romance

Adam LEDGEWAY  
University of Cambridge

Drawing on a number of corpus studies, including a considerable amount of data taken from my own corpora of textual and fieldwork studies on the dialects of Italy, I shall explore, in a manner which is accessible to both general scholars of the Romance languages and linguists, how the richly documented diachronic and synchronic variation exhibited by the Romance languages, and especially their dialectal varieties, offers an unparalleled wealth of linguistic data (often of a typologically exotic nature) of interest not just to Romance linguists, but also to general linguists. This perennially fertile and still under-utilized testing ground, I will show, has a central role to play in challenging linguistic orthodoxies and shaping and informing new ideas and perspectives about language change, structure and variation, and should therefore be at the forefront of linguistic research and accessible to the wider linguistic community. At the same time, a familiarity with current key ideas and assumptions in theoretical linguistics has an important role to play in understanding the structures and patterns of Romance, and, in particular, those known to us only through the texts of earlier periods where native speakers are not available to provide crucial grammaticality judgments and fill in the missing empirical pieces of the relevant puzzle.

Following a brief introduction in §1 to linguistic variation across Romance in relation to parameters, universals and language typology, I shall then explore in §2 some case studies of microvariation across Romance which highlight what Romance can do for syntactic theory by way of testing, challenging and expanding our theory of language and the empirical base. By the same token, the tools and insights of current theories of

Corpus n°12 « Dialectologie : corpus, atlas, analyses » (2013), 271-327
syntax can also be profitably used to throw light on many of the otherwise apparently inexplicable facts of Romance microvariation, the topic of the §3 where I shall highlight what syntactic theory can do for Romance through the exploration of a number of Romance case studies which have traditionally proven, at the very least, extremely difficult to interpret in a unitary and satisfactory fashion.

1. Linguistic variation

1.1 Parameters

One of the areas in which research into Romance has proven particularly influential is the investigation of linguistic parameters, those dimensions of linguistic variation along which natural languages are said to vary. To cite just one simple example, in his *Issues in Italian Syntax*, Rizzi (1982) observed that whereas in languages like English it is impossible to move a relative pronoun like *to whom* in (1a) from the embedded indirect question *which lies they told*... to the next clause up, such movement is possible in the corresponding Italian example (1b), where *a cui* successfully raises out of the embedded clause headed by *che bugie*.¹

[1]  a *your brother, [CP to whom [IP I wonder [CP which lies [IP they told to whom]]]] (Eng.)

b tuo fratello, [CP a cui [IP mi domando [CP che bugie [IP abbiano raccontato a cui]]]] (It.)

¹ The abbreviations for languages and localities cited in the following discussion are: Abr. (Abruzzese), Arl. (Ariellese, eastern Abruzzo), Bal. Cat. (Balearic Catalan), Cal. (Calabrian), Cat. (Catalan), Cic. (Cicagnino, Liguria), Cos. (Cosentino, northern Calabria), Crs. (Corsican), Ctz. (Catanzarese, central Calabria), Dmc. Sp. (Dominican Spanish), Eng. (English), Erv. (Ervedosa do Douro, north-eastern Portugal), Fr. (French), Ger. (German), Glc. (Galician), Gsc. (Gascon), It. (Italian), Leo. (Leonese), Lgd. (Lengadocien), Lul. (Lulese, Nuoro, Sardinia), Mes. (Messinese, north-eastern Sicily), Nap. (Neapolitan), Occ. (Occitan), Pal. (Palermitano), Pgl. (Pugliese), Pº. de At. (Póvoa de Atalaia, central-eastern Portugal), Rip. (Ripano, south-eastern Marches), Ro. (Romanian), Sal. (Salentino, southern Puglia), Sic. (Sicilian), Sp. (Spanish), Srd. (Sardinian), Tsc. (Tuscan), Umb. (Umbrian), Wal. (Walloon, southern Belgium).
Rizzi’s solution, subsequently widely adopted, was to propose that the difference between English and Italian observable in (1) is a consequence of a different parametric setting on the boundaries for movement in natural languages (viz. subjacency): in English inflected verb phrases (IPs) constitute boundaries for movement, of which a maximum of only one can be crossed in any one derivation, whereas in Italian complementizer phrases (CPs) count as boundaries for movement, of which again only one can at most be crossed. Therefore, the ungrammaticality of (1a) follows because the phrase to whom illicitly moves over two boundaries, namely I wonder and they told, whereas (1b) is grammatical because the relative a cui crosses only one boundary, the CP introduced by che bugie.

In more recent years the significance of Romance dialects for the generative enterprise has also been increasingly recognized, inasmuch as they offer fertile, and often virgin, territory in which to profitably study parametric variation. While neighbouring dialects tend to be closely related to each other, manifestly displaying in most cases a high degree of structural homogeneity, they do nonetheless often diverge minimally in significant ways which allow the linguist to isolate and observe what lies behind surface differences in particular parametric settings across a range of otherwise highly homogenized grammars. By drawing on such microvariation, it is possible to determine which phenomena are correlated with particular parametric options and how such relationships are mapped onto the syntax. A clear example of such reasoning comes from the so-called dative shift construction, a phenomenon attested in a number of Germanic languages whereby an underlying indirect object, such as to Mary in (2a), can be reanalyzed and promoted to direct object, where Mary now comes to precede the old direct object a book. Furthermore, it has been claimed that the possibility of dative shift is linked to another structural property, that of stranding prepositions in wh-questions and relative clauses, as demonstrated in (2b)

b  [Who] did John give a book [pp to [dp who]] ?
In contrast, it is claimed that Romance languages do not display either dative shift or preposition stranding (Kayne 1984; Larson 1988: 378; Demonte 1995: 8), as the sharp ungrammaticality of the Italian examples in (3) demonstrates:

[3]  a  Ugo diede un libro [a Maria] /*[Maria] un libro (It.)
     Ugo gave a book to Maria / Maria a book
  b *[Chi] diede Ugo un libro [pp a [De chi]] ? (It.)
     Who gave Ugo a book to

This apparent Germanic vs Romance parametric variation is complicated by the fact that something very similar, if not identical to dative shift, is found in many southern Italian dialects (Sornicola 1997: 35-36; Ledgeway 2000: 46-52; 2009a: 844-847), witness the representative Neapolitan and Cosentino examples in (4):

[4]  a  nce purtaje na torta [a Mario] (Nap.)
     to.him= he.brought a cake to Mario
     “he brought a cake to Mario”
  b  o purtaje [a Mario] na torta (Nap.)
     him= he.brought PA Mario a cake
     “he brought Mario a cake”
  c *[Chi] purtaje na torta [pp a [De chi]] ? (Nap.)
     who he.brought a cake to
     “Who did he take a cake to ?”
  d  cci scrivu (na littera) (Cos.)
     to.him= I.write a letter
  e  u scrivu (*na littera) (Cos.)
     him= I.write a letter
     “I’Il write him (a letter)”

Beginning with the Neapolitan example (4b), we see that the RECIPIENT argument a Mario, the underlying indirect object in (4a), has been advanced to direct object and is now marked by the prepositional accusative (PA), as shown by its position in front of the old direct object na torta and by the fact that a Mario is now referenced by an accusative clitic pronoun o “him”, and not the third person dative pronoun nce in (4a). Other neighbouring southern dialects such as Cosentino, by contrast, exhibit a more constrained type of dative shift, inasmuch as RECIPIENT arguments (cf. dative clitic cci “to him” in
Testing linguistic theory and variation to their limits

(4d)) may only surface as direct objects (cf. accusative clitic u “him” in (4e)) in monotransitive clauses.

These Italian dialects reveal therefore three important things. Firstly, dative shift is not a Germanic vs Romance parametric option. Secondly, the supposed link between dative shift and preposition stranding, argued to be derivable from a single parametric option, does not hold, the presence of both phenomena in languages like English simply representing a fortuitous combination rather than the principled outcome of a particular parameter setting, since preposition stranding is not found in these same southern dialects (cf. 4c). Thirdly, it is incorrect to subsume all instances of accusative marking of RECIPIENT arguments under the generic heading of dative shift, since some of the dialects of southern Italy prove sensitive to the mono- vs ditransitive distinction (cf. 4e). It follows therefore that what might otherwise be taken to represent the surface reflexes of a single parametric setting, in dialects like Neapolitan, namely the accusative marking of all RECIPIENT arguments irrespective of whether they occur in mono- or ditransitive clauses, turns out in fact to conceal two distinct structural operations in the light of evidence gleaned from dialects such as Cosentino.

1.2 Language universals

Romance varieties also have much to contribute in the area of so-called universal principles of language, essentially a system of rules forming part of the genetic endowment known as Universal Grammar which are believed to hold of all human languages. A good illustration of the valuable role that Romance can play in testing linguistic universals concerns the licensing of nominative Case. Within current theory, it is assumed that Infl, the locus of verbal inflection, may be specified as [+tense] or [-tense], featural specifications which in turn are argued to correlate respectively with the verb’s ability or otherwise to license a nominative Case-marked subject. This [+tense] distinction is supported by the evidence of many of the world’s languages, including French where tensed verbs license nominative subjects (5a), but untensed verbs such as infinitives and gerunds only allow null (Caseless) PRO subjects (5b):
Yet, the evidence of Romance dialects reveals that the supposed universal correlation between the specification of Infl and the availability of nominative Case is entirely spurious (Ledgeway 1998; 2000: ch. 4; Mensching 2000). In particular, dialects from the length and breadth of the Romance-speaking world demonstrate an abundant use of overt nominative subjects in conjunction with infinitival verbs:

1.3 Typological variation

Data like those exemplified in (6) also illustrate how investigations of Romance dialects frequently reveal that the extent of typological variation within Romance, and indeed even within Indo-European, can in particular cases prove to be considerably greater than is traditionally assumed. In this respect, one only has to think of such examples as the Romance inflected infinitives, gerunds and participles illustrated in Tables 1 and 2 with representative examples in (7)\(^2\), which, as intermediate cate-

---

Testing linguistic theory and variation to their limits

gories, clearly throw into turmoil traditionally narrow interpretations of finiteness in terms of a binary finite vs non-finite dichotomy (cf. Ledgeway 2007c):

Table 1. Romance inflected infinitives (cant- “sing”)

<table>
<thead>
<tr>
<th></th>
<th>Infinitive</th>
<th>1sg</th>
<th>2sg</th>
<th>3sg</th>
<th>1pl</th>
<th>2pl</th>
<th>3pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glc.</td>
<td>cantar</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>OLeo.</td>
<td>cantar</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>ONap.</td>
<td>cantar(e)</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>(E)Pt.</td>
<td>cantar</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>Srd.</td>
<td>cante(e)</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
</tbody>
</table>

Table 2. Romance inflected gerunds & participles (cant- “sing”)

<table>
<thead>
<tr>
<th></th>
<th>Gerund</th>
<th>Present/Past Participles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glc.</td>
<td>Pv. de At.</td>
</tr>
<tr>
<td></td>
<td>cantando</td>
<td>cantando(o)</td>
</tr>
<tr>
<td>1sg</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>2sg</td>
<td>-Ø</td>
<td>-s</td>
</tr>
<tr>
<td>3sg</td>
<td>-Ø</td>
<td>-Ø</td>
</tr>
<tr>
<td>1pl</td>
<td>-mos</td>
<td>-mos</td>
</tr>
<tr>
<td>2pl</td>
<td>-des</td>
<td>-eis</td>
</tr>
<tr>
<td>3pl</td>
<td>–</td>
<td>-em</td>
</tr>
</tbody>
</table>

[7] a é doado supoñer as cousas (Glc.)
   “it is easy for them to assume things”
   b não saindomos de casa, morrâmos à fome (Erv.)
   “if we don’t leave the house, we’ll starve to death”
   c dato no sasse insembla salute como convenne (ONap.)
   “after having greeted one another as was customary”

Another acute example comes from the unique infectious development of inflection in the Marchigiano dialect of Ripatransone. Simplifying somewhat, in addition to the usual person/number agreement, the Ripano finite verb simultaneously displays


masculine/feminine gender agreement with the subject, not to mention the possibility of agreement with a so-called third person neuter subject, by means of final inflectional vowel contrasts which, to all appearances, seem to have been grafted onto the verbal paradigm from the nominal paradigm. By way of example, we illustrate below in Table 3 the nominal paradigm for *fijj-* “son, daughter” and *prafutta* “ham” together with the present tense paradigm of the verb *magnà* “to eat” and *piovèr* “to rain”:

Table 3. Ripano nominal & present indicative paradigms

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg.</td>
<td>fijjì</td>
<td>fijjë</td>
<td>prafutta</td>
<td>1sg</td>
<td>magnà</td>
<td>magne</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2sg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3sg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pl.</td>
<td>fijjì</td>
<td>fijjë</td>
<td></td>
<td>1pl</td>
<td>magnèmì</td>
<td>magnèma</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2pl</td>
<td>magnìtì</td>
<td>magnìta</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3pl</td>
<td>magnì</td>
<td>magna</td>
<td></td>
</tr>
</tbody>
</table>

Far from being limited to finite verbs, such a rich and complex system of agreement has come to permeate even non-finite verb forms such as past participles (8a), gerunds (8b) and infinitives (8c), not to mention other parts of speech including predicative nominal complements of the copula *avè* “to have” (9a), wh-interrogatives (9b) and quite remarkably, on a par with such languages as Welsh, prepositions (9c):

[8] a lu fràki / le fràkine è ddàrmìtu / ddàrmìte (Rip.)
the boy the girl is slept.msg slept.fsg
“the boy / the girl has slept”

b li fràki stievì currenì (Rip.)
“the boys were running.mpl”

c sai skrivù/skrive? (Rip.)
you(msg/fsg).know to.write.msg/fsg
“do you know how to write?”

[9] a c’ajjì famù / c’ajjë famë (Rip.)
I.have.msg hunger/msg  I.have.fsg hunger.fsg
“I am hungry”

b quannù / quanne passu / passe îlùka (Rip.)
when.msg when.fsg pass.msg pass.fsg there
“whenever I/you(s)he pass(es) by there”
1.4 Interim conclusions

From the preceding introductory discussion, it is clear that the wealth of Romance standard and especially dialectal data, although frequently overlooked in the past, have a great deal to contribute to research into such areas as parametric variation, linguistic universals and typological variation. Nonetheless, the syntax of the dialects still represents a relatively poorly understood area of Romance linguistics, to the extent that there still remains a considerable amount of fieldwork to be done in recording and cataloguing the linguistic diversity within the Romania, as well as in bringing such facts to the attention of the wider linguistic community. With this in mind, in what follows we shall offer a number of valuable insights into the little-studied syntax of the Romance dialects in an attempt to highlight their significance for issues in general linguistic theory and their potential as relatively unexplored experimental territory in which to investigate new ideas about language structure, change and variation. Keeping the technical detail to a minimum, we shall discuss a number of issues relating to the syntax of Romance under the two broad headings of what Romance can do for syntactic theory and what syntactic theory can do for Romance. Under the former heading we shall review a number of assumptions about language structure and variation that have been proposed in the literature, demonstrating how in the specific cases examined the Romance dialect data contradict such principles and parameters, rendering them either invalid or in need of further elaboration. Under the latter heading, by contrast, we shall bring to light some of the less familiar and more problematic aspects of Romance syntax which can be shown to find an enlightening interpretation in light of current theoretical assumptions.
2. What Romance can do for syntactic theory

2.1 Pro-drop parameter

Undoubtedly, one of the best known and most widely studied parameters is the so-called pro-drop (or null subject) parameter (for a recent overview, see Biberauer, Holmberg, Roberts and Sheehan 2009). Limiting our attention to Romance and Germanic, it is traditionally claimed that, with the exception of modern French, morphological Agr(eement) for person and number on the Romance verb is sufficiently rich to license a null subject (11a), whereas in such languages as English Agr is so impoverished that that it is unable to recover the identity of a null pronominal subject which must instead be phonologically expressed (10a). By the same token, it is also assumed (cf. Chomsky 1981: 28; Rizzi 1986: 410; Haider 2001: 285) that expletive (or non-referential) pronouns are null in the former (11b) but overt in the latter (10b).

4 On the status and distribution of expletive subject clitics, see Oliviéri (2009) and Kaiser, Oliviéri and Palasis (in press).

On a par with others, Rizzi (1982: 143ff.) derives this supposed universal distinction from the pro-drop parameter, which he argues yields the four language types illustrated in Table 4:

Table 4. Typology of null subjects

<table>
<thead>
<tr>
<th>Pronoun</th>
<th>Type 1 English</th>
<th>Type 2 Spanish</th>
<th>Type 3 German</th>
<th>Type 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null referential</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Null expletive</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

Language types 1 and 2 are exemplified by English and Spanish, respectively. In Spanish both null expletives and null referential pronouns are licensed, whereas in English both types...
of null pronoun are excluded. Type 3 is argued to characterize German where, in contrast to referential pronouns which are invariably overt (12a), overt expletive pronouns are only licensed when they occur in clause-initial position (12b):

[12] a Er/*Ø weint (Ger.)
He/*Ø cries
b Es wird heute getanzt / Heute wird Ø/*es getanzt (Ger.)
it is today danced / today is Ø / it danced
“There is dancing going on today”

On the other hand, Rizzi (1982: 143) explicitly argues that type 4 languages with overt expletive subjects but null referential subjects are “excluded for intrinsic reasons”. However, the evidence of a number of non-standard Romance varieties demonstrates that type 4 languages do indeed exist. For example, although Neapolitan is a pro-drop language (13a), it also displays structures such as those in (13b) where the subject position is filled by the overt expletive chello “that” (Sornicola 1996; Ledgeway 2009a: 290-294; 2010), a pattern replicated by a number of other Romance varieties (14a-e).

[13] a Ø/Isso chiagne (Nap.)
Ø/He cries
b Ø/Chello chiove (Nap.)
Ø/That rains
[14] a El chove (Glc.)
“it rains”
b Ello estaba lloviznando un poco (Dmc. Sp.)
“it was drizzling a little”
c Ell és veritat! (Bal. Cat.)
“it is true!”
d Iddu cchi mm’ importa? (Ctz.)
it what me= matters
“What does it matter to me?”
e Ma Iddu chi cc’ è cosa? (Pal.)
but it what there= is what
“What is it?”
f ko pľọj (Corrèze, nth. Occ.)
that rains
“It is raining”

With the exception of the northern Occitan example in (16f), for which see Kaiser, Oliviéri and Palasis (in press), in most cases
the use of the overt expletive in such examples is associated with specific pragmatic functions (Sornicola 1996: 325-326; Ledgeway 2003; 2009a: 290-291; Hinzelin 2009), typically marking the illocutionary force of the clause as exclamative (cf. 14c) or interrogative (often with rhetorical overtones; cf. 14d-e), a usage which still requires much more detailed investigation. Summing up, the data considered here lead us to conclude, with Kaiser, Oliviéri and Palasis (in press), that the ability to drop referential pronouns and the availability of overt expletives are not necessarily mutually exclusive or, for that matter, two interrelated properties of a single pro-drop parameter. At the same time, however, the overt expletives of most of these Romance dialects cannot be equated tout court with those of languages like English and French in view of the marked pragmatic functions of the former and the purely syntactic nature of the latter.

The wealth and extent of the Romance evidence is such that our typology of null subjects does not, however, stop here. Turning our attention to medieval Romance, it has long been noted (Adams 1987; Roberts 1993: §3.2; Benincà 1994; 2006; 2010: §3.2.1; Salvi 2004: 16f., 26-31; Ledgeway 2012a: 74-75) that these varieties display an asymmetric distribution of null subjects, a pattern more robustly represented in Gallo-Romance (French, Occitan, and northern Italian dialects) than either in Ibero-Romance or central-southern Italo-Romance (Salvi 2004: 30f.). In particular, whereas in root clauses null subjects are freely licensed (15a-c), in subordinate clauses pronominal subjects must usually be phonologically expressed (16a-c), although not interpreted as emphatic or contrastively-focused. Illustrative in this respect is the old Umbro-Tuscan example in (16c), where, despite the coreference of main and embedded clause subjects, the latter is overtly realized yielding a structure which would be judged ungrammatical, for example, in modern Italian where the presence of an overt pronoun in the same context would typically signal switch reference.

[15] a Si Θ errent tant en tele maniere qu’ […] (OFr.)
so (they) wander so much in such way that
“they thus wandered so much in such a way that […]”
Testing linguistic theory and variation to their limits

b Del cor Ø sospire dels olhs Ø plor (OOcc.)
of the heart (I) sigh and of the eyes (I) cry
“I sigh from the heart and cry from the eyes”
c manifestamente Ø l’ho vedutonelle cose […] (OTsc.)
manifestly (I) it=have seen in the things
“I have seen this clearly in those things […]”

[16] a Endemiers que il regadoient les letres et la
while that they looked at the letters and the
damoisele, que il plagnoient sa mescheance (OFr.)
young lady that they regretted her misfortune
“While [that] they looked at the letters and the young lady,
[and that] they regretted her misfortune”
b Non es meravelha s’eu chan / Melhs de nul autre
not is wonder if I sing better of no other
chantador (OOcc.)
singer
“It is no wonder if I am a better singer than all others”
c Elli conosce certamente ch’elli avea ucciso lo migliore
he knows certainly that he had killed the best
cavaliere del mondo (OUmb.-Tsc.)
knight of the world
“He certainly knows that he had killed the best knight in the
land”

This asymmetrical distribution leads us to conclude that null subjects in medieval Romance were not licensed exclusively, if at all, by rich verb inflection for person and number, insofar as verbal Agr is equally rich in both root and subordinated clauses, but by a structural property (to be discussed below in §3.2) which aligns the verb with distinct positions in both clause types. Indeed, the unreliability of verb morphology as a diagnostic for correctly predicting the distribution of null subjects is further evidenced by Ripano, a pro-drop variety, whose verb was seen in Table 3 above, witness the three singular persons of the paradigm (namely. magna (m.) vs magné (f.),) to privilege the marking of gender over that of person, despite traditional claims that the licensing of null subjects is directly linked to the overt marking of the person feature.

In summary, the null pronominal types seen in medieval Romance and Ripano cannot be readily accommodated in terms of current theories of a binary null subject parameter (cf. also Oliviéri 2009; Kaiser, Oliviéri and Palasis in press) and, in
particular, the typology of null subjects predicted by such models as that observed in Table 4 above which ultimately reduce tout court the distribution of null pronominals to the availability of rich morphological Agr for person on the finite verb.

2.2 Verb positions

The confines of the verbal domain are generally taken to extend beyond the immediate v-VP shell, where the external and internal arguments are first-merged, to include a dedicated structural position Infl(ection) (Chomsky 1981; 1986) > T(ense) (Chomsky 1995). The latter represents the locus of verbal inflection and licenses the external argument’s Case and phi-features with which it enters into a Spec-head configuration upon raising of the subject to its associated specifier position (SpecIP > SpecTP). Not only can this head position be lexicalized by distinct auxiliaries (cf. epistemic use of Catalan auxiliary deure “must” to express supposition in (17a)) but, in the absence of the latter, may be overtly filled by the raised lexical verb where its finite inflectional features can be licensed (cf. epistemic use of the future in substandard Catalan to express supposition in (17b); see Badia I Margarit 1962, I: 391).

[17] a
   [Infl deu [v-VP tenir raó]] (Cat.)
   he.must to.have reason

b [Infl undra [v-VP undra raó]] (coll. Cat.)
   he.will.have reason
   “he must surely be right”

These structural assumptions provide a natural way of interpreting the observed contrast in the (unmarked) position of the finite verb in languages like English (18a) on the one hand and French (18b) on the other:

[18] a John [Infl Ø [v-VP often skis]] (Eng.)

b Jean [Infl skie [v-VP souvent skie]] (Fr.)

Exploiting the fixed positions of VP-adverbs like “always” as a diagnostic indicator of the left edge of the v-VP complex, we can now straightforwardly distinguish between overt verb-raising languages like French, where the finite verb raises to the Infl position to the left of VP-adverbs, and languages like English, where the verb remains in situ to the right of such VP-adverbs

284
and the Infl position is not overtly lexicalized in the syntax, a difference traditionally retraced to the respective richness of verbal inflection in the two languages (Emonds 1978; Pollock 1989; Belletti 1990).

Nonetheless, recent research has revealed a much more nuanced interpretation of Romance verb movement than these familiar broad-brush treatments which classify Romance tout court as having overt verb movement. Following the seminal work of Cinque (1999), Infl is now commonly interpreted as a general label for the rich inflectional area of the clause (the I-domain) made up of a series of distinct functional projections dedicated to marking various temporal, aspectual, modal and voice distinctions ranging over the lexical verb, its arguments, and possible adjuncts which can also be identified by the semantically corresponding adverbial modifiers they host (cf. also Cinque 2002; 2006; Belletti 2004; Rizzi 2004). Armed with these assumptions about a universal fixed hierarchy of adverb positions and corresponding functional projections, we can now construct a fine-grained typology of Romance varieties along the lines of (19):

<table>
<thead>
<tr>
<th></th>
<th>Asp\textsubscript{habit.}</th>
<th>Asp\textsubscript{cont.}</th>
<th>Asp\textsubscript{compl.}</th>
<th>VP…</th>
</tr>
</thead>
<tbody>
<tr>
<td>(19a)</td>
<td>normalmente</td>
<td>todavía</td>
<td>veo</td>
<td>todo</td>
</tr>
<tr>
<td>(19b)</td>
<td>di solito</td>
<td>vedo ancora</td>
<td>vedo tutto</td>
<td>vedo (It.)</td>
</tr>
<tr>
<td>(19c)</td>
<td>je \textit{d’habitude}</td>
<td>vois encore</td>
<td>vois tout</td>
<td>vois (Fr.)</td>
</tr>
</tbody>
</table>

Although in all three varieties exemplified in (19) the finite lexical verb invariably leaves its base position to vacate the verb phrase, witness its position to the left of the completive adverb “everything” immediately adjacent to the VP, it raises to different functional projections within the I-domain as illustrated by its differential position with respect to different adverb classes. For example, in Spanish (19a) the finite verb raises to the head position of the continuative aspectual projection immediately below the adverb “still”, whereas in Italian (19b) it appears to raise slightly higher to the head position of the habitual aspectual projection below the adverb “usually”, and in French (19c) it raises to the highest available position above all adverb classes.
On the basis of evidence like this, let us assume that the various adverbs and their associated functional projections making up the inflectional core of the clause can be broadly divided into two “spaces” termed the Lower Adverb Space (LAS) and the Higher Adverb Space (HAS). Broadly speaking, the HAS comprises modal functional projections variously spelt out by speaker-oriented adverbs marking such categories as evidential (e.g. “apparently”), epistemic (e.g. “presumably”) and irrealis (e.g. “perhaps”) modality, whereas the LAS principally includes aspeaturesal functional projections lexicalized by adverbial classes including perfective (e.g. “always”) and continuative (e.g. “still”) aspects, with temporal projections and associated adverbs (e.g. T\textsubscript{Past} “then”, T\textsubscript{Anterior} “already”) variously scattered across both spaces. Assuming in this way the clause to consist of a highly articulated functional structure, we can interpret a number of surface differences across Romance in terms of the varying extent of verb movement around different adverb classes in these two adverbial spaces, as illustrated in (20a-e):

\begin{verbatim}
[20] [HAS [las [v-VP...]]]
  a forze già sacciu sacciu a ricetta (Cos.)
  b poate de| è| t| u| u| re| t| e| t| a (Ro.)
  c tal vez ya sé sé la receta (Sp.)
  d forse so già so la ricetta (It.)
  e je sais peut-être déjà sais la recette (Fr.)
\end{verbatim}

I know perhaps already know the recipe

It can be seen that whereas in French (20e) the finite lexical verb raises to the highest position within the HAS (Rowlett 2007: 106f.), in Calabrian (20a), Romanian (20b) and Spanish (20c) they typically raise only as far as the LAS. Italian (20d) represents an intermediate case insofar as the finite lexical verb targets a clause-medial position sandwiched between the HAS and the LAS (Cinque 1999: 31, 110f., 180 fn. 80; Ledgeway and Lombardi 2005: 87f.).

---

These same assumptions about the fixed positions of adverbs also allow us to plot the differential position of non-finite verbs such as the active participle in the following examples:6

<table>
<thead>
<tr>
<th></th>
<th>[HAS]</th>
<th>[LAS AspPerfect]</th>
<th>Voice</th>
<th>[v-VP...]]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>j’ai</td>
<td>toujours</td>
<td>bien</td>
<td>mangé</td>
</tr>
<tr>
<td>b</td>
<td>apo</td>
<td>semper mandigadu</td>
<td>bene</td>
<td>mandigadu</td>
</tr>
<tr>
<td>c</td>
<td>ho</td>
<td>mangiato sempre</td>
<td>bene</td>
<td>mangiato</td>
</tr>
<tr>
<td>d</td>
<td>aju</td>
<td>mangiatu sempe</td>
<td>buonu</td>
<td>mangiato</td>
</tr>
<tr>
<td>e</td>
<td>am</td>
<td>mâncat mereu</td>
<td>bine</td>
<td>mâncat</td>
</tr>
<tr>
<td>f</td>
<td>he</td>
<td>comido siempre</td>
<td>bien</td>
<td>comido</td>
</tr>
</tbody>
</table>

“I’ve eaten always eaten well eaten

In contrast to what was observed with the finite verb (cf. 20e), the active participle moves least in French (21a), staying very low in the clause within the LAS to the right of the manner adverb “well” just above the v-VP, whereas in Calabrian (21d), Romanian (21e) and Spanish (21f) the active participle, unlike the finite verb in (20a-c), moves to the highest available position in the HAS beyond the perfective aspectual projection spelt out by the “always” adverb. In Sardinian and Italian, on the other hand, the position of the active participle displays greater freedom, raising in Sardinian (21b) no higher than “always” in the LAS, and possibly above “always” to a position in the HAS in Italian (21c).

In the literature there is no general consensus regarding the correct interpretation of V-movement to different clausal positions (for an overview, see Schifano 2011; in prep.), although traditionally there have been many attempts, as noted above, to relate the extent of movement to the richness or otherwise of the inflectional Agr of the verb,7 witness Baker’s (1985; 1988: 13) Mirror Principle and Bobaljik’s (2002) Rich Agreement Hypothesis. In essence, approaches of this type attempt to drive syntactic operations from crosslinguistic morphological differences in individual languages. Admittedly, this view finds some initial support in the Germanic vs Romance contrast in (17a-b) where


the Romance verb form, undoubtedly the inflectionally richest of the two, raises the highest. However, a brief comparison of the results reported in (20)-(21) above suffices to dispel such an approach, inasmuch as all the Romance varieties we have examined are what may be termed inflectionally rich languages, yet they display some quite marked differences in the extent of finite lexical and non-finite V-movement. This conclusion is further substantiated by the observation that much of the rich inflection of the modern French verb, unlike that of the other varieties examined, is predominantly orthographic, yet it shows higher verb movement of finite lexical verbs to the HAS than all other varieties.

Thus, to conclude, we have seen that the rich comparative evidence of multiple verb positions across Romance forces us to recognize a richer functional structure for the clause than has traditionally been assumed. More generally, we have established that from an empirical and a theoretical perspective broad-brush characterizations of Romance as invariably involving overt V-raising prove neither descriptively nor explanatory adequate, inasmuch as a more nuanced picture has to be recognized. At the same time, this same evidence has highlighted the danger of assuming a direct correlation between the richness of inflection and the extent of V-movement.

2.3 Mapping the left periphery of the clause
As seen in the previous section, one area of considerable interest in much recent syntactic research has been the investigation of the fine structure of the clause. Traditionally, the structure of the clause in a typical SVO language, of which almost all the modern Romance varieties are examples, has been taken to present (at least) the positions indicated in the linear template in (22a), exemplified from Italian in (22b):

[22] a Subject Aux Adverb Verb Object Adjunct(s)

b Ugo ha sempre fatto la pasta a cena (It.)
Ugo has always made the pasta at dinner
“Ugo has always made pasta for dinner”

The linear arrangement in (22a-b) reflects the standard idea that the confines of the clauses can therefore be identified with the preverbal subject position situated at the left edge of IP and the
complement or adjunct position situated at the right edge of VP (23a). However, in recent years research within generative syntax has been increasingly directed towards the investigation of the fine structure of the left periphery, the syntactic space immediately to the left of the sentential core (23b), culminating in the seminal work of Rizzi (1997) which has given rise to a widely-accepted view of the fundamental cartography of the C(omplementizer)-domain (cf. also Benincà and Munaro 2010).

Significantly, a considerable amount of work on the split C-domain has been conducted on the basis of the rich (dialectal) variation offered by Romance varieties (for an overview and relevant bibliography, see Ledgeway 2012a: 154-171), which in many cases provide invaluable overt evidence with which to map the fine structural organization of the left periphery. In particular, the left periphery, traditionally defined in terms of CP and its associated Spec(ifier) and head positions hosting wh-operators and complementizers (Chomsky 1986: §1), respectively, is now conceived as a split domain, hierarchically articulated into several fields and associated projections. Revealing in this respect are Italian topicalization examples such as (24a-c), where one of the constituents of the sentential core in (22b) conveying old/given information has been fronted to the left periphery and picked up, where available, by a resumptive clitic pronoun on the verb in the sentential core, a case of (clitic) left dislocation.8 This, however, does not exhaust all available structural possibilities, witness the examples in (25a-d) where as many as two constituents have been fronted under (clitic) left dislocation, illustrating the fact that topic is a recursive category and capable of multiple realizations within the same utterance.

8 In these and following examples we indicate topicalized constituents by underlining, contrastively focalized constituents by small caps, and informationally focalized constituents with bold.
A. LEDGEWAY

[Topic\text{OLD} [S V Adv O (X)]]

[24] a [La pasta, [Ugo l’ha sempre fatta a cena]]
the pasta Ugo it=has always made at dinner
b [Ugo, [lui ha sempre fatto la pasta a cena]]
Ugo he=it=has always made the pasta at dinner
c [A cena, [Ugo ha sempre fatto la pasta]]
at dinner Ugo has always made the pasta

[Topic\text{OLD} [S V Adv O (X)]]

[25] a [Ugo la pasta, [lui l’ha sempre fatta a cena]]
Ugo the pasta he=it=has always made at dinner
b [La pasta, Ugo, [lui l’ha sempre fatta a cena]]
the pasta Ugo he=it=has always made at dinner
c [Ugo a cena, [Ugo ha sempre fatto la pasta]]
Ugo at dinner he=it=has always made the pasta
d [A cena, la pasta, [Ugo l’ha sempre fatta]]
at dinner the pasta Ugo it=has always made

Further support for the richly articulated structure of the left periphery comes from a consideration of focus structures such as the Italian examples in (26), where a constituent of the sentential core in (22b) is fronted under contrastive focus to correct a previous assertion (cf. 26a-c). Significantly, however, fronted focused constituents are not in complementary distribution with fronted topicalized constituents, as the traditional simplex CP structure with its single left-peripheral position would lead us to believe, but can co-occur with left-dislocated topics in the strict order Topic + Focus, as shown in (27a-d).

[Focus\text{NEW} [S V O (X)]]

[26] a [SEMPRE [Ugo ha fatto la pasta a cena]]
always Ugo has made the pasta at dinner
b [A CENA [Ugo ha sempre fatto la pasta]]
at dinner Ugo has always made the pasta
c [LA PASTA [Ugo ha sempre fatto a cena]]
the pasta Ugo has always made at dinner
Testing linguistic theory and variation to their limits

[27] a  [Topic + Focus  [S V O (X)]]
    [Ugo, SEMPRE  [lui, ha fatto la pasta a cena]]
    Ugo, always, he has made the pasta at dinner
b [Ugo, A CENA  [lui, ha sempre fatto la pasta]]
    Ugo at dinner, he has always made the pasta
c [La pasta, SEMPRE  [Ugo l’i, ha fatta a cena]]
    the pasta, always, Ugo it has made at dinner
d *[SEMPRE la pasta,  [Ugo l’i, ha fatta a cena]]
    always the pasta, it has made at dinner

Evidence like this from Italian and other Romance varieties suggests that the existence of a single left-peripheral position is empirically inadequate. Rather, the relevant left-peripheral positions must be reconceived as distinct &egrave;rgmatic-syntactic spaces along the lines of Benincà & Poletto (2004), according to which we can identify from left to right at least two fields termed Topic and Focus, respectively (see also Cruschina 2012). Not only is this demarcation between Topic and Focus justified at a &egrave;rgmatic-semantic level, in that elements appearing in the Topic field are generally interpreted as “old” or “given” information whereas the Focus field is typically associated with informationally “new” elements, but it also finds confirmation at the syntactic level. For instance, in contrast to elements appearing within the Topic field, which often call for a resumptive pronominal (clitic) where available (cf. 24a-b, 25a-d), those appearing within Focus (cf. 26d) typically prove incompatible with a pronominal copy (Benincà 2001: 43ff.). Moreover, we have already observed that topic is a recursive syntactic category allowing several reiterations within a given utterance, whereas focus is restricted to just one occurrence per utterance.

Robust evidence like this demonstrates that topicalized and focused constituents indeed target distinct spaces within the left periphery, forcing us to recognize a representation of the C-domain along the lines of (28) below:

[28]  [CP Comp [TopP Topic [FocP FOCUS [IP ...]]]]

Complementizer positions
In addition to the Topic and Focus fields highlighted above, the left periphery of the clause also hosts complementizers. Now
alongside finite complementizers derived from QUOD/QUID and QU(ī)A “(because ->) that” which introduce tensed clauses (cf. (a) examples below), Romance varieties also present a series of non-finite complementizers derived from the prepositions DE “of, from” and AD “to” to introduce infinitival clauses (cf. (b) examples below), which to all intents and purposes parallel the use of their finite counterparts:

[29] a  Digues-li [IP Øi/*j vingui] (Cat.)
tell=him,  that  Øi/*j comes.sbjv
   “Tell him to come”

b  Digues-li [IP Øi/*j venir] (Cat.)
tell=him,  of  Øi/*j to.come
   “Tell him to come”

[30] a  Il a appris qu’ [IP elle est morte] (Fr.)
he has learnt that she is died
   “He has learnt that she has died”

b  Il a appris à [IP Øi/*j nager] (Fr.)
he, has learnt to Øi/*j to.swim
   “He has learnt to swim”

Despite appearances, the presumed parallelism between finite and non-finite complementizers is not, however, perfect, as revealed by their respective positions in relation to topics and foci: whereas finite complementizers precede topics and foci (31a), non-finite complementizers invariably follow both types of fronted constituent (31b).

[31] a  So che, la data, [IP l’ ho sbagliata] (It.)
I.know that, the date, it= I.have mistaken

b  So, la data di [IP averla sbagliata] (It.)
I.know, the date, of to.have=it mistaken
   “I know (that), the date, I got (it) wrong”

Evidence like this forces us to assume that the Topic and Focus fields outlined above are, in turn, closed off upwards by a higher complementizer position (termed Force) marking the illocutionary force of the clause and hosting such items as the Romance finite declarative complementizer que/che “that”, and downwards by a complementizer position (termed Fin(iteness)) specifying the modality and/or finiteness of the clause and hosting such items as the Romance infinitival complementizers de/di “of”, as schematicized in (32).
Indeed, some Romance varieties present dual finite complementizer systems which appear to exploit both the higher and lower complementizer positions within the left periphery. Such is the case in Romanian and many southern Italian dialects,9 which contrast an indicative/declarative complementizer (QU(i)A “because” >) că/ca that lexicalizes the higher complementizer position, and therefore precedes topics and foci (33a), and a subjunctive/irrealis complementizer (e.g. SI “if” > Ro. să) that lexicalizes the lower complementizer position, and therefore follows topics and foci (33b):

[33] a sunt sigură că pe Ion l-am văzut (Ro.)
    I am sure that on Ion him=I=have seen
    “I’m sure that I saw Ion”

    b vreau MÂINE să meargă la meci (Ro.)
    I want tomorrow that he.go sbjv to match
    “I want him to go to the match TOMORROW”

Such dual complementizer systems are widely found in the dialects of southern Italy, as illustrated in Table 5 (Rohlfs 1969: 190; Ledgeway 2004a; 2005; 2006; 2009b; 2012b):

Table 5. Southern Italian dual complementizer systems

<table>
<thead>
<tr>
<th>It.</th>
<th>penso che verrá “I think that he’ll come”</th>
<th>voglio che lui mangi “I want that he should eat”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sic.</td>
<td>pensu ca véri</td>
<td>vóghiu chi mmanciassi</td>
</tr>
<tr>
<td>Mes.</td>
<td>critu ca véri</td>
<td>óghiu mi mancia</td>
</tr>
<tr>
<td>S.Cal.</td>
<td>pensu ca véri</td>
<td>vóghiu mu (mi) mängia</td>
</tr>
<tr>
<td>N.Cal.</td>
<td>criju ca véri</td>
<td>vóghiu chi mmangia</td>
</tr>
<tr>
<td>Sal.</td>
<td>crisciu ca vêne</td>
<td>agghiu cu mmancia</td>
</tr>
<tr>
<td>Nap.</td>
<td>pensu ca vêna</td>
<td>vóghiu cha mmangà</td>
</tr>
<tr>
<td>N.Pgl.</td>
<td>pensu ca vêna</td>
<td>vóghiu cha mmangà</td>
</tr>
<tr>
<td>Abr.</td>
<td>pensu ca vêna</td>
<td>vóghiu cha mmangà</td>
</tr>
</tbody>
</table>

Further compelling evidence for these two complementizer positions comes from those varieties which allow the

---

simultaneous lexicalization of both positions around a fronted topic or focus constituent (a a case of so-called “recomplementary”), including many early Romance varieties (34a-b; Wanner 1998; Paoli 2003a; Ledgeway 2004a: §4.3.2.2; 2005: 380-389; Vincent 2006), and a number of modern Italian dialects (34c; Paoli 2002; 2003a,b; 2005; D’Alessandro and Ledgeway 2010) and modern Ibero-Romance varieties (34d; Demonte and Fernández-Soriano 2009; Gupton 2010: 227-234; Villa-García 2010).

[34] a je te adjure par le vray Dieu que ta fille
I you=beseech by the true God that your
Tarsienne, que tu ne la donnes a mariage a
Tarsienne that you nother=give to marriage to
autre que a moy (OFr.)
other that to me
“I beseech you before God that you may give your daughter
in marriage to me alone”

b èy manifesta cosa che HOMO CHE SE AVE A
it.is obvious thing that man that self=has to
DEFENDERE A LA PATRIA SOA INTRE LI AMICI E
to.defend to the fatherland his among the friends and
LI CANUSSIENTI SUOY cha ave a chesta parte gran
the acquaintances his that has at this part big
prerogativa e gran avantayo (ONap.)
prerogative and big advantage
“And it is abundantly clear that, a man who has to defend
himself in his own country among his friends and
acquaintances, has in this respect considerable privilege and
advantage”

c so ditte ca dumane, a Urtone, gni lu zie, ca
I.am said that tomorrow at Ortona with the uncle that
nin gi da’ ji (ArL.)
not there= he.must to.go
“I said that tomorrow he shouldn’t go to Ortona with his
uncle”

d Din que o irmán de Iago que sabe xogar moi
they.say that the brother of Iago that knows to play very
ben ao futbolin (Glc.)
well at.the football
“They say that Iago’s brother is very good at football”

Finally, unique within Romance is the situation encountered in Gascon since around the sixteenth century, where the
[+finite] feature of affirmative root clauses is exceptionally spelt out in the systematic lexicalization of the lower complementizer position Fin through the complementizer que "that". Firm proof that que spells out the lower complementizer head is provided by the observation that, apart from object clitics, nothing can intervene between que and the finite verb such that all preverbal lexical subjects must occur to the left of que. This latter observation highlights the fact that, unlike in other Romance varieties, there is apparently no preverbal subject position within the sentential core in Gascon, such that all subjects have to be syntactically fronted to a topicalized (typically if old and definite; 35a) or focalized (typically if new and indefinite; 35b) position within the left periphery. These facts find an immediate explanation in the assumption that que lexicalizes the Fin position, since finiteness is standardly assumed to license nominative Case, hence Gascon would appear to have grammaticalized the locus of finiteness and, by implication, the licensing of nominative subjects within the left periphery (as happens in V2 varieties), rather than in the sentential core.

\[35\] a \quad \{TopP \quad \text{ta pay} \quad \{FinP \quad qu’ \quad \text{ey arribat}]}\] (Gsc.)

your father that is arrived

"your father’s arrived"

b \quad \{FocP \quad quaunque trufandèr\} \quad \{FinP \quad que \quad \{\text{ip vos dirà…}]]\} (Gsc.)

whatever joker that you will say

"any joker will tell you…"

The fine structure of the left periphery can therefore be summarized as in (36):

\[36\] \[\text{force que/che} \quad \{\text{Topic} \quad \text{Foc} \quad \text{FOCUS} \quad \{\text{Fin de/di/che/cu/sà [ip…]}\}\}\]

2.4 Interim conclusions

The discussion so far has highlighted the importance of Romance data, and especially those from the still under-utilized non-standard varieties and dialects, in making us rethink some of our most basic assumptions about language structure and variation. In particular, we have seen that standard interpretations

---

of the null subject parameter prove insufficient in terms of the limited range of null and overt pronominal subject types they predict, the limited predictive power of rich inflection in the distribution of null subjects, and the non-universality of inflection for person in licensing null subjects. Within the verbal domain we have established that the now classical typological distinction between V-raising and non-V-raising languages, which places Romance among the former, proves empirically inadequate. By contrast, we have observed that the extent of V-movement across Romance varies enormously, revealing a whole host of different positions which can, in turn, be taken to spell out the vast range of functional projections that make up the rich architecture of the sentential core. In a similar vein, this rich functional design of the clause has been shown to extend to the left periphery, where the Romance evidence forces us to recognize a richly-articulated functional space composed, at the very least, of topic and focus fields sandwiched in turn, between two complementizer positions.

3. What linguistic theory can do for Romance

3.1 Word order
Undoubtedly one of the most striking differences between modern and medieval Romance varieties is manifested in the, often, radically differing word order patterns they permit in root clauses. By way of illustration, consider the old Neapolitan sentences in (37a-b) and their modern Italian translations in (38-b):

```
[37] a e viechy reduceva ad etate juvenile (ONap.)
   and old she.brought.back to age juvenile
   “and she could make the old young again”
   b de pov queste parole ademandao lo messayo
      after these words asked the messenger
      licencia (ONap.)
      leave
      “And following these words the messenger asked permission to leave”

[38] a e riportava i vecchi a età giovanile (It.)
   and she.brought.back the old to age juvenile
   b dopo queste parole il messaggiere domandò licenza (It.)
      after these words the messenger asked leave
```
(37a-b) exemplify a common early Romance structure in which
the preverbal position is occupied by some constituent other
than the subject, namely the direct object (viechy) and a non-
subcategorized adverbial phrase (de poy queste parole), res-
pectively. In the former case, the fronted rhematic direct object,
which conveys new information, constitutes an example of in-
formational focus (Lambrecht 1994: ch. 5; Cruschina 2012), and
contrasts sharply with many modern Romance varieties such as
Italian (cf. 26-27), where preposing of rhematic constituents is
only found under quite restrictive pragmatic conditions (namely,
to license contrastive focus), insofar as rhematic objects con-
voying informational focus canonically occur in postverbal po-
sition (cf. 38a). Similarly, example (37b) demonstrates how when
the preverbal position is occupied by a constituent distinct from
the subject, the latter, whenever overtly realized, is generally
required to follow the verb, giving rise to an apparent case of
verb-subject inversion. Significantly, in this and similar examples
of inversion the subject does not simply follow the verb but also
precedes any other sentential constituents (subcategorized or
otherwise), witness the order subject (lo messayo) + direct ob-
ject (licencia) in (37b). In modern Italian, by contrast, post-
verbal subjects generally follow their associated objects and
other sentential constituents, and in such cases are typically
associated with rhematic interpretations, whereas the postverb-
also subject in (37b) is clearly thematic. Consequently, in the mo-
dern Italian translation of (37b) illustrated in (38b) the thematic
subject obligatorily occurs in preverbal position.

Word order patterns such as those just considered, which
can be easily replicated for other early Romance varieties (cf.
39a-g), have led a number of linguists to argue that medieval
Romance word order is characterized by a verb second (V2)
constraint. 11 During this V2 stage sentences consist therefore of
two principal parts (39a), a sentential core (IP) with fixed S V O

Ledgeway (2007b; 2008).
Adv order on a par with what we have already witnessed in §2.3 for modern Romance, and a richly-articulated left periphery (CP) along the lines of (36) to whose lowest C(omplementizer) position (Fin) the finite verb raises in root clauses, where it is preceded by one or more elements fronted from the sentential core to the Topic and Focus fields to be assigned a pragmatically salient reading. In embedded clauses, by contrast, the left periphery generally hosts an overt Comp(lementizer) and the finite verb is consequently forced to remain within the sentential core, yielding the order S+V+O+Adv (39b). Thus, as the following representative early Romance examples demonstrate, alongside S+V+X (39c) we also frequently find in main clauses O+V(S) (39d), O+V(S) (39e), O+V(S) (39f), and Adv+V(S) (39g), whereas embedded clauses invariably display rigid S+V+O+Adv (39h):

[39] a [\text{Left Periphery/Top} / \text{Foc} V \quad [\text{Sentential Core} S \ X O \ Adv]]

b … [\text{Left Periphery/Comp} \quad [\text{Sentential Core} S \ V \ O \ Adv]]

c [CP Lo \text{ cavaliere} \text{ presse} \quad [IP \text{ lo cavaliere presse} \quad \text{the knight took the march}]\text{]} \text{ (OTsc.) marks}”

d [CP \text{ Grande duel} avien \quad [IP \text{ las yentes cristi}nas \text{ avien great sorrow had the peoples Christian grande duel}] \text{]} \text{ (OSp.) “The Christian peoples felt great sorrow”}

e [CP A \text{ ceste paroles} \text{ respont} \quad [IP \text{ la reine respont to these words replies the queen a ceste paroles}] \text{]} \text{ (OFr.) “The queen replied to these words”}

f [CP D’ \text{ ALGUÑAS COUSAS me calarei} \quad [IP \text{ me calarei of some things myself=I.will.fall.silent d’alguñas cousas}] \text{]} \text{ (OPl.) “I shall remain silent about certain matters”}

g [CP \text{ Luen} \text{ hes} \quad [IP \text{ lo castelli} es la tors es luen] \text{]} \text{ (OOcc.) “Far is the castle and the tower”}

h [CP q’ [IP \text{ ella fetz so q’ el volc}] \text{ (OOcc.) “And he honoured her such that she did what he wanted”]

In conclusion, we thus see that the fine structure of the sentential core (§2.3) and the left periphery (§2.3.1) independently
established above on the basis of modern Romance data provide us with the necessary pragmatico-syntactic tools to interpret the facts of medieval Romance word order.

3.2 Pro-drop parameter revisited

Returning to our observation regarding the distribution of null subjects in medieval Romance (cf. examples (15)-(16) in §2.1), this same generalized raising of the finite verb to C(omp) under V2 also provides us with an elegant and highly natural explanation for the asymmetric distribution of null subjects in early Romance. In particular, when the verb raises to the vacant C position, null subjects are freely licensed (40a), whereas in subordinate clauses, where the finite verb is forced to remain \textit{in situ} within the sentential core, pronominal subjects must be phonologically expressed (40b), although not interpreted as emphatic or contrastively-focused.

\begin{verbatim}
[40] a  Si \(\emptyset\) errent tant en tele maniere qu'[…] (OFr.)
        so (they) wander so much in such way that
        “they thus wandered so much in such a way that […]”

b  Endementiers que il regadoient les letres et la
     while that they looked at the letters and the
     damoisele, que il plagnoient sa mescheance (OFr.)
     young lady that they regretted her misfortune
     “While [that] they looked at the letters and the young lady,
      [and that] they regretted her misfortune”
\end{verbatim}

This asymmetrical distribution leads us to conclude that null subjects in medieval Romance were not licensed exclusively, if at all, by rich verb inflection for person and number, but, by a property which the finite verb uniquely acquires by raising to the vacant C position, presumably the locus of finiteness in medieval Romance (cf. Ledgeway 2012a: §3.4.1).

Interestingly, this archaic pattern has been exceptionally retained in modern Corsican (Marchetti 1974: 25, 51, 85, 94, 119), which behaves as a canonical Romance null subject variety in root clauses (41a-b), but which in embedded contexts usually requires referential (41a-c) and non-referential (41d) pronominal subjects to be overtly realized, albeit in reduced clitic form (cf. tonic/clitic \textit{eio/e, o} (1sg), \textit{tiu/tu} (2sg), \textit{ellu/ella / (e)llu/(e)lla} (3m/fsg), \textit{noi/no} (1pl), \textit{voi/vo} (2pl), \textit{elli/elle} /
This distribution of null and overt pronouns is however only superficially similar to medieval Romance, in that modern Corsican is not a V2 variety and the finite verb occurs in the sentential core in all cases, irrespective of the realization of the pronominal subject. However, when the finite verb does raise to the vacant C(omplementizer) position, as happens in polar interrogatives (41e), realization of the clitic subject pronoun is blocked in a similar fashion to what was seen for medieval Romance (cf. 40a).

[41]  

- **a**  Ø avemu da fà e divuzioni à tutti i (we) have from to do the devotions to all the saints that we= find  
  “we’ll pray to all the saints we meet”

- **b**  Ø e cumpretera à u scagnu quand’è vo (you) them= you will buy a the office when you= fàlate (Crs.)
  “you buy them at the office when you go down”

- **c**  eo ogni volta ch’o tornu in Cervioni, è ch’o l each time that I= return in Cervione and that I= u sentu u nostru chjoccu mi mette sottusopra (Crs.) it= hear the our chiming me= it puts under over
  “each time I go back to Cervione and hear our chiming bell I get overcome”

- **d**  Ch’ellu un ci sia troppu rumore (Crs.)
  “Make sure that there isn’t too much noise”

- **e**  [cP QUANDU fàlate [IP Ø fàlate quandu]] ? (Crs.)
  “When are you coming down?”

### 3.3 Placiti cassinesi

We now turn to the area of Romance philology and, in particular, the *Placiti cassinesi* (Bartoli 1944-45; Castellani 1973: 59-76; Bianchi, De Blasi and Librandi 1993: 211-212). Ignoring the *Indovinello veronese*, the Latin features of which are hardly negligible, these four brief, formulaic, sworn declarations composed in Capua, Sessa Aurunca and Teano (all situated in the modern-day southern Italian region of Campania) and dating between 960 and 963 are generally taken to represent the first documented attestations of the vernacular within the Italian
Peninsula. Below is reproduced the first of these, the Placito capuano from March 960, which relates to a land dispute between the abbot of Montecassino and a certain Rodelgrimo who claimed, through inheritance, ownership of the lands which the abbot maintained had been the property of the monastery of St. Benedict of Montecassino for thirty years. In the absence of any official documentation of ownership, the judge ordered the abbot to produce three witnesses to authenticate his claim, each of which was reported to have sworn the oath in (42):

[42]  Sao ko kelle terre, per kelle fini que ki contene,  
I know that those lands for those confines that here contains 
trenta anni le possette parte sancti Benedicti 
“three years there=possessed party saint.obl Benedict.obl 
“I know that, those lands, within those borders which are 
contained here [in the document/map before me], have 
belonged for thirty years to the part [= monastery] of 
St. Benedict [of Montecassino]”

Although a very short text, the Placito capuano is not without its problems, raising a number of qualitative and interpretive issues for the philologist. In particular, given the extremely brief and formulaic nature of the text, it is legitimate to question what is the value, if any, of the linguistic evidence that such a short piece can genuinely offer the historian of the language. Indeed, this is a problem which arises with many of the earliest attestations of the Romance vernacular including, for example, the earliest Romance text, the Strasbourg Oaths of 842, another short sworn oath produced in an early French dialect of disputed origin.\textsuperscript{12}

Another unresolved issue thrown up by the Placito capuano concerns the correct reconstruction of the pragmatico-semantic interpretation of the fronted constituents stacked up at the beginning of the embedded clause (namely, kelle terre, per kelle fini..., treinta anni), our reading of which is without doubt greatly hindered by the limited nature of our textual evidence. Again this is a frequent problem faced by philologists reading early texts, which in many cases only offer a rather brief glimpse of the language, especially when they only exist in fragmentary

form, and whose linguistic physiognomy is often deliberately limited by the specific style and register of the text type.

A final observation concerns the appearance of the complementizer *ko* “that” in (42). Contrary to our expectations in light of the discussion of dual complementizer systems above (cf. Table 5 in §2.3.1), the epistemic main clause predicate *sao* “I know” selects for an indicative clausal complement headed, not by the expected indicative complementizer *ca*, but by what appears to be the subjunctive complementizer *ko*. How are we then to interpret the appearance of *ko* in this instance? Is it an example of a scribal error, or should it be taken at face value? As we shall see below, a consideration of this question in light of our preceding discussion of the fine structure of the left periphery provides an illuminating solution.

Putting together the results of the discussions of the previous sections regarding the fine structure of the left periphery of the clause (§2.3) and the philological evidence of the *Placito capuano*, it becomes clear that a knowledge of the relevant linguistic and philological facts can profitably complement one another. In particular, we begin by observing how, despite only providing a glimpse of the early vernacular, the *Placito capuano* is of immense interest both to the historian of the language and the historical linguist since, although a very short text, it is nonetheless astonishingly rich in linguistic evidence. More specifically, it is quite remarkable that such a short text, and the first one from the Italian Peninsula no less, should provide such extensive early evidence of the fine structure of the left periphery and, above all, incontrovertible proof for the existence of the two left-peripheral fields postulated in §2.1, which, in turn, hold the key to a proper pragmatico-semantic interpretation of the fronted constituents stacked up at the beginning of the embedded clause. In particular, the rigid ordering of the *Topic + Focus* fields postulated in §2.1, together with our observations regarding the potential recursiveness of topics but not focus, which is limited to a single occurrence, allows us to infer that the left periphery of the embedded clause hosts two thematicizations within the Topic field, namely *kelle terre* “those lands” (picked up by the resumptive clitic pronoun *le*
“them”) and per kelle fini que ki contene “with those borders which are contained here”, and a contrastive focus treinta anni “(for) thirty years” within the Focus field, as illustrated in (43).

[43]  Sao [Force ko [Top kelle terre, per kelle fini que I.know that those lands for those confines that ki contene, [Foc treinta anni [Ip le possette parte here contains thirty years them= possessed part sancti Benedicti ]]]

saint.obl Benedict.obl

Turning now to the unexpected used of the complementizer ko (< QUOD) rather than ca, this too finds a principled explanation in terms of the structural organization of the complement clause. As demonstrated in Ledgeway (2004a; 2005), in the early dialects of southern Italy, including those of Campania, the distribution of the two complementizers co/che (< QUOD/ QUID) and ca (< QU(1)A) is not quite as neat as the traditional descriptions reviewed in §2.3.1 would lead us to expect (cf. Table 5). Whereas all types of subjunctive clause are introduced by che, indicative complement clauses are headed either by ca or che. Simplifying the facts somewhat (for detailed discussion, see Ledgeway 2005: §3), it will suffice to note here that either ca (44a) or che (44b) are employed when the left periphery of an embedded indicative clause does not contain any topics or foci, whereas che alone is found in the presence of fronted topics/foci (44c), as witnessed by the following old Neapolitan examples:

[44] a  Homero […] dice a li suoy libri ca [Ip foro nave Homer says to the his books that were ships MCLXXXVI] (ONap.)
MCLXXXVI “Homer […] says in his books that there were 1186 ships”

b  purriase ben dicere che [Ip fo causa mutto could=self well to.say that it.was cause very legiere] (ONap.)
light “it could indeed be said that there was very little reason for it”

c  Considerava che [Top a quista insula de Colcosa […], [Ip he.considered that to this island of Colchis non se nce poteva gire se non permare] (ONap.)
not one=there=could to.go if not by sea
“He thought that, this island of Colchis […], was only accessible by sea”
In view of these distributional facts, it is possible to argue that old Neapolitan had just one indicative complementizer *ca* generated in the lowest complementizer position (*Fin*) which, whenever raised to the higher complementizer position (*Force*), as proves obligatory whenever topics or foci are present, is morphologically spelt out in the form *che* (namely, \(ca_{Fin} \Rightarrow che_{Force}\)). In short, the overt form assumed by indicative *co/che* is interpreted as nothing more than the surface morphological reflex of raising *ca* from its base position to the higher complementizer position within the left periphery. Indeed, this analysis is directly supported by old Neapolitan recombination examples such as (34b) above where, crucially, the higher complementizer invariably surfaces in the morphological form *co/che* but the lower position is always spelt out as *c(h)a*, and never *vice versa*.

In this light, we can now return to the apparently erroneous selection of *ko* in the *Placito* in (42). It turns out after all that the use of *ko* here is not a scribal error or a singular mistake, but reflects the availability and use of two different complementizer positions determined by the informational structure of the embedded clause (cf. 45a) which, as we have seen in (43), contains a contrastively focused constituent preceded by two topicalized constituents. It is the activation of these *Topic-Focus* fields in the embedded left periphery which is directly responsible for the presence of *ko*, inasmuch as the complementizer *ca* is forced to move to the higher complementizer position in the presence of fronted topics or foci where it is spelt out as *ko* (45b). It now comes as no surprise therefore that the complementizer *ca* (or its graphic variants *ka, cha*) should not be employed in (42).

\begin{enumerate}
\item[a] \ldots \quad [CP \ ko/c(h)a \ Topic \ + \ FOCUS \ ka/c(h)a \ [ip\ldots]]
\item[b] \quad Sao \quad [CP \ ko, kelle\ldots, per kelle\ldots \ TRENTA ANNI \ ka, [ip\ldots]]
\end{enumerate}

This example also highlights the dangers of dismissing too hastily the linguistic attestations offered to us by even the most meagre of philological evidence. In this particular case, the *Placito capuano* provides an invaluable early example of the *ca \Rightarrow che* alternation, which is not attested again in our textual record for the Campania region for at least another 350 years. Thanks however to this early attestation, we can conclude with
Testing linguistic theory and variation to their limits

certainty that the relevant complementizer alternation and associated positions licensed by the absence/presence of fronted topics and foci, otherwise richly attested in texts from the fourteenth century onwards (Ledgeway 2004a; 2005; Vincent 2006), must date back to at least the tenth century. Of course, it might be objected that the linguistic evidence of the texts is so formulaic that its value for the linguist is questionable on a number of counts. Indeed, a comparison of the Placito capuano with the other three Placiti cassinesi produced three years later in the nearby localities of Sessa Aurunca (46a) and Teano (46b-c) reveals such an extraordinarily high level of structural, discourse and lexical uniformity across all four texts that it would be naïve to imagine that all four sworn oaths represent authentic testimonies of the spontaneous spoken vernacular of the time.

[46] a Sao cco kelle terre, per kelle fini que tebe
I know that those lands for those confines that to you I showed Pergoaldi foro, que ki contene, et
I showed Pergoaldi. obl were that here contains and
tranta anni le possette
thirty years them= possessed
“I know that, those lands, within those borders which I have shown you, and which are contained here [in the document/map before me], belonged to Pergoaldo [the abbot of the monastery of San Salvatore] and that he owned them for thirty years”

b Kella terra, per kelle fini que bobe mostrai,
that land for those confines that to you I showed
sancte Marie ò, et trenta anni la possette
saint. obl Maria. obl is and thirty years it= possessed
parte sancte Marie
part saint. obl Maria
“That land, within those borders which I have shown you [in the document/map before me], belong to [the monastery of] Santa Maria [di Cengla], and were possessed by the part [= monastery] of Santa Maria [di Cengla] for thirty years”

c Sao cco kelle terre, per kelle fini que tebe
I know that those lands for those confines that to you I showed, trenta anni le possette parte sancte
I showed thirty years them= possessed part saint. obl
Maria
Maria. obl
“I know that those lands, within those borders which I have shown you [in the document/map before me], have belonged for thirty years to the part [= monastery] of Santa Maria [of Cengla].”

On the contrary, given the nature of the four oaths, which all had the specific purpose of persuading the court that a set of disputed lands had been in the possession of a given monastery for thirty years and consequently a legitimate part of the latter’s estate, it is more than likely that the illiterate peasants enlisted by the Church to serve as “independent” witnesses, presumably under the promise of personal financial reward, were given very precise instructions regarding what they were required to swear under oath. Nonetheless, it would be rash to disregard the evidence of these four short, highly formulaic written testimonies on these grounds alone; rather, given the Church’s deliberate efforts to place presumably authentic-sounding words and structures of the vernacular in the mouths of their witnesses, it is still possible to see in the language of these four texts a deliberate hypercharacterization of some of the most salient traits of the spoken language of the time such as the fronting of topicalized and focused constituents, which still constitutes to this day a characteristic feature of the spoken, rather than written, registers of Romance (cf. Duranti and Ochs 1979).

The discussion in this section has demonstrated that, when theory and philological evidence are considered together, the results of traditional philological and linguistic scholarship can be considerably enhanced. More specifically, we have seen that current theoretical assumptions about the design and architecture of the left periphery of the clause provide us with some novel and powerful tools to shed light on the interpretation and linguistic choices of one of the earliest Romance texts. At the same time, the Placiti cassinesi provide independent and robust evidence for these same syntactic assumptions. In short, and as abundantly demonstrated in this case study, linguistics and philology should complement each other to produce enlightening results, rather than be seen as alternatives to be pursued in isolation from each other.
Dual complementizer systems

Above in §2.3.1 we observed how Romanian and a number of southern Italian dialects employ a dual complementizer system (cf. Table 5) which distinguishes between a realis complementizer derived from *qua* (*ca, că*) and an irrealis complementizer derived from *quid* (*che, chī*, *quod* (*cu*), *modo* (*mu, ma, mi*) or *si* (*sā*): while the former heads clauses selected by declarative and epistemic predicates (typically marked by the indicative; 47a), the latter is employed after predicates such as volitionals that characterize the state or events of their complements as unrealized at the time of speaking (typically marked by the subjunctive; 47b).

We also saw that both complementizers differ with respect to their relative positions in conjunction with fronted topicalized and focused constituents. For instance, as the preceding Salentino examples demonstrate, while *ca* precedes all such fronted elements (47a), *cu* must follow (47b). Facts like these led us to propose in terms of Rizzi’s (1997) split CP perspective that the realis complementizer (e.g. *ca*) lexicalizes the higher Force head and the irrealis complementizer (e.g. *cu*) the lower Fin head, as sketched in (48):

We also saw that both complementizers differ with respect to their relative positions in conjunction with fronted topicalized and focused constituents. For instance, as the preceding Salentino examples demonstrate, while *ca* precedes all such fronted elements (47a), *cu* must follow (47b). Facts like these led us to propose in terms of Rizzi’s (1997) split CP perspective that the realis complementizer (e.g. *ca*) lexicalizes the higher Force head and the irrealis complementizer (e.g. *cu*) the lower Fin head, as sketched in (48):

A further distinction between the two complementizers which, at first sight, does not seem to immediately follow from (48) concerns their phonological realization (Calabrese 1993; Terzi 1994; 1996; Ledgeway in press). For example, while in Salentino *ca* must always be pronounced (49a), *cu* may optionally remain unpronounced (49b), a case of what we may term C-drop.

A further distinction between the two complementizers which, at first sight, does not seem to immediately follow from (48) concerns their phonological realization (Calabrese 1993; Terzi 1994; 1996; Ledgeway in press). For example, while in Salentino *ca* must always be pronounced (49a), *cu* may optionally remain unpronounced (49b), a case of what we may term C-drop.
A. LEDGEWAY

b ogghiù (cu) llu faci stare cittù (Sal.)
I want that it= you.make to.be quiet
“I want you to shut him up”

Strikingly, even in those Romance varieties which do not display a dual complementizer system it is not infrequent for the finite complementizer (viz. que/che) to remain phonologically unpronounced, on condition that it introduces an irrealis complement (50b, 51b, 52), but not, crucially, when it introduces a realis complement (50a, 51a).13

[50] a Digo *(que) puede durar mucho (Sp.)
I.say that it.can.ind to.last much
“I say that it can last a long time”
b Temo (que) pueda durar mucho (Sp.)
I.fear that it.can.sbjv to.last much
“I fear that it may last a long time”

[51] a Dice *(che) vuole piovere (It.)
he.says that it.wants.ind to.rain
“He says that it is going to rain”
b Sembra (che) voglia piovere (It.)
it.seems that it.wants.sbjv to.rain
“It seems to want to rain”

[52] Non vuel (Ø) dormas plus (OOcc.)
not I.want (that) you.sleep.sbjv more
“I don’t want you to sleep any longer”

In view of the superficial similarity in the conditions regulating the distribution of C-drop in varieties with dual complementizer systems and languages such as Spanish, Italian and Occitan with apparently only one complementizer, it is tempting to extend the analysis of dual complementizer systems to languages of the latter type. In particular, we have established on the basis of dual complementizer systems that C-drop is a property uniquely licensed by the lowest C-related position, presumably involving a residual V2 effect with raising of the modally-marked finite verb to \(C_{Fin}\) in the absence of the irrealis complementizer (Poletto 2001), but not by the realis complementizer which lexicalizes the highest C-related position Force and hence unavailable to

finite verb raising under V2. By the same token, in its irrealis uses standard Romance *que/*che must also lexicalize *Fin*, thereby explaining its complementary distribution with finite verb raising under C-drop, in contrast to its realis uses where it lexicalizes the highest position *Force*.

In short, we are led to conclude that the dual complementizer system explicitly attested for such varieties as Romanian and the dialects of southern Italy must also be assumed to hold more generally for Romance where, despite the difference between the two complementizers not being overtly lexicalized, the relevant distinction between the two homophonous complementizers (*que/*che) is marked indirectly by their differential positions within the left periphery and their compatibility with C-drop (53a-c).

\[
\begin{align*}
&[\text{Force } \emptyset] & & [\text{Topic, Focus } [\text{Fin } \text{que/che-sbjv } [\text{ip } \text{v-sbjv}]]] \\
&[\text{Force } \emptyset] & & [\text{Topic, Focus } [\text{Fin } \text{v-sbjv } [\text{ip } \text{v-sbjv}]]] \\
&[\text{Force } \text{que/che-ind}] & & [\text{Topic, Focus } [\text{Fin } \emptyset [\text{ip } \text{v-ind}]]]
\end{align*}
\]

3.5 Noun positions

Oversimplifying somewhat, prenominal and postnominal adjectives positions typically correlate with the following respective interpretations in the modern standard varieties of Romance:14 (i) inherent / non-inherent (54a); (ii) descriptive / distinguishing (54b); (iii) subjective / objective (54c); and (iv) figurative / literal (54d).

\[
\begin{align*}
&\text{su británica reserva vs la Embajada británica (Sp.)} \\
&\text{his British reserve the Embassy British} \\
&\text{“his British reserve” vs “the British Embassy”} \\
&\text{une courte lettre vs une jupe courte (Fr.)} \\
&\text{a short letter a skirt short} \\
&\text{“a short letter” vs “a skirt that is short”} \\
&\text{un (formidable) om (formidable) (Ro.)} \\
&\text{a tremendous person tremendous} \\
&\text{“a tremendous person”}
\end{align*}
\]

A. Ledgeway

d la grisa quotidianitat vs la camisa grisa (Cat.)
the grey daily routine the shirt grey
“the grey (= dull) daily routine” vs “the grey shirt”

Matters were, however, quite different in early Romance, insofar as contrastive readings were not necessarily associated with the postnominal position as in modern Romance, but could equally be licensed in the prenominal position (Ledgeway 2007a; 2009a: 214f; Vincent 2007; Thiella 2008: ch. 4; Giusti 2010: 599-609).15 This is illustrated in the following near minimal pairs, where the adjective in each case invariably gives rise to a contrastive reading irrespective of its pre- or postnominal position.

[55] a pro christian poblò vs lo nom christiën (OFr.)
for Christian people the name Christian
“for the Christian people” vs “the Christian name”

b con tondo giro vs uno tagliere tondo (OTsc.)
with round circle a chopping board round
“with a round circle” vs “a round chopping board”

c li spangnoli soldati vs le compagnie spanghols (ONap.)
the Spanish soldiers the companies Spanish
“the Spanish soldiers” vs “the Spanish troops”

Strikingly similar is the situation found in modern Walloon (Bernstein 1991), where all adjectival classes (apart from those expressing nationality) occur in prenominal position (56a-c), perhaps representing a conservative feature, although one cannot also exclude the (reinforcing) influence of neighbouring Flemish varieties.

[56] a dès r’châfés crompies (Wal.)
“some reheated potatoes”

b du l’ corante ëwe (Wal.)
of the running water
“some running water”

c lès cuits pans (Wal.)
“the cooked loaves”

By contrast, in a number of, especially non-standard, Romance varieties including Occitan (57a; cf. Wheeler 1988: 268), Sardinian (57b; cf. Jones 1993: 42) and central-southern Italian dialects

Testing linguistic theory and variation to their limits

(57c; cf. Rohlfs 1969: 330; Ledgeway 2007a), the prenominal adjectival position is extremely restricted and generally replaced by the postnominal position, which is neutral to the contrastive vs non-contrastive distinction:16

a lo (*viélh) pönt viélh d’Avinhon (Lgd.)
the old bridge old of Avignon
“the old Avignon bridge”
b na (*piccerella) patana piccerella (Nap.)
a small potato small
“a small potato”
c una (*nova) mákkina nova (Srd.)
a new car new
“a new car”

Building on our discussion and analysis in §2.2 of variable verb positions around different adverbial classes distributed across the clause, it is possible to make sense of the variation witnessed in (54)-(57) along parallel lines in terms of the varying extent of N(oun)-movement in different Romance varieties. In particular, distinct adjectival positions can be reinterpreted as the surface reflex of the head noun variously moving across a rigidly-ordered series of different adjectival classes situated immediately above the NP, as shown in the structural representation in (58):

[58]  (N) [AP$_1$ Adj (N) [AP$_2$ Adj (N) [AP$_3$ Adj (N) [AP$_4$ Adj [NP]]]]]

Assuming the much-simplified structure in (59a) in which AP$_1$ and AP$_2$ can be broadly understood as the “areas” in which non-contrastive and contrastive adjectives, respectively, are generated, we can now formally capture in a highly simple manner the differences between the non-standard varieties in (57a-c) on the one hand and standard Romance varieties (54a-d) on the other: in the former the nominal head typically raises to the highest available position above the highest adjectival projection (AP$_1$), which hosts non-contrastive adjectives from where it precedes both non-contrastive and contrastive adjectives (59b), whereas in the latter the noun only targets the higher adjectival

---

16 Cf. the postnominal position of the adjectives in (57a-c) with the corresponding prenominal position in their French (le vieux pont d’Avignon) and Italian (una piccola patata; una nuova macchina) translations.
projection (AP'), from where it precedes non-contrastive adjectives but follows those with a contrastive reading (59c):

\[[59] \begin{align*}
  a & \quad \ldots \quad \text{(N)} [\text{AP}_1 \text{Adj}_1 \text{(N)}] [\text{AP}_2 \text{Adj}_2 [\text{NP \ N}]] \\
  b & \quad \ldots \text{lo pont} [\text{AP}_1 \text{vièlh pont}] [\text{AP}_2 [\text{NP pont}]] \text{ (Lgd.)} \\
  c & \quad \ldots \text{le vieux pont} [\text{AP}_1 \text{vieux pont} [\text{AP}_2 [\text{NP pont}]] \text{ (Fr.)}
\end{align*}\]

By the same line of reasoning, we can explain the frequent pre-nominal position of contrastive adjectives in early Romance (cf. 55a-c) by assuming that N-raising is only optional in the early varieties (60a-b), an archaic pattern still preserved to the present-day in Walloon where the nominal head barely moves at all (60c):

\[[60] \begin{align*}
  a & \quad \text{li} [\text{AP}_1] [\text{AP}_2 \text{spagnoli [NP soldati]]}] \text{ (ONap.)} \\
  b & \quad \text{le} [\text{AP}_1 \text{compagnie [AP}_2 \text{spangnole [NP comp.]}}] \text{ (ONap.)} \\
  c & \quad \text{dès} [\text{AP}_1 \text{r'tchâfés [AP}_2 \text{crompîres]}] \text{ (Wal.)}
\end{align*}\]

In conclusion, we see that there are truly striking parallels between the nominal and verbal domains and the functional structures associated with these, as revealed by the fixed positions of distinct classes of adjectival and adverbial modifiers, respectively. In particular, we have seen how different diachronic and diatopic varieties of Romance provide clear evidence for some considerable microvariation in terms of the extent of N-movement. Among the various patterns observed, we also identified a strong diachronic tendency for Romance nouns to climb progressively higher within the available nominal functional structure, the end result of which has given rise in a number of non-standard varieties to a semantic neutralization of the erstwhile interpretive difference between contrastive and non-contrastive adjectival readings.

**Conclusion**

Although modern linguistics as a discipline developed in large part out of the philological study of individual languages and language families such as Romance, it is a striking characteristic of contemporary research in both Linguistics and in Romance Studies that the traditional link between the two disciplines is
often not as strong and prominent as it might be. To be sure, analyses within Romance linguistics which fail to take account of the most enlightening ideas and principles of linguistic theory risk overlooking and/or misconstruing the relevance of all or part of the available empirical evidence they are so at pains to correctly reconstruct, evaluate and interpret. By the same token, linguistic analyses which are blindly driven by theory-internal considerations with little or no real interest in actual data such as those offered by textual corpora and the numerous dialectal varieties of Romance run the risk of presenting a largely idealized and, by definition, necessarily selective representation of the available linguistic evidence. The result is a partial theory which is only capable of accounting for a subset of the available data that largely ignores the imperfections and irregularities characteristic of authentic linguistic productions. The discussion of the preceding sections, by contrast, has demonstrated that when theory and Romance evidence are considered together, the results of traditional Romance and linguistic scholarship can be considerably enhanced.

Bibliographical references


Bec P. (1967). La langue occitane. Montpellier: PUF.


Testing linguistic theory and variation to their limits


Testing linguistic theory and variation to their limits

_Linguistics. Papers in Honor of Mario Saltarelli._ Amsterdam/Philadelphia: Benjamins, 45-76.


A. LEDGEWAY


Testing linguistic theory and variation to their limits


A. LEDGEWAY


Testing linguistic theory and variation to their limits


Testing linguistic theory and variation to their limits
