

CHAPTER 5

Romance

A typological approach

PAOLO RAMAT AND DAVIDE RICCA

5.1 Typologies of Romance languages

Nowadays there exist many proposals for classifying languages from a typological point of view. In principle all the proposed typological approaches are acceptable. In what follows we shall base our description on the viewpoints of the major contemporary proposals. These in fact complement each other, since they refer to different aspects of language. Thus, word order typology is not in contrast with the actancy-based approach nor head/dependent-marking typology. Since a 'holistic typology', capable of explaining all the facts a language exhibits via a unique principle, does not yet exist, we will choose from time to time the approaches that are the most fitting to describe the various phenomena at issue.

One typological taxonomy is the 'Actancy Typology': Lazard (1997) has sketched a typology according to the behaviour of the semantic roles of Agent and Patient in relation to the verb. He assumes as a basic semantic structure that which contains a Predicate, an Agent, and a Patient (or Undergoer, Experiencer) such as (1) which he calls 'phrase d'action' ('process'), 'construction biactancielle majeure' ('major biactancy construction'). The grammatical relations between the predicate and its arguments may be expressed by their order in the sentence, by special markings on the predicate or on the arguments or even on the predicate and (parts of) its arguments.

- (1) Dominus seruum uerberat. (Lat.)
 master.NOM servant.ACC whip.IND.PRS.3SG
 'The master whips the servant.'

Lazard makes use of very neutral symbols such as X, Y, and V(erb) and provides these symbols with indices which express their mutual relations. Thus the abstract structure $X_0Y_0V_{xy}$ symbolizes a structure where the Agent and Patient have no marking, but the verb is marked both for the Agent and the Patient; $X_iY_0V_{xy}$ will, in turn, refer to a structure where the Agent has marking but the Patient does not,

while the verb continues to reference both Agent and Patient.

It is thus possible to characterize different types of constructions via this symbolism. Example (1) will be represented by

$X_nY_aV_x$ ip3Sg [n = NOM, a = ACC, ip3Sg = Present Indicative, 3rd person singular]

All members of the sentence have morphological marking, and there is no morphological difference between animate and inanimate participants. Compare:

- (2) Non omnis arbusta iuuant
 not all.ACC.MPL bushes.NOM please.IND.PRS.3PL
 humilesque myricae (Lat., Virg. Ecl. IV, 2)
 humble.NOM.FPL=and tamarisks.NOM
 'not everyone likes bushes and humble tamarisks'

The intransitive sentence (3) will be X_nV_x ip3Sg, whereby the Agent of the one-actant verb behaves like that of the two-actant sentence (1).

- (3) Dominus uenit. (Lat.)
 master.NOM come.IND.PRS.3SG
 'The master is coming.'

This kind of notation enables us to classify different linguistic alignments and constructions (the active, the ergative, the antipassive, the inactive, etc.). What makes the cross-linguistic comparison possible is the invariable semantic content and the functional relations obtaining between the predicate and its two basic arguments, namely Agent and Patient.

Romance languages have predominantly constructions of the types $X_0Y_0V_x$ ip3Sg and X_0V_x ip3Sg:

- (4) Le maître bat le serviteur. (Fr.)
 the master beat.IND.PRS.3SG the servant
 'The master beats the servant.'

- (5) Le maître arrive. (Fr.)
 the master come.IND.PRS.3SG
 'The master is coming.'

The unmarked word order is subject+verb+object (SVO) in declarative sentences (cf. also §§31.1-2, 34.3.1, 62.2). Latin, on the contrary, had SOV as basic word order, though other linearizations were frequently used according to focalizing strategies. Example (2), for instance, has the subject *humilesque myricae* in focus position at the end of the sentence.

The word order typology inspired by Greenberg (1966) unifies under a general principle many linguistic phenomena such as the relation between verb and object, noun and adjective, noun and genitive, noun and relative clause, standard and comparative. The unifying principle is the relation between Determinee (Det^{ee}) and Determiner (Det^{er}): the OV order is Det^{er}+Det^{ee}, the VO order is the reverse. The transition from Latin to Romance sees a drift from OV, which was prevalent (but not compulsory) in Classical Latin, to VO. However, Classical Latin already had some VO features (e.g. it had mostly prepositions instead of postpositions, and the order noun-genitive was already prevailing over the more archaic genitive-noun order, cf. Ledgeway 2012a:213).

The typology proposed by Nichols (1986) is based on the 'head-marking' vs 'dependent-marking' dichotomy. It is still a debated question which is the head of a noun phrase. According to 'categorial grammar', the constituent whose category coincides with the category of the entire phrase is the Determinee ('operand'), and the Determiner ('operator') is the adjoined, specifying element (see Bartsch and Vennemann 1982:39). In the adpositional phrase (6) the head is the preposition *ab* and the dependent NP *urbe condita* marks its relation with the head via a special case, namely the ablative. However, in a language in which prepositions do not govern case (like most Romance languages),¹ the PPs are neither head- nor dependent-marked, i.e. they are neutral with respect to the distinction.

- (6) ab urbe condita (Lat.)
 [[from]_{PREP} [[city.ABL]_N [founded.ABL]_{PTCP}]_{NP}]_{PP}
 'from the foundation of the city [Rome]'

In a noun phrase expressing the possession relationship, such as (7), *domus* (and *house*) are the determined 'heads' ('operands') and *hominis* (and *the man's*) are the 'dependent'

determiners ('operators'). The relation between the two elements is specified on the dependent one (*hominis/the man's*).

- (7) hominis domus (Lat.)
 man.GEN house.NOM
 'the man's house'

The Romance languages, where genitival prepositions are linked (prosodically and syntactically) with the possessor, still belong to the dependent-marking type: cf. Fr. *la maison [de/(à) Pierre]_{PP}* 'the house of/(to) Pierre'.²

Gender agreement in adjectives is another good instance of dependent marking both in Latin and Romance languages:

- (8) a. altus mons bona sors (Lat.)
 high.NOM.MSG mountain.NOM.(M) good.NOM.FSG fortune.NOM.(F)
 'high mountain' 'good luck'
 b. monte alto fonte fresca (It.)
 mountain.NOM.(M) high.NOM.MSG spring.NOM.(F) fresh.NOM.FSG
 'high mountain' 'fresh spring'

In (8), nouns with the same phonetic structure (*mons-sors* and *monte-fonte*) have no sign of their gender, and only the dependent adjectives tell us that *mons* and *monte* are masculine while *sors* and *fonte* are feminine. On the contrary, subject agreement inflection on the verb is a clear head-marking feature both in Latin and in Romance. This is not surprising, since this is probably the most widespread head-marking feature cross-linguistically (Nichols 1986:77). On the whole, Latin and Romance approach more closely the dependent- than the head-marking type, but Romance languages much less so than Latin, as will be seen.

Dependent-marking languages seem to disprefer VO order (cf. Nichols 1986:79). However, neither Latin nor—much less so—Romance fully comply with this tendency (see §5.3.3).

In the verb system tense, mood, aspect, and person morphemes, as usual in fusional languages, follow the base (cf. Lat. *CANTAUERUNT*, Fr. *chantèrent*, Sp. *cantaron*, etc., 'sing.IND.PST.PFV.3PL') and determine the verbal nature of the word while the base is not necessarily verbal (*CANTUS*, Fr. *chant*, Sp. *canto* mean 'song'). At any rate, in the evolution from Latin to Romance we note 'a progressive reversal of the directionality

¹ Exceptions are Romanian and medieval Gallo-Romance, which keep bicasual systems in nouns, adjectives, and determiners. More instances of case-governing prepositions are of course found with personal pronouns, with up to four-case systems still attested in some varieties of southern Italy and Sardinia (Loporcaro 2008; 2009).

² There are some cases of prepositionless possessors in Romance (cf. §§16.3.1.1, 56.2.14), in modern French remnants like *Hôtel-Dieu*, lit. 'house God', or in southern Italian dialects, e.g. Verbecchese [a 'kasa u provv'ssurə], lit. 'the house the professor' (Silvestri 2012:569). These obviously qualify as neutral with respect to the head/dependent typology.

parameter from a regular head-final setting towards a head-initial setting' (Ledgeway 2012a:236), and the synthetic form (*ils chantèrent* '(they) sing.IND.PST.PFV.3PL' has been progressively replaced by the analytic (*ils ont chanté* (lit. '(they) have.3PL sung'), where the grammatical information (*ont* 'they have') has been fronted.

Coseriu (1987; 1988) noticed the difference between nominal and verbal systems in the Romance languages. Verbs tendentially have an internal, synthetic determination: Fr. (*nous*) *chantons/chanterons/chanterions*, It. *cantiamo/canteremo/canteremmo*, Pt. *cantamos/cantaremos/cantariamos* 'we (are) sing(ing)/we will sing/we would sing'. Non-relational functions such as number, person, tense, or mood are expressed internally to the verb. Relational functions that establish a relation between two terms tend, on the contrary, to be expressed in Romance languages by analytic means: Lat. *DULCIOR* 'sweet.COMPR' vs Sp. *más dulce*, Ro. *mai dulce*, It. *più dolce*, Nap. [kkju d'dotʃə] lit. 'more sweet', Lat. *MARCI FILIUS* 'MARCUS.GEN SONI.NOM' > It. *figlio di Marco*, Pt. *filho do Marco*, Fr. *fil de Marc* 'son of Mark'.

5.2 Areal typology: Standard Average European and the Romance languages

A different approach within the typological perspective deals with the geographical dimension of interlinguistic contact and the possible rise of linguistic areas. In this domain, a much-discussed topic in the last decade concerns the reshaping of Whorf's (1941) impressionistic intuition of Standard Average European (SAE) into a workable and even measurable/gradable concept to identify a linguistic area embracing the core of western Europe (cf. van der Auwera 1998; Haspelmath 1998; 2001; for a more recent assessment, see van der Auwera 2011; see also §4.4). Romance languages all participate, more or less intensively, in the phenomenon, as do Germanic languages, western Slavonic, and Balkan languages; however, the very core of the SAE linguistic area appears to be centred on the felicitously named 'Charlemagne Sprachbund' (van der Auwera 1998:824), which involves German, Dutch, and, among the Romance languages, just French and northern Italian dialects.

Most areally shared features of SAE appear indeed to have originated in the early Middle Ages (Haspelmath 2001), plausibly in a period of widespread bilingualism, especially between Latinized populations and Germanic newcomers in central Europe. Historically, the progressive emergence of SAE may have had its critical period at the time of the expansion of the Franks in central Europe (Pippin and his son Charlemagne) and the diffusion of the Holy Scriptures as well of their vulgar translations into Germanic, Romance, and even Slavonic languages.

Several of these features probably spread later and gradually in a wider area in the following centuries, often via the mediation of high-prestige varieties (Latin first of all, but later some established national languages as well, e.g. German for their eastern neighbours such as western Slavonic and Hungarian). Therefore, it is still an open question how pervasive SAE features are at the level of non-standard varieties, even those strictly related to the 'core' SAE languages (cf. Murelli and Kortmann 2011). A clear example of a defining SAE feature strongly influenced by sociolinguistic factors is given by the relative clause construction employing a relative pronoun, i.e. a quite peculiar subordinator which marks the syntactic role it assumes in the relative clause, but agrees at the same time in gender/number with the relativized head, e.g. Pt. *As senhoras das quais falei são professoras* 'The women of whom I spoke are professors'. Since the relative pronoun so defined is extremely rare outside Europe but nearly universal in standard west European languages (Comrie and Kuteva 2011), it appears to be one of the best diagnostics for SAE (Haspelmath 1998:279). However, looking in more depth beyond standard languages, the construction turns out to be rare or nonexistent, e.g. in most dialects of Italy (cf. §64.4) and Germany (Fleischer 2004). In Italo-Romance dialects, the most common strategy is pronoun retention (Cennamo 1997b:194): the relative subordinator is invariable (most often [ke] 'that') and the argument role of the relativized noun is coded by a resumptive clitic within the relative clause.

- (9) [el 'tozo ke ge go pre'sta el 'libro] (Pad.)
the boy that DAT.3SG= I.have lent the book
'The boy I lent the book to.'

The limitation of relative pronouns essentially to standard varieties is readily explained by interpreting its spread as a reflex of the Latin model on the rhetorical/communicative strategies of prestigious European languages, but sheds some doubts on the definition of SAE if we do not take into account its sociolinguistic context.

5.3 From Latin to Romance: typologically significant category losses and innovations

5.3.1 Determiners

One of the most important innovations characterizing the Romance languages is the introduction of the definite and indefinite articles (for detailed discussion see §§30.3.3, 46.3.1.1). As is known, the definite article derives from

Latin demonstratives used as deictic reinforcement, mainly at the level of the spoken language; witness in (10) the use of distal demonstrative *ILLE* (> Cat./Sp. *el~la*, Fr. *le~la*, It. *il/lo~la*, Ara./Cor./Glc./Pt. and SitR. *o~a*, Occ. *lo(u)/el~la*, Ro. *-(u)l~a*).³

- (10) Phaselus ille quem uidetis
little.boat.NOM that.NOM.M who.ACC.MSG see.2PL
hospites [...] (Lat., Catullus 4.1)
guests.VOC
'That little boat you see, strangers...'

Other Romance languages derive their definite article from Lat. *IPSE*, originally a topic marker and intensifier (e.g. Lat. *ipse Caesar* 'Caesar himself', Cic. *Fam.* 6.10.2), which yields Srd. *su~sa*, and Balearic and Costa Brava Cat. *es~sa*.

As for the indefinite article, it continues the Latin cardinal number *UNU(M)~UNA(M)* 'one' (> Cat./It./Sp. *un~una*, Fr. *un~une*, Occ. *u(n)~uno*, Pt. *um~uma*, Ro. *un~o*, Cal. *nu~na*), as happens in many other linguistic traditions (cf. Dryer 2011a). As in the case of the definite article, an article-like use of *UNUS* is already attested in Classical Latin, although rarely and in very limited contexts:

- (11) non [...] Pompeium tamquam unus
not Pompey.ACC like a/one.NOM.MSG
manipularis secutus sim. (Lat., Cic. *Att.* 9.10.2)
soldier.NOM having.followed be.SBJV.PRS.1SG
'I did not follow Pompey like a private soldier.' (De la Villa 2010:228)

The simultaneous presence of the indefinite and definite article in a language is considered by Haspelmath (2001:1494) one of the most significant features of SAE as a linguistic area. In the overall picture, all Romance languages do in fact show this feature, although, as noted above, not all Romance languages belong to the SAE core. Nevertheless, conditions of use of the articles differ sharply across Romance languages, and cannot be discussed here in detail (but see e.g. Stark et al. 2007). In some varieties, the definite article may not occur with the inherently definite personal names, in others it is possible or obligatory (some varieties of Catalan even have a dedicated form: *m en/na*). Variation among the Romance languages is found regarding co-occurrence with possessives: cf. Fr. *(**la) ma valise* vs It. *(**la) mia valigia* '(the) my suitcase'. Also the possibility of a zero-article or 'bare' NPs (chiefly mass nouns, abstract nouns, and/or postverbal plural indefinites) varies greatly, with a

³ Also in other linguistic traditions (e.g. Semitic) the source of definite articles is to be sought in the spoken style of non-literary texts (cf. Putzu and Ramat 2001).

minimum in French: cf. It. *ho comprato (del) pane e (delle) cipolle* vs Fr. *j'ai acheté du pain et des oignons* 'I bought (some) bread and (some) onions'. Finally, for mass nouns a central area within Romance languages—excluding the whole of the Iberian Peninsula, and southern Italo-Romance (but for old Neapolitan see Ledgeway 2009a:189-91)—has developed a partitive article (cf. §6.4.2), which is obligatory in French while it may alternate with zero in Italian. In some Italo-Romance varieties it may coincide with the simple preposition 'of' rather than articulated forms: cf. Tur. [i l aj ka'ta d paj] 'I bought some (lit. 'of') bread'.

5.3.2 Auxiliaries

Another important change in Romance concerns the emergence of a new system of auxiliaries (cf. §46.3.2.1). According to Benveniste (1968) we have here a 'conservative change', whereas in the previous case of the articles we can speak of a 'innovating change' inasmuch we can observe the emergence of a new category (albeit not completely unknown to the mother language: see (11)). The Romance verb system introduces many periphrastic forms that have roughly the same function as the synthetic forms of the mother language (cf. §46.1). The Latin first person singular perfect *DIXI* corresponds both to It. *dissi* 'I said' (simple past, the so-called 'passato remoto') and *ho detto* 'I have said' (the so-called 'passato prossimo'). The same holds for many Romance languages (cf. §58.2.3), though the frequency and the functions of the periphrastic vs the simple forms may vary from language to language.

The bridging contexts which set the conditions for the change to develop are already found in Latin. Compare (12) with (13), (14):

- (12) neive quis in eo agro agrum oqupatum
and.not someone in that field land.M.ACC occupied.ACC.MSG
habeto (Lat., *CIL* I² 585, 25; 111 BC)
have.IMP.FUT.3SG
'and nobody must keep occupied the land in that field'

In (12) the 'immediate constituent' analysis is [[*agrum oqupatum*]_{NP} [*habeto*]_V]_{VP} (with OV, but Det^{ee}+Det^{er} in the NP). Hence, the meaning of (12) is not that of a perfective action in the past ('someone occupied'); rather it refers to a universally valid state of affairs enshrined in the law.

On the other hand, in (13) and (14) the past participles *exquisitum* and *probatum* may, and probably should, be interpreted as belonging not to the NP but to the VP, giving it a perfect tense value (see Ramat 1987:142-7; Ledgeway 2012a:130-33). Note that *probatum* does not even agree with *haec omnia*.

- (13) dicam de istis graecis [...] quid Athenis
I.will.say about these Greeks what Athens.LOC
exquisitum habeam (Lat., Cato *ad fil.* frg. 1)
discovered.ACC.NSG have.SBJV.PRS.1SG
'I will say about these Greeks [...] what I have found
out in Athens/(what I hold as discovered in Athens)'
- (14) haec omnia probatum habemus
this.ACC.NPL all.ACC.NPL tried.ACC.NSG we.have
(Orib. *Synopsis* 7.48; Lat. transl. 6th c.)
'we have tried all these things'

In the course of history of the Romance languages the periphrastic construction became more and more grammaticalized, and in some contemporary varieties it has almost replaced the ancient perfect in all its functions (cf. §58.3.2). The extreme cases are spoken French, most dialects of northern Italy, Sardinian, and spoken Romanian, where the simple past has disappeared altogether.

Typologically, the use of a 'have + past perfect participle' construction to form a perfect tense (whether or not it further grammaticalizes into a general perfective past) is another strong diagnostic of SAE, according to Haspelmath (2001:1495). Together with Romance, it is found in all Germanic and some Balkan languages: Greek, Albanian, and Macedonian. On the other hand, it is practically unknown outside Europe (Dahl 1995; note, however, that a transitive possession verb 'have' is by itself relatively uncommon across the world). The parallel constructions in the different branches of European languages, arising roughly in the same period, point to a complex contact picture (cf. Giacalone Ramat 2008:135-43). The case seems to exemplify well the situation of '(contact-induced) replica grammaticalization' or even a 'grammaticalization area', in the words of Heine and Kuteva (2005:92-4, 182-5).

Alongside the perfect periphrastic constructions with HABERE there also exist those with auxiliary ESSE 'be', the distribution of which is discussed in detail in §49.3.

Other verbs, too, progressively assumed the role of auxiliary, and the Romance languages offer a large palette of different functions for the new auxiliaries. For example, Spanish distinguishes between a dynamic passive auxiliary (*ser*) and a stative passive auxiliary (*estar*):

- (15) El coche era pintado. vs
the car *ser*.IND.IPFV.3SG painted
'The car was being painted'
El coche estaba pintado. (Sp.)
the car *estar*.IND.IPFV.3SG painted
'The car was painted.'

Reflexes of STARE 'to stand' are the most widespread auxiliary for expressing the progressive when followed by the

gerund: Sp. *estaba trabajando*, It. *stavo lavorando* 'I was working', although other strategies may occur. Many of them are based on a location schema (cf. Heine 1993:32): Gen. [sun de're: a 'skri:ve/sun ki ke 'skri:vu 'i:na 'letja] (Toso 1997:217) 'I'm writing a letter' (lit. 'I'm behind to write'/'I'm here that I write a letter').

It. *andare* 'to go' may be considered a passive auxiliary in sentences where the main verb has a meaning such as 'lose, destroy' (see Giacalone Ramat 2000:126f.):

- (16) La memoria di questi fatti va/ è andata
the memory of these facts goes/ is gone
perduta. (It.)
lost
'The memory of these things is being/has been lost.'

In non-perfective tenses, *andare* may freely combine with the passive past participle of all transitive verbs; however, it acquires a deontic meaning with the passive, as in (17):

- (17) La memoria di questi fatti va/andava/andrà
the memory of these facts *go*.PRS/IPFV/FUT.3SG
mantenuta/salvata. (It.)
maintained/saved
'The memory of these things must be/had to be/will have to be maintained/saved.'

In several Romance languages, the same movement verb has been grammaticalized as a future auxiliary, in an allative construction quite parallel to Eng. *going to*. In instances such as (18), there can still be ambiguity with the original meaning as a verb of motion:

- (18) Las señoras van a dormir. (Sp.)
the ladies go to sleep.INF
'The ladies go to sleep/will (soon) sleep.'

Similarly, Fr. *aller* and Pt. (especially Brazilian) *ir* 'go' are used as auxiliaries for expressing future, especially if established as certain:

- (19) Je vais [**à] lui téléphoner. (Fr.)
I go.IND.PRS.1SG to him.DAT= phone.INF
'I'm going to call him.'
- (20) Vou [**a] ver um filme no cinema. (Pt.)
go.IND.PRS.1SG to see.INF a film in.the cinema
'I'm going to see a film at the cinema.'

Catalan also has the construction *anar a* 'to go to' + infinitive. However, its meaning is not future proper, but rather imminent: its limited uses as a future proper are mostly considered as due to Castilian interference (cf. Fabra

1986:88; Badia i Margarit 1962, I:394, n. 12; Gavarró and Laca 2002:2693f.). On the other hand, in Catalan a very similar construction (*anar* 'to go' + infinitive, without a 'to') has the opposite meaning: it is the current punctual past tense in the contemporary language. On the discourse-based motivations which led to such an unexpected development, see Detges (2004). Example (21) shows both periphrases in the same sentence (Gavarró and Laca 2002:2693):

- (21) El fiscal anava a esternudar, però va
the prosecutor *go*.IPFV.3SG to sneeze.INF but *go*.PRS.3SG
aguantar-se. (Cat.)
restrain.INF.=himself
'The public prosecutor was about to sneeze, but he restrained himself.'

In a few Gallo-Romance varieties the verb 'come' may be used as an auxiliary for the immediate past, but in this case the construction involves, much less unexpectedly, an ablative instead of an allative preposition:

- (22) Elle vient de le voir. (Fr.)
she comes from him= see.INF
'She has just seen him.'
- (23) Venh de me lavar. (Lgd.)
I.come from me= wash.INF
'I have just washed.' (Ledgeway 2012a:123).

Coming back to future periphrases, Romance displays a host of different patterns. Typologically, they match almost perfectly the main lexical sources for future 'grams' attested worldwide (cf. Bybee et al. 1994:251-71; Dahl 2000b), which comprise verbs of motion (see above) and modal verbs expressing volition (Ro. *voi* (originally meaning) 'I want' + infinitive, Jones 1993:90, and the many reflexes of the type HABERE DE/AD 'to have of/to'). Perhaps the only major type lacking is change-of-state futures (as in Ger. *ich werde* 'I become' + infinitive). However, the Romansh ventive futures (Dahl 2000b:321; Ledgeway 2012a:123) might fill the gap, given the very close semantic connection between 'come' and 'become', and the fact that Romance reflexes of UENIRE 'to come' may be often used to express just change of state, with no movement involved, e.g. It. *viene brutto*, lit. '(the weather) is coming bad'.

The change-of-state meaning is indeed the most plausible starting point for a third pattern of grammaticalization of 'come': namely, its use as a passive auxiliary with the passive past participle in Italian and various other Romance languages (and, because of language contact, in some Alpine Germanic dialects such as Cimbrian and Bavarian; Ramat 1998:227f.; Heine and Kuteva 2005:186; for non-IE parallels

see Giacalone Ramat and Sansò 2014). Note the imminent meaning in (25):

- (24) El ei vegnius cattaus. (Srs.)
he is come found
'He has been found.'
- (25) Masa vine pusă de ei în colț. (reg. Ro.)
table=the comes put.FSG by them in corner
'The table is going to be put by them in the corner.'

It is impossible to describe here all the verbs which function as auxiliaries in the Romance languages (cf. §46.3.2.1 and the list in Ledgeway 2012a:122-4). What matters is that the periphrastic forms have the auxiliary (=Det^{ee}) first and the main verb (=Det^{er}) following it. This is consistent with the general trend we observed in §5.1: the transition from Latin to Romance languages involves a tendential drift from Det^{er}+Det^{ee} to the reverse order. However, inversion is possible in the Romanian future and conditional, in archaizing usage (cf. 26) where the auxiliary is an eroded form of the verb *a vrea* 'want'.

- (26) Adormi-vom. (Ro.)
fall.asleep.INF=AUX.FUT.1PL
'We shall fall asleep.'

This inversion reminds one of the western Romance future grammaticalization pattern CANTARE HABEO 'sing.INF I have' with OV order. Traces of the periphrasis that gave rise very early in the history of the Romance languages to the new synthetic forms are still found in the (literary) Portuguese type *lavar-me-ei* lit. 'wash.INF=me=I.have (=I shall wash myself)', alongside *lavarei-me* lit. 'wash.INF.I.have=me', where the clitic pronoun is inserted between the basic verb and the ending that derives from Lat. HABERE 'to have'. Similarly, the conditional may be *lavar-se-iam* lit. 'wash.INF=selves=they.had (=they would wash themselves)'.

5.3.3 Word order change

Although Ledgeway (2012a:201, n. 29) observes that it is difficult to provide a consistent definition of the term 'head' as used by Nichols (1986:56f.) and others, there is a general agreement that the change from OV (i.e. Det^{er}+Det^{ee}) to VO (i.e. Det^{ee}+Det^{er}) is one of the main features characterizing the shift from Latin to the Romance languages. However, it has to be noted that Latin was not a rigid OV language with obligatory final position of the verb, like Turkish or Japanese. As a basically inflectional language, its word order was rather free. For instance, in

the very same author we find with an SOV order (27) and with OSV (28).

- (27) Caesar suas copias in proximum
 Caesar.NOM his.ACC.FPL troops.ACC in next.ACC.MSG
 collem subducit (Lat., Caes. B.G.1.22.3)
 hill.ACC withdraws
 'Caesar withdraws his troops to the next hill'
- (28) copias suas Caesar in proximum
 troops.ACC his.ACC.FPL Caesar.NOM in next.ACC.MSG
 collem subducit (Lat., Caes. B.G. 1.24.1)
 hill.ACC withdraws
 'Caesar withdraws his troops to the next hill'

Compare also (29) with (30), and these in turn with (31) where the verb interrupts the object NP:

- (29) maximas gratias agit (Lat., Cic. Att. 3.8.4) = OV
 greatest thanks does
 'He thanks very much'
- (30) agit maximas gratias (Lat., Cic. Att. 1.20.7) = VO
 does greatest thanks
 'He thanks very much'
- (31) maximas agit gratias (Lat., Cic. Att. 3.5.1)
 greatest does thanks
 'He thanks very much'

Examples (27) and (28) carry the same information (refer to the same state of affairs), but they focus on different components of the information. Even a formulaic sentence of epistolary style may have variants, as shown in (29)-(31).

The fact is that in strong inflectional languages almost every word (excepting adverbs, adpositions, and most numerals) has a morphological mark that assigns to the word its semantic function in the sentential syntax: the word is, so to say, self-sufficient, independently of its position (cf. Marouzeau 1922:1; Meillet 1964:439).

Romance languages progressively lost this freedom and word order became more rigid. One may find Lat. *Platonis libri* as well as *libri Platonis* 'the books of Plato/Plato's books' but not, say, It. ***di Platone i libri*, Fr. ***de Platon les livres*, or Pt. ***de Platão os livros*. The basic transitive sentence (see §5.1) e.g. 'John loves Mary'—with S, O, and V—shows an SVO order, just as in English where 'Mary loves John' has a totally different meaning. Generally, it can be said that in Romance the order within the NP/DP is much more rigid than that within the VP and sentential modifiers: in particular, sentence adverbs may surface more or less everywhere, although they show some preferential sequences.

Most variation within Romance does not involve the basic orders, but rather their rigidity, both diatopically and diachronically. While the unmarked order is SVO everywhere, marked sentence orders reflecting differences in informational structure are marginal in French, but quite common, for example, in Italian and Spanish: in Italian VOS and OVS are possible, and Spanish, besides VOS order pragmatically similar to Italian, has also VSO for all-rhematic sentences (Zubizarreta 1999:4232-4). Historically, both old Italian and old French displayed less rigidity in many constructions than their contemporary descendants (cf. §§31.3.3, 62.5).⁴

The increasing rigidity which developed from Latin to Romance languages has been traditionally connected with the loss of case endings. However, reduction of case endings is attested already in ancient Latin: the famous funeral inscription of Lucius Cornelius Scipio (*CIL* I², 9, second half of the third century BC) illustrated in (32) would be as in (33) in 'Classical' Latin.

- (32) honc oino ploirume cosentiont
 this.ACC.M one.ACC.M most.NOM.MPL agree.IND.PRS.3PL
 R[omane] / duonoro optumo fuise
 Roman.NOM.PL good.GEN.MPL best.ACC.MSG be.INF.PST
 uiro (OLat.)
 man.ACC
 'the majority of Romans agree that this one has been the best of the good men'

- (33) hunc unum plurimi consentiunt Romani/bonorum
 optimum fuisse uirum (CLat.).

We can see that final *-m* is regularly omitted in *oino*, *duonoro*, *optumo*, and *uiro*. This may involve neutralization with the nominative singular, because also the nominative ending *-s* is not fully stable: for instance, it does not appear in *Fourio*, from the inscription from Tusculum (*CIL* I², 49, third century BC):

- (34) M. Fourio C. F. tribunus militare de
 M. FURIUS.NOM C. F. tribune.NOM. military.ABL.SG from
 praidad Maurte dedet (OLat.)
 plunder.ABL Mars.DAT gave.3SG
 'The tribune M. Furius, son of Caius, gave (this object) to Mars from a military plunder'

⁴ To give some old Italian illustrations, from Salvi and Renzi (2010): degree adverbs could follow the modified adjective (*bella molto* lit. 'beautiful very', *lucidi troppo* lit. 'bright too.much', Ricca 2010:736); preverbal predicate adverbs were commonplace: *allora il cavaliere dolcemente le parlò* lit. 'then the knight sweetly spoke to her' (Ricca 2010:717); direct objects could freely occupy the post-auxiliary position: *i nimici avessero già il passo pigliato* lit. 'the enemies had already the pass taken' (Benincà and Poletto 2010:71).

Consequently, it cannot be said that the loss of final consonants automatically generated the use of prepositions, nor vice versa. In (34) we see that an ablative (*praidad*) is preceded by a preposition (*de*): case marking and prepositions are often used simultaneously. There has been a conspiracy between the two factors and the process leading from *Platonis libri* 'Plato.GEN books' to *i libri di Platone* 'the books of Plato' developed along a *continuum* lasting centuries. In (34) *Fourio* coexists with nominative case-marked *tribunos* (CLat. *tribunus*).

Note that in some Romance varieties, after the loss of the distinction between nominative and accusative (*tribuno* = nominative and accusative), a new direct object marker was developed much later via the prepositions *a* 'to' or (in Romanian) *pe* 'on':

- (35) Chiamàu a Micheli. (SCal.)
 he.called a Michele
 'He called Michele.'
- (36) Ana a văzut-o pe Maria. (Ro.)
 Ana has seen=her pe Maria
 'Ana has seen Maria.'

At this stage, the object relation is marked similarly to the other non-subject roles in the sentence. Moreover, the reintroduction of a direct object marker makes inversion easier: *A Micheli chiamàu* (though with *Micheli* now focalized), since the object is clearly marked. In no Romance language, however, has this prepositional marking of the object been generalized: it is limited to a subset of NPs, whose extension varies, but is generally identifiable as an upper segment of the well-known typological hierarchies of animacy and/or definiteness (cf. Bossong 1998:218-30). For instance, in (37) we have both the direct object with *a* and without *a*, depending on the different status of the two NPs along the hierarchy (a pronoun vs a NP with a lower degree of referentiality):

- (37) Più nobili cosa esti vinciri a se
 more noble thing is win.INF a REFL.3
 medemmi ca vinciri li jnimici
 selves than win.INF the enemies
 (OSic., [1321-37], Iemmolo 2009:202)
 'It is nobler is to overcome oneself than to overcome the enemies'

As for the other features related to the OV or VO order (Greenberg's 'universals'), we may observe, again, a freedom which progressively diminished in the Romance languages.

Latin, especially old Latin, could have the relative clause preposed to its referent head, thus with *Det^{er}+Det^{ee}*:

- (38) mandatae quae sunt uolo deferre epistulas
 given which are I.want deliver letters
 (Lat., Pl. *Persa* 694)
 'I want to deliver the letters that were given (scil. to me).'

However, the most frequent order was already N + relative clause (i.e. *Det^{ee}+Det^{er}*). In French, as well as in the other Romance languages, it would be impossible to have ***qui (m) ont été consignées je veux livrer les lettres* 'which (to. me=) have been given I want to deliver the letters'. What is possible is *les lettres qu'on (m) a consignées je veux les livrer* 'the letters which were given to me I want to deliver them', with a pronominal anaphoric resumption of 'letters' in the main clause.

The same holds for comparative constructions: Latin had both comparative + standard and standard + comparative, whereas Romance languages have just comparative + standard: Cor. *Antone hé più vechju ché/cà Filippu* 'Antone is older than Filippu'.

Finally, in Latin adjectives may precede (39) or follow (40) their head:

- (39) lex est recta ratio
 law.NOM.SG is right.NOM.FSG way.NOM.SG
 imperandi (Lat., Cic. *Leg. Man.* 1, 42)
 rule.GER.GEN
 '(the) law is the right way to rule'
- (40) in senatus populique Romani
 in Senate.GEN.SG people.GEN.SG=and Roman.GEN.MSG
 potestate (Lat., Cic. *Phil.* 6.4)
 power.ABL.SG
 'under control of the Roman Senate and people'

The traditional view is that the adjective is preposed when it is strictly connected to the noun as an epithet, as in (39), and postposed if it delimits a particular subclass, as in (40) (cf. Ernout and Thomas 1989:162f.): *urbanus praetor* 'an urbane praetor' (epithetic) vs *praetor urbanus* 'praetor of the town' (classifying). However, there are many counterexamples, such as *nauibus [...] onerariis vs onerariae naues* (Caes. B.G. 4.22.3/4.22.4) 'transport ships'. Other authors claim that the N-adjective order was already the unmarked in Classical Latin (see Ledgeway 2012a:210-13).

In Romance languages we have instances of clear semantic oppositions: Fr. *un pauvre causeur* 'a poor speaker' is a person who is not gifted for speeches, whereas *un causeur pauvre* is a speaker who is poor ('not rich'); It. *numerose famiglie* means 'many families' but *famiglie numerose* 'large families'; Sp. *un pobre hombre* is a 'poor fellow' whereas *un hombre pobre* is a 'poor man'. Apart from such cases of strong

contrast, the adjective–noun order is still quite generally available in some Romance languages, for example Italian, although the prenominal position has a restricted semantic range: it cannot have classifying function (relational adjectives like **il regionale treno* ‘the regional train’ are excluded).

From a general typological perspective, the relative flexibility of order of noun and adjective is not surprising, given that this order—contrary to Greenberg’s (1966) expectations—does not really correlate with the VO/OV parameter, as shown by Dryer (1988).

Reasons of space do not permit us to describe other facts related to word order(s) such as the usage of prepositions and postpositions (prepositions are in principle consistent with VO, postpositions with OV; Latin already shows a clear preponderance for prepositions). Summing up, Latin was a predominantly OV language, but not a rigid one, i.e. not consistently OV. The passage from Latin to Romance saw a drift from OV to VO, again not in a rigid way, as we have seen in the case of adjectives.

5.3.4 Sentence negation

A significant instance of differentiation in word order among the Romance languages is offered by sentence negation, reflecting the different stages of the well-known Jespersen’s Cycle (Jespersen 1917:4; see e.g. Schwegler 1990:153–74; Bernini and Ramat 1996; van der Auwera and Neuckermans 2004:458f; cf. also §51.2.1). The inherited order, with the Latin sentence negator *NON* ‘not’ preceding the finite verb in main declarative sentences, has been maintained, together with the marker itself, in the lateral areas (Iberian Peninsula, central and southern Italy, and Romania). Central regions—apart from Liguria and Veneto in northern Italy—have seen a host of basically parallel, but independent, processes of renewal, by which a multitude of different items (chiefly: (i) the negative quantifier ‘nothing’; (ii) lexical items originally denoting minimal quantities, such as ‘step’, ‘crumb’, ‘drop’, etc., so-called ‘negative polarity items’; and (iii) expressions of holophrastic negation originally occurring sentence-finally) have grammaticalized as sentence negators.

Most of the above processes give rise to a new *V-NEG* order, although many varieties (standard Fr. *ne...pas* is obviously the most familiar) display the intermediate stage of discontinuous negation. However, the new postverbal negators, due to their different historical origins, cannot be subsumed under a single syntactic position: for a detailed analysis, see Parry (1996) for Italo-Romance, and §51.2.2.

At any rate, the resulting *V-NEG* order taken as a whole is clearly marked from a general typological perspective: in Dryer’s (2011b) sample of 1,326 languages, the *NEG-V* type

(including both word-like and affixal negators, and even negative verbs, which do not have much in common with the process outlined above) is by far the most common worldwide, with about 52% of the total, while *V-NEG* types amount to about 28%, and the discontinuous ‘double negator’ type to about 15% (with several minor or hybrid types accounting for the remaining 5%).

As a matter of fact, French-based creoles in Louisiana and Guadeloupe have restored the typologically unmarked preverbal position of negation, thus completing Jespersen’s Cycle (Ramat 2006):

- | | | |
|------|-------------------|-------------------|
| (41) | Mo pa kup. (Lou.) | Nu pa vle. (Gua.) |
| | I NEG cut | WE NEG want |
| | ‘I don’t cut.’ | ‘We don’t want.’ |

A different typological perspective on sentence negation is taken in Miestamo (2005), whose main defining parameter is the dichotomy between ‘symmetric’ and ‘asymmetric’ negation. The negative construction is labelled symmetric if it does not differ from the corresponding affirmative in any other meaningful way than by the presence of negative markers (Miestamo 2005:61).

In particular, asymmetry may involve: (i) neutralizations in the negative paradigm (e.g. fewer tense–mood distinctions, or no person inflection in the negative constructions); (ii) the occurrence of a dedicated ‘negative verb’ which carries inflection; (iii) some sort of specialized tense–aspect–mood marking for negative constructions.

On the whole, despite the relevant functional motivations for the different kinds of asymmetric negation, symmetric negation turns out to be the most widespread type in Miestamo’s balanced language sample, which deals with main declarative sentences only. Romance languages all display the symmetric type in such sentences, although they often show (as Latin did) asymmetry in imperatives (especially in the second person singular), a favourite locus for asymmetry cross-linguistically (van der Auwera and Lejeune 2011). Negative imperatives may require a different verb form as in (42) and (43), or the use of a dedicated auxiliary, as [sta] ‘stay’ in (44):

- | | | |
|------|--------------|--------------------------------|
| (42) | Canta! | Non cantare! (It.) |
| | sing.IMP.2SG | NEG sing.INF |
| | ‘Sing!’ | ‘Don’t sing!’ |
| (43) | ¡Canta! | ¡No cantes! (Sp.) |
| | sing.IMP.2SG | NEG sing.SBJV.PRS.2SG |
| | ‘Sing!’ | ‘Don’t sing!’ |
| (44) | [ˈkanta] | [nu sta a kanˈta] (Gen.) |
| | sing.IMP.2SG | NEG stay.INF to sing.INF |
| | ‘Sing!’ | ‘Don’t sing!’ (Toso 1997:214). |

Note that in all three cases the prohibitive negation is the same as in declaratives. Cross-linguistically, this combination (special imperative + normal negative) is the less common of the four logical possibilities (van der Auwera and Lejeune 2011), and within Europe it is almost exclusively peculiar to Romance. Moreover, within Romance there is a very strong correlation between the two parameters discussed in this section: preverbal negation patterns with asymmetric imperatives, and vice versa. For a discussion within the generative framework, see Zanuttini (1997:105–54).

5.3.5 The clitic pronoun system and its grammaticalization potential

Another major innovation in the Romance system with respect to Latin is undoubtedly the development of clitic pronouns. These are found in all Romance varieties for the object and the dative function, and less extensively for oblique functions.

A series of clitic subject pronouns is also found in a central area of the Romance domain, comprising French, northern Occitan, Francoprovençal, Romansh except Surselvan, Ladin, Friulian, and northern Italo-Romance, including Florentine. This latter series is often not complete for all persons, depending on the variety. The literature on both subject and non-subject clitics is huge (see Chs 45, 47, 48). We limit ourselves to some brief considerations about the typological significance of this innovation.

Cross-linguistically, Romance non-subject pronominal clitics have rather close parallels in other Indo-European languages of Europe, especially Slavonic (cf. Dimitrova-Vulchanova 1999) and Greek (cf. Anagnostopoulou 1999). They do not fit particularly well into the picture of Standard Average European, however, and they have never been proposed as an identifying and typical feature for this linguistic area, although they were the object of a dedicated volume within the ‘Eurotyp’ project (van Riemsdijk 1999).

For all three families, perhaps the most significant typological issue is the extent to which they can be considered an instance of a head-marking strategy for non-subject roles in the sentence. This would mean in the case of Romance a significant change with respect to Latin, which is usually characterized as overwhelmingly dependent-marking, as seen in §5.1 (cf. Nichols 1986:89). Related to this issue, and applicable to subject clitics as well, is the question of the level of grammaticalization of clitics along the syntax–morphology continuum. A strictly head-marking interpretation of clitics would imply their turning into obligatory agreement markers on the verb, fully entering the domain of inflectional morphology.

From this perspective the relevant data involve so-called left- and right-dislocations (we use the term simply as a useful label, without implying that they are the result of any kind of movement rule), in which the full NP and its clitic resumption (or anticipation) co-occur in the same utterance, as in the following Italian examples where the object clitic *le* co-occurs with the full NP object:

- | | |
|------|-------------------------------------|
| (45) | Le chiavi le prendo io. (It.) |
| | The keys them= I.take I |
| | ‘I’ll take the keys myself.’ |
| (46) | a. Le ho prese le chiavi. (It.) |
| | them= I.have taken.FPL the keys |
| | ‘I’ve taken the keys.’ |
| | b. Le ho prese io, le chiavi. (It.) |
| | them= I.have taken.FPL I the keys |
| | ‘It’s me that took the keys.’ |

Sentences (45) and (46) are not pragmatically equivalent to those without a clitic such as (47) because (simplifying considerably) they tend to imply topicalization of the object NP, which in the unmarked sentence (47) is normally part of the rheme.

- | | |
|------|------------------------------------|
| (47) | Prendo/ho preso le chiavi. (It.) |
| | I.take/I.have taken.MSG the keys |
| | ‘I(’ll) take/I’ve taken the keys.’ |

As a matter of fact, and despite complications in the range of their pragmatic value(s) (Berruto 1985; Frascarelli 2003), dislocation constructions are never obligatory in Italian, from a strictly syntactic point of view. This means that the clitics cannot be considered (yet?) merely as agreement markers. Right dislocations are probably a better starting point than left dislocations for a full grammaticalization of the construction, because the transition to an obligatory ‘objective conjugation’, with the clitic as an agreement marker (cf. Berretta 1989), would be unproblematic in terms of the word order of the ‘major constituents’.

The grammaticalization process of the clitic towards an agreement marker is more advanced in Spanish, where constructions displaying clitic obligatoriness do exist, and therefore cannot be dealt with in terms of some marked pragmatic value (cf. Lehmann 1982:238). For instance, there is no alternative in Spanish if the object is a full personal pronoun, as in (48), or the quantifier *todo* ‘everything’, as in (49):

- | | |
|------|---|
| (48) | Yo lo/le vi a él. vs <i>**Yo vi a él.</i> (Sp.) |
| | I him= saw to him I saw to him |
| | ‘I saw him.’ |

- (49) Yo lo sabía todo. vs **Yo sabía todo. (Sp.)
 I it= knew all I knew all
 'I knew everything.'

Moreover, in many varieties of Spanish the dative clitic is obligatory in ditransitive constructions, as in (50) (from Company 2001:21):

- (50) Juan le /**∅ dio un libro a Pedro. (Sp.)
 Juan him.DAT= gave a book to Pedro
 'Juan gave a book to Pedro.'

The same obligatoriness for dative clitics is the rule in many Italo-Romance dialects, especially of the north (Poletto 1997:141):

- (51) [Ma'ria a I a 'da-je /**dajt vj
 Maria scl3sg= ∅= has given=DAT:3/given a
 'liber a dʒu'an] (Tur.)
 book to Gioan
 'Maria gave a book to Gioan.'

Although left and right dislocations are attested throughout the history of Romance, some diachronic studies on Spanish corpora point to a gradually increasing relative frequency of dative doubling constructions like (50) from the sixteenth century onwards, until becoming nearly obligatory in a twentieth-century Mexican written and spoken corpus (cf. Company 2001:23).

It may be puzzling that the obligatoriness of the dative clitic seems to be much more widespread across Romance than that of the object clitic. From the point of view of the grammaticalization target, one would expect the contrary, because languages which mark three arguments on the verb are much less common cross-linguistically than those which mark subject and object (or, more generally speaking, the three core roles Agent, Patient, and (intransitive) subject only; see §5.1). However, considering things from the other end of the scale, the dative argument (often expressing the Beneficiary or Experiencer) is usually much higher on the animacy hierarchy with respect to the Patient. This makes it a better candidate for individuation, saliency, and availability for anaphoric reference.

There is no space to deal even cursorily with the complex issue of subject clitics, for which see §45.2.2 and Ch. 47. Their behaviour is by no means uniform. Although they all originate from the Latin subject personal pronouns *EGO* 'I', *TU* 'you.sg', *ILLE* 'that.one', etc., subject clitics fill a continuum which has free (weak) personal pronouns at one end and fully bound morphological agreement markers at the other. Standard French is located at one end of the continuum, to the extent that in some approaches (e.g. Cardinaletti and

Starke 1999:167) *il* 'he' is not a clitic, but a weak pronoun, since it does not allow for doubling (*Jean **il chante* 'John sings') and allows for gapping in coordination (*Il chante et danse* 'he sings and dances'). Some informal varieties of spoken French, however, display exactly the reverse pattern, suggesting that in these varieties *i(l)* has turned into a 'true' clitic: *Jean i chante*.

In northern Italian dialects subject clitics display the highest level of complexity and variation, leading Poletto (2000:11-40) to propose four distinct positions in the preverbal subject clitic field. Moreover, in some varieties further postverbal clitics exist (e.g. Bol. [a 'si:] 'you.PL are', but [si:v] 'are you.PL?'; Hajek 1997b:277), which show a higher level of bondedness with the verb stem, and quite different paradigms, so that they have often been described as part of inflectional morphology (in terms of 'interrogative conjugation'). For a discussion, see Poletto (1998; 2000:42-55) and §47.4.2.

5.3.6 Gender and number categories

Although both gender and number will be dealt with in separate chapters (Chs 57, 42), it seems useful to mention some points of typological import.

At first sight, it might be thought that the two categories have remained substantially stable in the transition from Latin to Romance, as well as the *loci* of their expression (not necessarily the forms, of course), apart from the loss of one of the three values of the gender category, namely the neuter gender. However, a closer look yields a less straightforward picture.

As for gender, the analysis of data depends very much on how the category is defined. Referring to 'target genders', i.e. to the number of contrasting inflectional markers found in the paradigm of the agreement targets (adjectives, determiners, etc., cf. Corbett 1991), it is indeed the case that targets in Romance inflect for two values only (apart from the case of 'mass neuter' in central-southern Italy and Asturian, itself a disputed counterexample: see below). If, however, we deal with nominal gender ('controller gender' in Corbett 1991) and we take it as an inherent property of each nominal lexeme which divides nouns into distinct agreement classes, then Romanian indisputably has three of them, since most inanimate nouns select masculine targets in the singular and feminine targets in the plural (Maiden 2011a:172). To designate this third agreement class, the label of 'alternating gender', as used by Loporcaro and Paciaroni (2011), among many others, is probably less misleading than the traditional one of 'neuter', but does not change the need for a tripartition among nominals (cf. also Acquaviva 2008:136). The issue, however, is currently much

disputed: for a different approach, which postulates just two genders but assigns independent controller capability to the single nominal inflectional markers, see Maiden (2011a:701, n. 36; 2013b).

The Italian case for a distinct alternating gender is less clear. Nouns like *l'uovo* 'the egg', plural *le uova*, which select masculine targets in the singular and feminine ones in the plural, as in Romanian, do exist, but they are very limited in number (twenty or so), are not productive, and represent a residual class. For them, Acquaviva (2002; 2008:123-61) proposes a different description. Relying on the fact that several of such nouns admit two plural forms, an *-a* feminine and an *-i* masculine, often with difference in meaning (e.g. MSG *braccio* 'arm', FPL *braccia* 'arms' [of persons], MPL *bracci* 'arms' [of objects, e.g. *bracci di un fiume* 'arms of a river']), and relying further on a syntactic argument about gender resolution in coordinate phrases, Acquaviva suggests that the relationship between *il braccio* and *le braccia* is not inflectional, but derivational: according to him, *braccio/-i* is a regular masculine noun, and *le braccia* is a different lexeme, a feminine *plurale tantum* related to the former by a conversion process. Clearly, if *braccia* no longer belongs to the same inflectional paradigm of *braccio*, there is no need for an alternating gender at all. But see further discussion in §42.4.2. This argument cannot, however, be automatically extended to older phases of Italian; nor to several Italo-Romance varieties in central and southern Italy, where the type is much more widespread and the agreement patterns in coordination may work like in Romanian, as shown by Loporcaro and Paciaroni (2011:406-09).

Typologically, many of the Italo-Romance varieties mentioned above are more interesting than Romanian, because, besides a still vital alternating gender, they also display a different phenomenon, also a remnant of the Latin neuter, but kept fully separate synchronically: the so-called mass neuter. In these varieties, many mass nouns (not only original Latin neuters) require a set of agreement targets different from masculines, i.e. they belong to a different agreement class and therefore to a different gender. The opposition is reflected generally in the determiners (e.g. Mac. [lo pa] 'bread', mass neuter, vs [lu ka] 'the dog') and may extend to other targets like adjectives (Paciaroni and Loporcaro 2010). A similar, though not identical, phenomenon is present in Asturian and Leonese (Ojeda 1992). Leaving the details to the discussion in §§16.3.1.2 and §57.4, it is important to stress that this 'fourth gender' is productive (it is assigned to borrowings like 'sport', and applies e.g. to nominalized infinitives), and, unlike the alternating gender discussed above, it is also a target gender. As discussed in Loporcaro and Paciaroni (2011), four-gender systems are a rarity in Indo-European, where the dominant tendency has been

reducing the original three-value system or at most keeping it intact.⁵

It is fair to say that the very nature of 'mass neuter' as a phenomenon pertaining to gender does not meet universal consensus. For instance, Corbett (2000:124f.) considers the Asturian data as an instance of 'mass number' and Ledgeway (2009a:150) speaks of a third independent inflectional category [±num] for Neapolitan.

Coming now to number distinctions proper, the apparent typological uniformity of Romance data needs some qualification. Certainly all Romance varieties distinguish just two values for this category, singular and plural, as did Latin; but the expression of the category is not that uniform (for further discussion of the typological issues from a Romance-internal perspective, see also §42.1). A significant dimension of variation regards the form—or the bare existence—of morphological marking on the noun. Despite the overwhelming cross-linguistic prevalence for suffixation, for several Romance varieties suffixation turns out to be only a recessive strategy. In French, nouns are basically uninflected for number (despite orthography and liaison phenomena in very limited contexts, e.g. *les jeux olympiques* [le ʒøzələ'pik] 'the Olympic games'); only a small minority of nouns keep the number distinction in all contexts (essentially some words ending in [-al], with plural [-o], such as *cheval~chevaux* 'horse(s)'). In many colloquial varieties of Spanish, as in Andalusia and many areas of Latin America, the plural suffix *-s* is being lost, leaving—in the extreme case of full disappearance—most nouns uninflected for number (Penny 2000:122-5, 148-50). But probably the greatest wealth of cross-linguistically marked solutions is found in northern Italy. Invariability tends to be the rule—apart from small, phonologically conditioned classes and remnants of metaphonetic stem alternants—for all masculines (i.e. for the majority of nouns) in Piedmontese, Lombard, and Emilian varieties, while in Romagna the dominant strategy appears to be the typologically marked stem-change process, due to the relevance of morphologized metaphonetic processes. See, for instance, the examples—with both nouns and adjectives—from Lugo (Pellicciardi 1977, quoted in Maiden 1997b:21):

- (52) [spos] 'husband' - PL [spus]
 [mo^art] 'dead' - PL [murt]
 [a'me^ar] 'bitter' - PL [a'mer]
 [mes] 'month' - PL [mis]
 [bɛl] 'beautiful' - PL [be^al].

⁵ However, wider surveys would be needed here, especially concerning the possible emergence of alternating genders in languages which also keep the original Indo-European tripartite masculine/feminine/neuter distinction (cf. Loporcaro and Paciaroni 2011:413).

However, the most 'exotic' instance is found in western Lombard and Emilian varieties which couple the invariable or metaphonetic masculines with a subtractive (or at least anti-iconic) process for the -a feminines: Mil./Bol. [la 'skarpa], pl. [i 'skarp] 'the shoe(s)' (a very rare phenomenon cross-linguistically).

Although these data are basically the consequence of 'blind' phonetic changes, they led to a recurrent evolutionary tendency concerning the marking site(s) of plural information. Unlike Latin, number in Romance can be viewed as a stable inflectional category only if we look beyond the noun. Clearly, number is always a property of the NP (or DP) as a whole, but the dominant strategy in Latin, a typical fusional language, was the redundant marking of the feature on both the noun and its agreement targets (determiners, quantifiers, possessives, adjectives, and anaphoric pronouns). In the Romance varieties in which many or most nouns are invariable, very little has changed at the level of the phrase: nearly all phrases still carry unambiguous number information, due to the fact that number inflection has been generally preserved in determiners, quantifiers, and other grammatical items (and also verbs).

For instance, in Andalusian Spanish we have (53a,b) and in Torinese (54):

- (53) a. El/E(h)te perro peligroso e(h)tá
 the/this.MSG dog dangerous is
 encadenao. (And.)
 chained
 'The/this dangerous dog is chained up.'

- b. Lo/E(h)to perro peligroso e(h)tán
 the/this.MPL dog dangerous are
 encadenao. (And.)
 enchained
 'The/these dangerous dogs are chained up.'

- (54) a. [ɛl/ɛŋ/kust/kul 'liber] (Tur.)
 'the/a/this/that book.'

- b. [i/ɛd/'kusti/kuj 'liber] (Tur.)
 'the/some/these/those books.'

Some instances of preservation of inflectional marking cannot be easily accounted for on purely phonetic grounds, and may be restorations/conservations reflecting a morphosyntactic principle of generalizing the preferential locus for number marking outside the noun. In Torinese, this is the case for the masculine plural ending -i in the demonstrative ['kusti] 'these' above, and in quantifiers: ['tanti/'poki 'liber] 'many/few books.'

Similar instances of selective marking of plural on determiners only are reported for several other Romance varieties (e.g. Auvergnat and Brazilian Portuguese) by Ledgeway (2012a:290), who takes them as reflecting a more general tendency towards increase of head-marking strategies in the transition from Latin to Romance. This is certainly possible, but ultimately depends on the still problematic definition of head: articles are clearly heads with respect to N in the current DP approach of generative models, but would be considered as modifiers in other models, including that adopted by Nichols (1986) in her original proposal of the head/dependent typology.