

4 The Romance languages

Johannes Kabatek and Claus D. Pusch

1. Introduction
2. Phonology
 - 2.1. Vowel system
 - 2.2. Consonants
 - 2.3. Prosody
3. Morphology and syntax: some preliminary remarks
4. Morphology
 - 4.1. Inflectional morphology
 - 4.2. Derivational morphology
5. Syntax: General remarks
 - 5.1. Basic word order and alignment
 - 5.2. The noun phrase
 - 5.3. The verb phrase
 - 5.4. Complex syntactic structures
 - 5.5. Synopsis of some typological features

1. Introduction

Of all European language families Romance has the largest number of native speakers worldwide. A branch of Indo-European, it subsumes all the languages that historically evolved from Latin. Most prominent among these are global languages such as Spanish, Portuguese and French as well as European languages such as Italian, Romanian, Catalan. But there are also “lesser used” languages such as Occitan, Rhaeto-Romance, Sardinian and Galician. The number of Romance languages varies due to the notoriously imprecise boundary between language and dialect. In recent years, attempts to protect linguistic minorities have been accompanied by tendencies to classify linguistic varieties as languages rather than as dialects, and in some classifications of the Romance languages varieties such as Asturian, Aragonese, Valencian, Gascon, Corsican, Piemontese, Venetian or Aromanian appear as languages. We will not go further into this discussion, but it seems important to point out that all the lesser-used Romance languages and varieties must be considered, at least to some degree, as endangered.

A general distinction can be established between the “Old Romania”, i.e. the Romance languages spoken in Europe, and the “New Romania”, Romance spoken due to colonization or migration in other parts of the world, above all in the Americas and in Africa. Traditional, i.e. genetically and geographically based, classifications distinguish between Western and Eastern Romance languages (with a line

from La Spezia to Rimini, across the centre of the Italian peninsula, separating both according to phonetic and morphological differences); another classification distinguishes areas such as Balkan-Romance, Italo-Romance, Gallo-Romance and Ibero-Romance (with Catalan and Dalmatian, a language that died out at the end of the 19th century, as so-called “bridge languages” between the areas).

Table 1: Areal classification of the major Romance languages

Ibero-Romance	Portuguese
	Galician
	Spanish
Gallo-Romance	Catalan
	Occitan
	French
Italo-Romance	Franco-Provençal
	Italian
	Sardinian
Balkan-Romance	Rhaeto-Romance
	Dalmatian †
	Romanian

The following two maps show the worldwide distribution of the Romance languages and the European Romance language areas. Figure 1 shows the areas where Romance languages are spoken as L1 but also those areas (particularly in Africa) where Romance languages are official / national languages or lingua francas for a large part of the population:

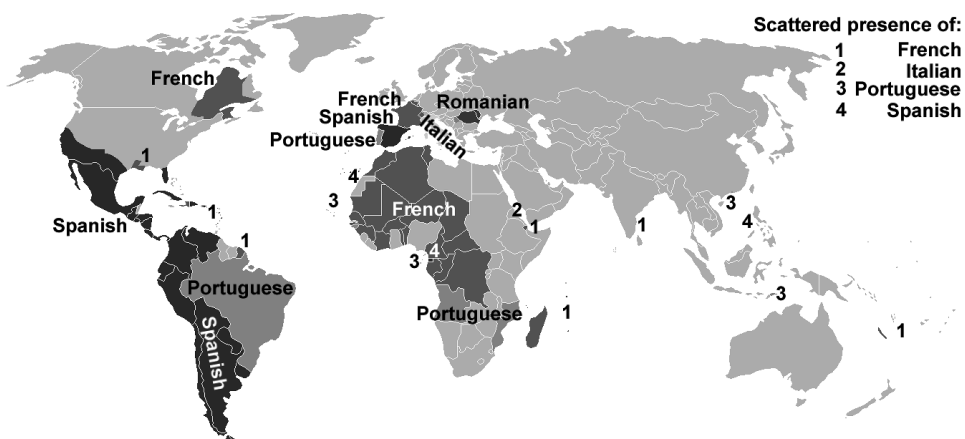


Figure 1: The Romance language areas of the world

Figure 2 gives an overview of the traditional language territories of the different Romance languages according to dialectological criteria. Striped areas indicate territories where Romance languages are nowadays used as official / national languages but which do not belong to their traditional dialectologically defined area of extension:



Figure 2: The Romance language areas in Europe

Considering the Romance languages as a typologically coherent object of study is justified not only on genetic grounds; areal and purely synchronic reasons can also be given. The genetic perspective is implied by the mere fact of classifying the Romance languages as a family. As for areal reasons, there exists an “alliance” relationship between the Romance languages, as Bossong (1998: 1005) baptized it: apart from their historical kinship, the Romance languages have also had a shared contact history during their independent evolution – above all with classical Latin, but also with other languages. Some scholars even claim (with evident shortcomings) that after the post-Latin period of divergence, there are now dominant tendencies towards convergence observable within the family. Apart from the inner-Romance areal convergence, areal parallels have been observed with Germanic (for the Gallo-Romance and partly the Italo-Romance area) and with Balkan languages (for Balkan-Romance within the “Balkan-Sprachbund”). In other cases, such as Brazilian Portuguese, areal arguments (claiming indigenous or African influence) seem more speculative.

If we look at the Romance languages from a morphological, syntactic or content-oriented synchronic perspective, there are several features common to all of them that justify the assumption of a more or less coherent Romance type different from Latin. Within this common Romance type, various subtypes can be identified. Traditionally, these subtypes have been related to different stages of evolution: whereas all the Romance languages have preserved until today a series of common typological features since the Middle Ages, there are a number of typological characteristics which emerged in the course of the last centuries, with French (contrary to Old French, which still belongs to the common type) as the most advanced and most innovative language among the Romance literary languages. However, it would be too simplistic to reduce the different subtypes to some sort of panchronic positions on a single line of evolution in the sense of a Sapirian drift. As will be shown in section 5, some of the fundamental evolutionary tendencies within the Romance family have clearly been divergent since the Middle Ages and led to results that must be described separately.

The Romance language family is one of the best documented ones among the language families of the world. Romance Philology as a discipline goes back to the 19th century, and even the evolution of typology itself is closely linked to Romance linguistics (see Bossong 2001). An historiographical overview of Romance Linguistics is offered by the monumental *Lexikon der Romanistischen Linguistik* (Holtus, Metzeltin and Schmitt 1988–2005); for a recent synopsis on historical and typological aspects see Maiden, Smith and Ledgeway 2011.

Almost all Romance languages are written in Latin orthography. Until the 19th century, Romanian used Cyrillic spelling, as did Moldavian, the Romanian variety spoken in Moldova, even in recent times when Moldova was part of the Soviet Union. Ladino, the Spanish varieties preserved by Sephardic Jews after their expulsion from Spain in 1492, is partly written in Hebrew spelling.

2. Phonology

2.1. Vowel system

Classical Latin had a rather simple vowel system with five cardinal vowel phonemes /a/, /e/, /i/, /o/, /u/ and two quantities. However, the common base for the Romance languages is not classical Latin but so-called *Vulgar Latin*, a term referring to the partly reconstructed spoken varieties in the Roman Empire. In Vulgar Latin, the classical quantities collapsed; the emerging system distinguished seven vowel phonemes (/a/, /e/, /ɛ/, /i/, /o/, /ɔ/, /u/); this system can still be found in several Romance languages such as Italian, Sardinian and Galician. The vowel system is reduced to the classical five vowels in Spanish (due to diphthongization of Vulgar Latin stressed E and O to *ie* and *ue* [BENE>*bien*; BONU>*bueno*]); in all the other

Romance languages, the systems have evolved into more complex ones. Central (Standard) Catalan has an eight vowel system, adding a central /ə/ to the common Vulgar Latin base; Occitan adds a rounded /y/ to the Catalan system; European Portuguese adds to the Vulgar Latin system two centralized vowels (/ɐ/ and /i/) and five nasalized vowels that emerged from the contact with nasal consonants as well as nasal diphthongs. Romanian presents a seven-vowel system with /a/, /ɛ/, /ɔ/, /i/, /u/ plus the marked /i/ and a Schwa /ə/. In Rhaeto-Romance varieties, systems between six and ten vowels can be found. The most complex Romance vowel system is that of European French, with 12 oral and four nasal vowels.

2.2. Consonants

Most initial consonants in Latin were preserved in Romance. Some consonants – with the exception of some Sardinian dialects – were palatalized in contact with front vowels, like *k^{e.i} g^{e.i}*. Initial clusters were partly reduced. Intervocalic voiceless stops were preserved in Eastern Romance and sonorized in Western Romance; intervocalic voiced consonants were partly lost in Western Romance. The tendency of Vulgar Latin to prefer vocalic word endings “remains very strong in Italian, is less pronounced in Spanish and Portuguese, even less so in Catalan, Occitan and Romanian, and it is weakest in French” (Jensen 1999: 299, see also section 4.2.).

“Geminated” consonants are characteristic for Italian and Sardinian, where stops in the syllable coda are generally long. The Ibero-Romance languages have preserved the Latin distinction between /r/ (a tap) and /r/ (a multiple vibrant); in other cases, the geminated consonants of Latin have generally been simplified.

2.3. Prosody

If we distinguish three dimensions of syllable structure – universal naturalness, L1-types and learned structures of later, L2-like acquisition – we can observe that, in contrast to their Latin ancestor, the Romance languages tend to simplify their syllable structures and to dominantly prefer the (universally “more natural”) CV type (cf. Lehmann 2005; like Lat. *TEMPUS* > Sardinian Logudorese *tempusu*). In the course of the history of the Romance languages, however, the contact with Classical Latin has repeatedly introduced more complex syllables associated with written language and oral standard, i.e. with “distance” varieties (see Koch/Oesterreicher 2001, Heinz 2010). Frequently this has led to a difference between more complex syllables in the standard variety and reduced and simplified syllables in the more colloquial varieties (like Standard Spanish *examen* [eksamen] vs. popular [esame]). Some Romance languages are syllable-timing and tend to preserve the syllable structure. Others (like French, European Portuguese or Central Mexican Spanish) show some characteristics of the stress-timing type (or rather of the word rhythm type in the sense of Auer’s 1993 distinction between syllable rhythm and

word rhythm), with tendencies towards vowel syncopations which, in turn, lead to consonant clustering and more complex syllables, like in European Portuguese (e.g. *desprezar* [dʃpr.zar]) or in Mexican Spanish (e.g. *bloques para apuntes* [bloks.pa.ra.punts]). Stress is phonologically relevant in several languages (e.g. Spanish *término* ‘border’ – *termino* ‘I finish’ – *terminó* ‘he/she/it finished’); in French, the accent position is fixed at the end of the phonological word.

3. Morphology and syntax: some preliminary remarks

Before describing some of the morphological and syntactic categories of the Romance languages, some general remarks on the Romance type seem to be necessary. As is commonly known, Latin is a highly synthetic language with a very rich morphology, a fact that allows an almost free word order in Latin and a “flat” syntactic structure (see Ledgeway 2011). The evolution from Latin to Romance is marked by the emergence of a series of syntagmatic functions that replace the Latin inflectional morphemes. This has traditionally been described as the shift from a synthetic (or more paradigmatic) to an analytic (or more syntagmatic) type, an idea already present in linguistics since August Wilhelm Schlegel at the beginning of the 19th century (see Bossong 2001: 719) and repeatedly mentioned in descriptions of the main transformations from Latin to Romance. This shift with a series of fundamental changes can be exemplified by a Latin sentence and its Romance translations, where almost every Latin word is replaced by at least two words in the Romance languages (see Table 2):

Table 2: Some major changes from Latin to Romance (adapted from Tecavčić 1972: 20; cf. also Ledgeway 2011)

Latin	VENDIDI	LIBRUM	MAIOREM	PETRO	MINOREM	TIBI
Italian	ho venduto	il libro	più grande	a Pietro	quello più piccolo	a te
French	j’ai vendu	le livre	plus grand	à Pierre	le plus petit	à toi
Spanish	he vendido	el libro	más grande	a Pedro	el menos grande	a ti
Romanian	am vândut	cartea	mai mare	lui Petru	cea mai mică	ție
	I have sold	the book	bigger	to Peter	the smaller	to you
	<i>I have sold Peter the bigger book and you the smaller one.</i>					

Some comments must be added to this general observation. The first point was repeatedly claimed by Eugenio Coseriu from the 1960s onwards (see Coseriu 1988), stressing that a distinction should be made between those functions expressed syntagmatically in Romance – thus corresponding to Schlegel’s shift –, and others that were not affected by it. Several inflectional functions are still morphologically expressed via nominal or verbal endings in Romance (with the exception of French; see section 4.1), especially, person, tense and number for verbs and gender and

number for nouns (see Table 3 for some examples of Spanish). This could simply be regarded as a casual effect of categories not yet affected by the shift, but in fact gender and number are even productive categories in the emergence of the Romance languages, and in several cases Romance languages have developed overt distinctions where they were opaque in Latin or Vulgar Latin. (Contrast, for example, ambiguous Latin *CASAE* with Spanish *casa / casas* or Classical Latin *ERAM, ERAS, ERAT*; Vulgar Latin *era, era, era* with Italian *ero, eri, era*). According to Coseriu (1988), the distribution of analytic and synthetic constructions, historically possibly related to Ancient Greek influence on Vulgar Latin, follows a typological principle that expresses formally a more general, inner principle of grammatical organization (in the sense of Humboldt's or Klimov's viewpoints on typology): in contrast to Latin, Romance codes "inner" functions, i.e. functions referring to the category itself, by inflection, whereas "external" relationships between several categories are coded by analytic constructions. (For French as a "purely" analytic language, see section 5)

Table 3: Some synthetic and analytic functions in Romance (here: in Spanish) according to Coseriu (1988)

synthetic, "inner functions"		
number	perro, 'dog'	perros 'dogs'
gender	perro 'dog'	perra 'female dog'
relative (without direct comparison)	grandísimo 'very big'	
analytic, "external functions"		
case	la casa del padre 'the father's house' se lo doy a Juan 'I give it to John'	
superlative (with comparison)	más grande que 'bigger than'	

Two further examples of this typological principle are, on the one hand, the emergence of Romance verbal periphrasis expressing relational tense functions (i.e. relations between two temporal points like in Spanish *voy a amar* 'I will love' relating present and future) and, on the other hand, the fusion of non-relational periphrastic forms to new synthetic forms, as in the case of the Vulgar Latin future *AMARE HABEO* > *amar aio* > Fr *j'aimerai*, It *amero*, Spn *amaré*.

This general organization principle leads to a kind of syntactic iconicity: inner functions that determine a referent or the category of the referent are expressed by different means than *syntactic* relations between various referents, where the analytic construction 'mirrors' the external relationship. This is illustrated in Figure 3: the diminutive in Spanish *casita* 'small house' is expressed by a suffix whereas the relationship between the doghouse and the dog is expressed by the preposition *de*:

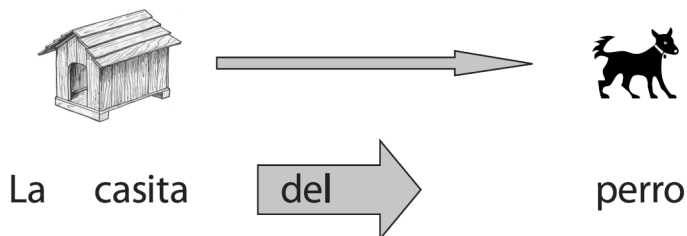


Figure 3: “Iconic” effect of the Romance distinction between synthetic and analytic forms

Another, more recent viewpoint considers the numerous shifts from analytic to synthetic constructions within a formal framework as epiphenomena related to a deeper change. This view denies the necessary existence of an organized syntactic linearity in (classical) Latin and postulates rather that the emergence of a fixed word order accompanied the shift towards Romance. This shift included the rise of a DP, a linearly structured NP and a structured VP, with a default order as described in Figure 4:

- a NP: (DET) (QUANT) (*ADJ) N (*ADJ) (*PP)
 b VP: (AUX) V (*ADV) (*OBJ) (*ADV)

Figure 4: Romance syntactic ordering contrasting with Latin (see Ledgeway 2011)

As several authors have pointed out, the deeper change underlying all these phenomena can be described as the shift towards a “configurational” structure, with an almost English-like linear syntax completely different from the rather “flat” structure of Latin. This means that both syntactic and morphological features need to be taken into consideration at the same time since diachronic changes made Romance syntax develop categories which were expressed by morphology in Latin. Other categories such as the article did not exist in Latin and were newly generated in Romance.

4. Morphology

4.1. Inflectional morphology

4.1.1. Verbs

The Romance languages have preserved conjugation systems, but the rich Latin verbal morphology is generally reduced in Romance. It is extremely reduced in French, where only tense is systematically marked by endings while person, number and mood distinctions have partly disappeared. For subject conjugation,

French employs preverbal subject clitics that grammaticalized from subject pronouns, whereas the personal endings have been preserved only in writing but not in the spoken language. The loss of endings in the Middle French period (between the 14th and the 16th century) affected not only the verb, but also the noun (see 4.1.2) and led to a complete reorganization of the French system: formerly postmodifying elements move to a premodifier position. Contemporary spoken Brazilian Portuguese also tends towards an overall reduction of endings and substitution by preverbal and pronominal elements. In the other Romance languages, apart from some local or partial tendencies towards the loss of endings, tense, person, number and mood are generally distinguished by verbal suffixes:

(1)	Latin	(1')	Spanish	(1'')	French
	<i>Leg-o.</i>		<i>Le-o.</i>		<i>Je lis.</i>
	read-1SG		read-1SG		1SG read
					'I read.'

Object conjugation is grammaticalized to different degrees and marked by clitic pronouns in some Romance languages. In European Spanish, object conjugation distinguishes human objects from non-human objects (the so-called *leísmo de persona*); this correlates with differential object marking by prepositions (see 5.3.2). In Portuguese, Galician, and a number of other Romance varieties (Sardinian, Sicilian, Ancient Napolitan, Ancient Romanian), inflectional endings can be added to infinitives.

Tense is the basic category of the Romance verb systems. In the history of the individual languages, secondary periphrastic systems have emerged which grammaticalized temporal verbal periphrasis to different degrees. This happened to a high degree in French, Northern Italian and Romanian, where the simple past perfective forms have been replaced by periphrastic forms. In other languages like Canadian French, Southern Italian or Spanish, synthetic and periphrastic perfect tense forms coexist with more or less differentiated meanings. The archaic verbal system in Galician is characterized by the non-existence, or perhaps rather marginal existence, of temporal verbal periphrases. Catalan has – as the only Romance language – grammaticalized a past perfect periphrasis formed on the basis of the movement verb *anar* 'to go' + infinitive:

(2)	Catalan									
	<i>Ahir</i>	<i>vaig</i>	<i>anar</i>	<i>a casa</i>	<i>de=l</i>	<i>meu</i>	<i>amic.</i>			
	yesterday	go.1SG	go.INF	at house	of=DET	my.M	friend			
										'Yesterday, I went to my friend's house.'

All verbal periphrases in Romance have secondary aspectual meanings deriving from the interaction of the semantics of the auxiliary and the non-finite form. The distinction between an imperfective and a perfective aspect is generally present in the Romance verb, but it is considered secondary compared with tense distinctions.

4.1.2. *Nouns*

By means of suffixes, the Romance languages distinguish morphologically different genders (male and female, with female gender generally being the marked form) and numbers (singular and plural, plural being generally marked). French again is an exception: in Middle French, endings were lost and the task of coding the gender/number distinction passed over to the determiner, which at that time began to become a gender/number prefix. (A similar tendency can be observed in contemporary Brazilian Portuguese.) The complex Latin case system was replaced by a reduced ‘dual case’ system in the ancient Gallo-Romance languages, and morphological case marking disappeared completely in modern Gallo-Romance as it did in the other Romance areas at even earlier diachronic stages. The only exception in this respect is Romanian, where a distinction between NOM and ACC, on the one hand, and GEN and DAT, on the other hand, is still marked morphologically, as in Table 4:

Table 4: Case marking in Latin and Romance

	NOM	ACC	DAT	GEN
Latin	homo	hominem	homi	hominis
Romanian		omul		omului
Spanish	el hombre	al hombre	al hombre	del hombre

In the other Romance languages, case is marked by word order or by prepositions, as in the Spanish examples for DAT and GEN in Table 4.

The Latin neuter gender, however, was lost in Romance. Some vestiges may be found in pronominal systems (like French *ça* ‘this’). In some languages and varieties, the Latin neuter has taken on new functions, like in the case of the so-called Spanish neuter article *lo*. This element no longer denotes a third gender, but has rather assumed the function of denoting imprecise, non-discrete reference (Pomino and Stark 2009). A similar case is the Asturian ‘material neuter’: apparently masculine adjectives accompanying female nouns indicate that the nouns are mass nouns, e.g. *mantega fresco* ‘fresh butter’ (see Fernández-Ordóñez 2006).

4.2. Derivational Morphology

The most productive procedure of word formation in Romance is derivation, above all suffixation and, to a lesser degree, prefixation. Categorial conversion without overt marking, almost impossible in Latin, is less generalized than in English or Chinese, but possible in several Romance languages (Adj → Noun Fr *noir* → *le noir*; Verb → Noun It *parlare* → *il parlare*).

The dominant type of compounding consists in preposing the determined element to the determining element, so that the general direction of compounding is the opposite of that to be found in English or other Germanic languages (e.g. Fr *wagon-lits*, Spn *coche-cama*, ‘sleeping car’; Fr/Spn *OTAN Organisation du traité de l’Atlantique Nord / Organización del Tratado del Atlántico Norte*; Engl NATO ‘North Atlantic Treaty Organisation’). However, some loanwords are Germanic-type compounds (like Fr *télévision*/Spn *televisión* ‘television’). Apocopations are common in several Romance languages, primarily in informal speech, and very common in spoken French, where they have phonotactic consequences when leading to final stops (*baccalauréat* → *bac* ‘high school graduation’). For an overview on lexical typology, including word formation, see Koch (2001).

5. Syntax: General remarks

The traditional classifications leading to clear-cut, discrete subcategorizations of the Romance languages are primarily geographic and generally based on phonetic and/or morphological features. More recent views prefer to consider the Romance languages as forming a “continuum with clines” (Bossong 2008: 273). We have seen that the general morphosyntactic evolution from Latin to Romance clearly allows us to distinguish a Romance type different from the type represented by Latin (section 3). At first glance, this seems to be confirmed by syntactic properties, where the Romance type appears to develop towards (S)VO, to prefer analytic, periphrastic forms instead of morphological case marking, to show overt determiners, and to place nominal modifiers to the right of the noun. This overall view is certainly justified; upon closer inspection, though, arguments can be, and have been, presented for a subclassification into two fundamental syntactic types (what Körner 1987 called “A-languages” and “De-languages”). An exhaustive list of typological features justifying such a classification would include features such as: non-obligatory vs. obligatory subject clitics, a more flexible word order vs. strict SVO, differential object marking vs. partitives, participle invariability vs. participle agreement, generalized auxiliaries vs. auxiliary selection, no partitive clitics vs. partitive clitics, etc. Spanish and French might be regarded as the prototypical exponents of the two Romance sub-types, and it is worth observing that these features, although having medieval roots, emerged or were enhanced after the Middle Ages, with some of the changes still going on (see Figure 5):

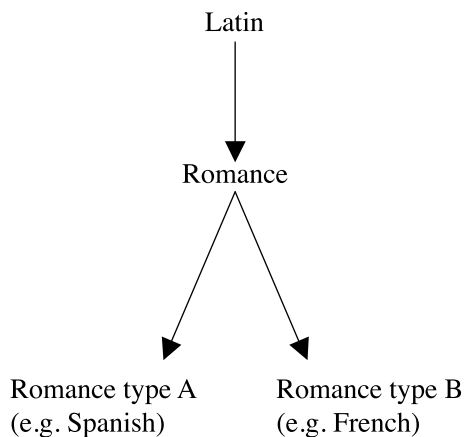


Figure 5: Typological shift from Latin to modern Romance

However, a critical view will have to add two further aspects: On the one hand, departing from the “prototypes” Spanish and French, several Romance languages and varieties cannot easily be placed in either category. A differentiation into several continua would show a less clear, but more adequate picture (see the synoptic Table 5 at the end of this chapter). On the other hand, even if correlations between the different features are evident, it is much more difficult to find an overall explanation that would allow us to discover the deeper principles underlying all the different historical developments. It has been claimed that if we look for a hierarchical order among the various continua, we should consider different means of expressing transitivity (Hopper and Thompson 1980; Fiorentino 2003) along the ergativity-accusativity continuum as the leading principle. Other semantic categories such as aspectuality or specificity also seem to play a crucial role. In sections 5.1 to 5.4, we will take a closer look at the most prominent typological features marking the contrast between different Romance subtypes.

5.1. Basic word order and alignment

All Romance languages are basically nominative-accusative languages, coding the distinction between nominative and accusative predominantly by word order. There is a rich literature on the overall changes in basic word order from Latin to Romance. The general evolution, related to the loss of Latin case marking, is towards fixation of word order. Latin, despite having a rather free word order, tended to prefer (S)OV. For the contemporary Romance languages, it can be stated that SVO is the most generalized type; it is rather strict in French and tends to be clearly dominant in those languages and varieties that have obligatory, or almost obligatory, subject clitics. For Old French, it has been claimed that the functional principle underlying the different overt structures was TVX. For Spanish and

other Romance languages, V1-sentences are unmarked for certain constructions. If we consider (3) and (4), several observations on Romance word order can be made:

- (3) Spanish
Juan trabaj-a.
 John work-3SG.PRS
 ‘John works / is working.’
- (4) *Entr-a Juan.*
 come_in-3SG.PRS John
 ‘John is coming in.’

Both sentences represent unmarked cases in Spanish word order, and the syntactic difference between them has been accounted for in a series of theoretical approaches. One of the basic assumptions is that word order in Romance is governed by information structure. As in most of the world’s languages, rhematic elements tend to follow thematic elements, and (3) is interpreted as a sentence where the “new” and relevant information is John’s action of working, whereas in (4) the relevant information is the person who comes in. But it would be too simple to restrict the basic word order difference to a purely discourse-pragmatic one: in fact, as has been claimed repeatedly, both sentences are semantically different. The proposition in (3) is of a “categorical” nature (i.e. it is a real proposition with two members), while it is of a “thetic” nature (with only one member) in (4) (Ulrich 1985). The logical argumentation for universally different kinds of propositions that may lead to different expressions in the languages of the world distinguishes scene-like “events”, such as a person coming in as in (4), from predications as in (3), where something is stated about someone. Following Burzio’s (1981) claim, a lexical difference between three fundamentally different classes of verbs is considered to be responsible for this syntactic differentiation (for discussion see Mackenzie 2006); the traditional distinction between transitive verbs like *to build* and intransitive verbs is reformulated with the addition of the distinction between two classes of intransitive verbs: *unergative* verbs like *to sleep*, where no agentive semantic subject can be identified, and *ergative* intransitive (or unaccusative) verbs like *to arrive*, where what is superficially a subject occupies the object position due to its being semantically a deep object of the verb in a (semantically) subjectless sentence (like *Juan* in example (4)). What can be claimed for modern Romance – with the exception of French and French-type varieties – is that basic word order is SVO for transitive and intransitive-unergative sentences, and that there exists the possibility, in unaccusative constructions, of moving the (syntactic) subject into object position when no real “semantic subject” is present.

Whereas in languages like Spanish this distinction is expressed by changes in word order, French has fixed word order but allows the choice between two differ-

ent auxiliaries: the auxiliary selected in unaccusative constructions is *être* ‘to be’, while it is *avoir* ‘to have’ in unergative intransitive constructions. Consider the following examples:

- (5) French
Jean est arriv-é.
 John is arrive-PSTPTCP
 ‘John has arrived.’
- (6) *Jean a dorm-i.*
 John has sleep-PSTPTCP.
 ‘John has slept.’

Spanish does not select different auxiliaries; other languages, like Italian or some Catalan varieties, combine both auxiliary selection and word order change. A series of further features can be related to this distinction, such as participle agreement, partitive object cliticization (only possible with unaccusative intransitives in French and Italian, impossible with unergatives) and blocking of passive transformations with ergative verbs (Gabriel and Müller 2008: 64–70). Whilst the first and the second actant are marked positionally in Romance, the third actant (as other case functions) is marked by prepositions and thus behaves in a more flexible way with regard to its position.

5.2. The noun phrase

All Romance languages are article languages with a definite article that emerged from the Latin demonstrative *ILLE* (exceptionally, in Sardinian and in some Catalan dialects, from *IPSE*) and an indefinite article that emerged from the quantifier *UNUS*. The grammaticalization of the article seems to be a proto-Romance phenomenon since all Romance languages have article systems. What is frequently ignored is that the “positive” grammaticalization of the Latin demonstratives, as illustrated in (7), is not an isolated fact and leads, in turn, to the grammaticalization of zero as a marker for mass nouns, as in (8).

- (7) Spanish
Com-o la/un-a manzana.
 eat-1SG.PRS det[F]/one-F apple
 ‘I eat the/an apple.’
- (8) *Com-o ∅ manzana.*
 eat-1SG.PRS DET apple
 ‘I eat apple.’ [+mass.]

This distinction makes it possible to mark syntactically the conversion of mass nouns into count nouns, and vice-versa (for partitives see 5.3.4).

Most of the Romance languages place determiners to the left of the noun. Only Romanian, in harmony with other Balkan languages and as a member of the *Balkan Sprachbund* (see chapter 16 by Tomić this volume), has an enclitic article: Rom *om* ‘man’ / *omul* ‘the man’ (but also: *cel om*) – Fr *homme ll’homme*; It *uomo ll’uomo*.

Attributive adjectives generally follow their head nouns; marked, pre-nominal adjectives tend to form units with the noun and to express idiomatic senses (Spn *buen chico* ‘good chap’), postnominal adjectives are of specifying nature (Spn *un chico bueno* ‘a boy who is good’), with some exceptions due to Germanic influence or due to a certain stylistic effect of the marked structure (Spn *la moderna filosofía* ‘modern philosophy’, *la madrileña plaza mayor* ‘Madrid’s major square’).

Sentential attributes in the form of relative clauses follow the noun and are introduced either by neutral relative pronouns (*que/che*) or by relative pronouns with gender and number agreement (Fr *lequel*, Sp. *el cual*, It *il quale* etc., see also 5.4).

Predicative nominals without copulas were common in Latin, but are rare in Romance. In most Romance languages, there is a general copula deriving from Lat. ESSE (with participles partly derived from STARE ‘to stand’, ‘to stay’). However, Ibero-Romance distinguishes two copular aspects, with a stative copula *estar* and an identifying copula *ser*, as in (9) and (10) respectively:

(9) Spanish

Juan es nervios-o.
John COP.3SG.PRS NERVOUS-M
‘John is a nervous person.’

(10) *Juan est-á nervios-o.*

John COP-3SG.PRS NERVOUS-M
‘John appears to be nervous (in this moment).’

5.3. The verb phrase

5.3.1. Subject clitics

Like their Latin ancestor, the medieval Romance languages did not require preverbal subject marking. However, in some Romance varieties, nominal or pronominal preverbal overt subjects became more and more obligatory, as in French. During the Middle French period, verbal and nominal endings got lost and the function of subject marking was more and more assumed by preverbal markers: Spn *trabajo* / It *lavoro* / Fr *je travaille* (see section 4.1). Obligatory subject marking can also be found in Rhaeto-Romance and Northern Italian and is becoming more and more common in spoken Brazilian Portuguese. A tendency towards obligatory subject markers has also been observed in Caribbean Spanish. Whereas some authors classify French as a strongly analytic language, it has also been claimed since the early

1960s that French has rather developed towards a synthetic language after the cliticization of the preverbal subject pronouns, the main difference to the other Romance languages consisting in the pre-modification of French in contrast to Southern and traditional Romance post-modification (see Weinrich 1962). In generative terms, this means that French would still be a pro-drop language (Kaiser 1992), with the former pronouns *je, tu, il, elle, on, vous, ils* being preverbal inflectional affixes that can be reinforced by diachronically more recent overt subjects *moi, toi, lui/elle, nous, vous, eux*:

Old French: *travaille* > Middle French/standard written modern French *je travaille* (obligatory pronoun) > modern spoken French *moi, je travaille* (with a tendency to phonic reduction of the “affix” *je*: mwaʃtr'vaj)

Figure 6: Evolution of subject marking in French

5.3.2. Differential object marking

In his early 20th century ‘neolinguistic’ areal-typological differentiation Matteo Bartoli (1925) distinguished between “inner Romance” and “lateral Romance” languages, the latter being diachronically more conservative since late innovations spreading from Rome did not affect these peripheral areas. Some lexical phenomena such as the existence of archaic words like Prt *fermoso*, Spn *hermoso* and Rom *frumos*, which contrast with It *bello*, Fr *beau/bel*, served to support this hypothesis. (Note that Spn *bello* is an Italianism which was incorporated later.)

A grammatical parallel between peripheral Ibero-Romance and Balkan-Romance is the existence of differential object marking, as in (11) and (12):

(11) Spanish

- a. *¿Has visto el tren?*
 have.2SG.PRS see.PSTPTCP.M DET train
 ‘Have you seen the train?’
- b. *¿Has visto a papá?*
 have.2SG.PRS see.PSTPTCP.M DOM dad
 ‘Have you seen dad?’

(12) Romanian

- a. *Ai văzut tren-ul?*
 have.2SG.PRS see.PSTPTCP train-DET
 ‘Have you seen the train?’
- b. *L-ai văzut pe tata?*
 OBJ-have.2SG.PRS see.PSTPTCP DOM dad
 ‘Have you seen dad?’

Differential object marking (henceforth short DOM) can also be found, with different degrees of grammaticalization, in varieties of Occitan, in Southern Italian

varieties and in Sardinian. It is generally obligatory with indirect objects and serves to distinguish human or human-like animate direct objects from other objects, allowing a series of subtle distinctions along the animacy-inanimacy continuum.

A phenomenon that has been related to DOM is the neutralization of the opposition between direct and indirect human pronominal objects and a preference for a human – non-human distinction (so-called *leísmo de persona*) in some European Spanish varieties (13a). By contrast, other European varieties and generally Latin American Spanish preserved the etymological distinction between male and female objects and indirect objects without gender distinction (13b).

- (13) a. European Spanish (with “*leísmo de persona*”; accepted as standard in the male case)

Le ve-o a Juan. Le ve-o a María.
 OBJ.HUM see-1SG.PRS DOM John OBJ.HUM see-1SG.PRS DOM Mary
 ‘I see John. I see Mary.’

- b. European Spanish (without “*leísmo de persona*”) / Standard American Spanish

Lo ve-o a Juan. La ve-o a María.
 OBJ.ACC.M see-1SG.PRS DOM John OBJ.ACC.F see-1SG.PRS DOM
 Mary
 ‘I see John. I see Mary.’

5.3.3. Clitic doubling

An important typological characteristic of the Romance languages is the tendency towards cliticization. Apart from articles, possessives, some prepositions and conjunctions, mainly object pronouns and, to a lesser extent (and in some languages) subject pronouns tend to appear as bound clitics. A phenomenon that like DOM can be related to the animacy hierarchy is *clitic doubling*, i.e. object clitics appear simultaneously with the full noun phrases they refer to (see 12b and 13). Clitic doubling is common in all those Romance languages and varieties that feature DOM and can further be found in colloquial Italian and in Northern Italian dialects.

5.3.4. Partitives

Those Romance languages that show DOM mark mass nouns in object position by zero (see 5.2.). Ongoing grammaticalization of DOM parallels the loss of partitives in these languages. Other languages have overt partitive elements derived from the Latin preposition *DE* or the pronominal adverb *INDE* (Fr/Ctl *en*; It *ne*). Partitives are

characteristic of contemporary Gallo-Romance and Italo-Romance. In French, the partitive determiner functions as a plural indefinite article, and therefore considerably reduces pre-nominal zero marking in this language. According to the animacy hierarchy, partitives are linked to inanimate objects and yield a [+mass] interpretation of the objects.

5.3.5. *Passive*

In some Romance languages (e.g. French), a periphrastic “real passive” (like in 14a) is more frequent than in other languages, like Spanish, where a “reflexive passive” (14b) is the most frequent form and “real passives” are practically restricted to written language. The “reflexive passive”, only possible in third person, is used, above all, for impersonal expressions since it almost precludes the explicit mention of the agent. French, on the other hand, prefers impersonal active constructions as in (14c):

(14) a. French

Le bâtiment est construit au centre ville.
 ART.M building is construct.PSTPTCP to.ART centre town

b. Spanish

El edificio se construy-e en el centro de la ciudad.
 ART.M building REFL build-3SG.PRS in ART.M centre of
 ART.F town

c. French

On construit le bâtiment au centre ville.
 SBJ.N.3SG construct.3SG.PRS ART.M building to.ART
 centre ville.
 centre town

‘The building is being constructed in the centre of town.’

5.3.6. *Negation*

Basically three types of sentential negation exist in the Romance languages: a) preverbal negation, b) discontinuous negation with a pre- and a postverbal element and, finally, c) postverbal negation (Bernini and Ramat 1996: 17). Type a), NEG+V, can be considered the “common Romance ground”. It is found in Latin and in the medieval Romance languages, and still is the type to be found in present-day Southern Romance (Portuguese, Spanish, Italian and Romanian), where it is accompanied by negative concord in postverbal elements. “Discontinuous” negation (type b) is generally a result of grammaticalization of postverbal reinforcing elements with an originally positive meaning: Fr/Ctl/Occitan *pas*, Piedmontese *pä*

(‘step’); Emilian, Lombard *miga* or *mia*, Ctl *mica*, Old French *mie* (‘crumb’); Piedmontese *ren*, Ctl *res* (‘thing’); Occitan *ges*, Ctl *gens* (‘people’), Fr *point* (‘point’). In spoken Brazilian Portuguese, discontinuous negation is common with repetition of the negation marker, in contrast to European Portuguese where preverbal negation is the rule:

- (15) a. European Portuguese / Standard Brazilian Portuguese:

Não quer-o.

NEG want-1SG.PRS

- b. spoken colloquial Brazilian Portuguese:

Não quer-o não.

NEG want-1SG.PRS NEG

‘I don’t want (it).’

This double negation with postverbal repetition is also common in several Romance-based Creoles. The third type (c) is the result of the elimination of preverbal elements in type-b)-areas and coexists with type b) as a colloquial variant (see also chapter 15 by van der Auwera in this volume and van der Auwera 2009).

- (16) a. Standard French

Je n’aime pas le concombres.

SBJ.1SG neg=like.PRS NEG ART.M cucumber

- b. Colloquial French

J’aime pas le concombres.

SBJ.1SG=like.PRS NEG ART.M cucumber

‘I don’t like cucumber.’

- (17) a. Standard Italian (Bernini/Ramat 1996: 21)

Ma non c’era niente da fare.

but NEG it=was.3SG nothing from do.INF

- b. Colloquial Italian

Ma c’era niente da fare.

but it=was.3SG nothing from do.INF

‘But there was nothing we could do.’

5.4. Complex syntactic structures

Clause-linking in Romance is carried out via different conjunctive morphemes, some of which evolved from Latin conjunctions whereas others were newly created in later periods. As in other European languages, many of these morphemes are used both as conjunctions and interrogative markers and are therefore grouped under the common label of *wh*-words. In the case of syntactic subordination, traditional grammar distinguishes between complement clauses, relative clauses and adverbial clauses. For complement clauses, a pan-Romance complementizer mor-

pHEME *que* (*che*, *că*, etc.) is used, which is a conflation of the Latin subordinating morphemes *QUEM*, *QUOD* and *QUIA*. Complement-clause formation being a way of nominalizing sentential strings in order to use them as verbal or nominal complements, the complementizer *que*, in a number of Romance languages, must be preceded by a preposition according to the valency of the governing matrix-clause element, as in (18a–b) from Spanish:

(18) Spanish

- a. *Tiene la convicción de que Dios*
 hold.3SG.PRS ART.F conviction from COMP God
exist-e (cf. *tiene la convicción de esta cosa*)
 exist-3SG.PRS
 ‘he is convinced that God exists.’
- b. *Me acuerdo de que es su*
 REFL remember-1SG.PRS from COMP is POSS.3
cumpleaños (cf. *me acuerdo de esta cosa*)
 birthday
 ‘I remember that it is his / her birthday.’

In colloquial varieties of Spanish, but also in some other Romance languages, there is a tendency to extend this type of prepositionally supplemented complementation to verbs which do not call for it from a valency point of view. This tendency, called *dequeísmo* in Spanish and rejected by normative grammar, is considered to express certain modalizing nuances. Consider (19):

(19) Spanish

- Dice de que es el cumpleaños de su*
 say.3SG.PRS from COMP is ART.M birthday of POSS.3
hermano (cf. *#dice de esta cosa*)
 brother
 ‘he says that it is his brother’s birthday.’

In some Romance varieties, including several Italian dialects, dual complementizer systems (i.e. *que/che* altering with another subordinating morpheme in the same slot) and complementizer-doubling, as illustrated by (20a–b) from Romanian, are attested. These phenomena of complementizer alternation correlate with mood contrasts and mirror different modal values of the subordinate clause.

(20) Romanian

- a. *Zicea că mâine nu se duce la câmp.*
 say.3SG.PRS COMP tomorrow NEG REFL go.3SG.PRS(IND) to field
 ‘he says that he doesn’t go to the field tomorrow.’
- b. *Zicea că mâine să nu se ducă la câmp.*
 say.3SG.PRS COMP tomorrow COMP NEG REFL go.3SG.PRS.SBJV to field
 ‘he says that tomorrow he should not go to the field.’

Complementizer deletion, on the other hand, is less frequent than in Germanic languages such as German or English and often controversial in normative terms but still found in several Romance languages.

Adverbial clauses are introduced by different conjunctions according to the semantics of the subordinate clause (cause, purpose, result, temporal or locative meaning). Some of these conjunctions have evolved directly from the corresponding Latin *wh*-words (e.g. Lat QUANDO > Spn *cuando*, Prt/It *quando* ‘when’; Lat QUARE [< QUA RE ‘which thing’] > Fr/Ctl *car* ‘because’, Lat UNDE > Fr *où*, Ctl *on*) but many other Latin forms have been replaced in medieval and modern times by complex conjunctions based on prepositional, nominal, adverbial or – less frequently – verbal elements and the complementizer morpheme *que*, such as Fr *afin que* ‘in order that’ (lit. ‘to the end that’), Ctl *tot i que* ‘although’ (lit. ‘all and that’), Spn *a pesar de que* ‘although’ (lit. ‘at weighing of that’), It *allorché* ‘when’ (lit. ‘at the hour that’). Some colloquial varieties of French in Canada tend to add the complementizer *que* even to simple, etymological adverbial conjunctions, yielding forms like *quand que*, *où que*, etc.

As for relative clause formation, Romance shares with many other European languages the typologically rather infrequent feature of relativizing conjunctive elements that carry inflectional features of the head noun in the matrix clause and, at the same time, inflectional features indicating the syntactic function of this noun in the subordinate clause (Cristofaro and Giacalone Ramat 2007). Such ‘true’ (dedicated) relative pronouns are therefore both anaphoric and cataphoric. In (21a), for instance, Fr *duquel* shows gender and number agreement with the head noun *le journaliste* and indicates that this noun functions as an oblique argument of the verb *recevoir* in the relative clause. With the alternative form *de qui*, illustrated in (21b), the anaphoric pronominal value corresponds to the indication of the animacy of the head noun:

(21) French

a. *Le journaliste du-quel j’ai*
 ART.M journalist from.ART.M-which.M SBJ.1SG=have.1SG.PRS
reçu cette information est normale-ment bien
 receive.PSTPTCP dem.F information is normal.F-ADV well
renseigné.
 inform-PSTPTCP

b. *Le journaliste de qui j’ai*
 ART.M journalist from who SBJ.1SG=have.1SG.PRS
reçu cette information est normale-ment bien
 receive.PSTPTCP DEM.F information is normal.F-ADV well
renseigné-é.
 inform-PSTPTCP

‘the journalist from whom I have received this information is normally well informed.’

However, upon closer inspection the cataphoric function (i.e. that of a syntactic slot-filler) turns out to be fulfilled by the prepositional element *de*, whereas the pronominal element (*lequel/qui*) is in charge of the anaphoric function. This holds primarily for those cases where the relativized element has a syntactic function which ranks low on the Noun Phrase Accessibility Hierarchy for relativization (cf. Keenan and Comrie 1977). In spoken Romance varieties, this tendency to disentangle and morphologically separate the anaphoric and the cataphoric function frequently leads to the use of the above-mentioned generalized neuter *que* as a marker of subordination and as a syntactic link with the head noun, while the syntactic slot in the dependent clause is filled by a separate (non-relative) pronoun, as in (22):

(22) a. Colloquial French

C'est le seul prof qu'on peut
 it=is ART.M only teacher REL=SBJ.N.3SG can.3SG.PRS
lui faire confiance.
 OBJ.DAT.3SG do.INF confidence

b. Standard French

C'est le seul prof à qui on peut faire
 it=is ART.M only teacher to who SBJ.N.3SG can.3SG.PRS do.INF
confiance.
 confidence
 'that's the only teacher you can be confident in.'

Que thus plays an important role in relative-clause formation, too, although the array of functions varies considerably between languages and registers. In Standard French, relative *que* can mark the relativized item as being the second actant only (in opposition to bare *qui*, indicating first-actant status). In Spanish or Italian, *que/che* alone may relativize both syntactic functions ranking uppermost on the Accessibility Hierarchy. For relativizations of lower-ranked syntactic functions the above-mentioned disentangling strategies involving *que* are frequently used. To sum up, the complementizer/relativizer *que* must be attributed a crucial role as a subordinating device in Romance.

The second central device for marking syntactic (and ensuing semantic/pragmatic) subordination in Romance is the use of the subjunctive mood. As compared with Classical Latin, the subjunctive paradigm of Romance has undergone heavy restructuring and simplification (see 4.1.1). On the one hand, many subjunctive forms have become formally identical to indicative forms, with the notable exception of irregular high-frequency verbs. On the other hand, only two tense forms survive, namely the present and the past subjunctive, which are of variable vitality in the different Romance languages and / or according to whether written or spoken language is taken into account. However, contrary to English and other Germanic languages, the subjunctive is an important feature in Romance and its loss from the verbal system is far from being imminent. Its distributive patterns are

complex and, furthermore, vary remarkably between languages. Generally speaking, the subjunctive is a modalizing device through which the speaker makes manifest that a statement depends on some external fact or subjective point of view, which then is encoded through syntactic dependency. In languages that closely stick to the Romance type, the subjunctive still has a certain degree of pragmatic availability and optionality, which may lead to variable use along with indicative or conditional verb forms, depending on the pragmatic context. However, in all Romance tongues there is a tendency towards fixation of the subjunctive after certain triggers in the matrix clause. Again, French proves to be most advanced in this process since it has almost fully grammaticalized the obligatory use of the subjunctive as a means of indicating syntactic subordination after a closed inventory of matrix-clause triggers. Only a limited degree of variability remains, e.g. in restrictive relative clauses where the subjunctive is used to distinguish between a factive and a non-factive reading (a usage found in other Romance languages, too):

(23) French

- a. *Je cherche un étudiant qui sait
parl-er chinois.*
SBJ.1SG search.PRS one(M) student who know.SG.PRS(IND)
speak-INF Chinese
'I am looking for a student who speaks Chinese.'
- b. *Je cherche un étudiant qui sache
parl-er chinois.*
SBJ.1SG search.PRS one(M) student who know.PRS.SBJV
speak-INF Chinese
'I am looking for a student who speaks (= should be able to speak) Chinese.'

Syntactic dependency correlates with reduced finiteness features of the verbal element in the subordinate clause (Raible 1992). As mentioned before, the Romance subjunctive, as a verbal paradigm typical for subordinate clauses, cannot express the same array of tense forms as the indicative and therefore lacks some finiteness features. Frequently, however, sentential subordination is replaced by even less finite verbal forms, namely infinitives and gerunds / present participles (the latter two being formally undistinguishable for most regular verbs). The replacement of restrictive relative clauses by an uninflected participle/gerund, well-known from English, occurs frequently in French (cf. 24) but remains marginal in other Romance languages. However, on top of that the gerund often replaces adverbial clauses expressing simultaneous action and manner / method, but also – particularly in Ibero-Romance – purpose and cause, as in (25) from Spanish:

(24) French

Je cherche un étudiant sachant
 SBJ.1SG search.PRS one(M) student know.GER
parl-er chinois (= qui sache / sait parler chinois; cf. 23)
 speak-INF Chinese

(25) Spanish

Siendo el responsable, tengo que preocup-ar=me
 be.GER ART.M responsible hold.1SG.PRS COMP WORRY-INF=REFL
de esto.
 from DEM

‘since I am the responsible person, I have to be worried about this.’

The infinitive replaces a finite subordinate clause if the logical subjects of the subordinate clause and of the matrix clause are identical (cf. 26); this, however, does not hold true for Balkan-Romance and Southern Italian dialects, where due to Balkan-Sprachbund influence (see 5.2) the infinitive is rarely used.

(26) Catalan

a. *Prefereix torn-ar a casa.*
 prefer.3SG.PRS return-INF to house
 ‘he prefers to go back home.’

b. *Prefereix que torn-i a casa.*
 prefer.3SG.PRS COMP return.PRS.SBJV(3SG) to house
 ‘he_i prefers that he_k goes back home.’ (but not: ‘he_i prefers that he_i goes back home’)

Since syntactic subordination always is a type of nominalization, the infinitive may display nominal features when used instead of a subordinate clause. This is the case in Spanish where temporal adverbial clauses may be replaced by an infinitive preceded by the preposition *a* and the enclitic definite article (cf. 27a). This construction may even appear with non-coreferential subjects in the matrix and the subordinate clause, as in (27b):

(27) Spanish

a. *A=l lleg-ar a casa, me di cuenta de mi error*
 at=ART.M arrive-INF to house REFL gave.1SG account from my error
 ‘when I arrived at home, I became aware of my mistake.’

b. *A=l lleg-ar la policia a casa, me di cuenta de mi error.*
 at=ART.M arrive-INF ART.F police to house REFL gave.1SG account
 from my error
 ‘when the police arrived at home, I became aware of my mistake.’

Even though the subjunctive may occasionally be used in main clauses, it is generally a good criterion to attribute subordinate status to the clause it appears in. European Portuguese, Galician and Asturian offer another morphosyntactic criterion for the main-clause / subordinate-clause distinction, comparable e.g. to the word order variation between main and subordinate clauses in German: in these Ibero-Romance languages, clitic pronouns are enclitic in main clauses and proclitic in subordinate clauses (cf. 28). However, this general enclisis/proclisis distribution may be altered by preposed sentence adverbs, negation markers and other focus-attracting or stance-expressing elements.

(28) Portuguese (European)

- a. *Ele compr-a=o hoje mesmo.*
 SBJ.3SG.M buy-3SG.PRS=OBJ.3SG today REFL
 'he buys it today'
- b. *Quer-o que ele o compre hoje mesmo.*
 want-1SG.PRS COMP SBJ.3SG.M OBJ.3SG buy.PRS.SBJV.3SG today REFL
 'I want that he buys it today.'

5.4. Synopsis of some typological features

Table 5 concludes this chapter with a synopsis of major typological features of the Romance languages.

Table 5: Romance: synopsis of major typological features

Feature → Language ↓	1	2	3	4	5	6	7	8	9
Romanian	+	+	-	-	-	+	+	-	-
Italian	-	-	+	+	-	+	+	-(+)	-
Rhaetoromance	-	(+)	+/-	+	+/-	-	+/-	-(+)	-
Sardinian	-	+	-(+)	+	-	+	+	-	-
Catalan	-	-	-(+)	-(+)	-	+	+	+	-
Spanish	-	+	-	-	-	+	+	+	+
Galician	-	+	-	-	-	+	+	+	+
Portuguese	-	+	-	-	-	+	+	+	+
Occitan	-	(+)	+	+	-	+	+/-	-	(+)
French	-	-	+	+	+	-	-	-	-

1 Morphological case marking

2 DOM

3 Partitives

4 Auxiliary type: *be/have*

5 Overt subject clitics

6 SV-VS-variation

7 *Se*-passives

8 *Be/stay*-copula

9 Simple past tense dominant

+ = general; - = inexistent; (+) = rare; +/- = existent

See also the following chapters in this volume: 15 by van der Auwera, 16 by Tomić, 18 by Sansò and 28 by Murelli and Kortmann.

References

- Auer, Peter
 1993 *Is a Rhythm-Based Typology Possible? A Study of the Role of Prosody in Phonological Typology*. Hamburg: KontRI.
- Bartoli, Matteo
 1925 *Introduzione alla neolinguistica*. Genève: Olschki.
- Bernini, Giuliano and Paolo Ramat
 1996 *Negative Sentences in the Languages of Europe. A Typological Approach*. Berlin/New York: Mouton de Gruyter.
- Bossong, Georg
 1998 Typologie des langues romanes. In: Günter Holtus, Michael Metzeltin and Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. 7, *Kontakt, Migration und Kunstsprachen. Kontrastivität, Klassifikation und Typologie*, 1003–1019. Tübingen: Niemeyer.
- Bossong, Georg
 2001 Typologie linguistique. In: Günter Holtus, Michael Metzeltin and Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. 1.2, *Methodologie (Sprache in der Gesellschaft / Sprache und Klassifikation / Datensammlung und -verarbeitung)*, 718–730. Tübingen: Niemeyer.
- Bossong, Georg
 2008 *Die romanischen Sprachen. Eine vergleichende Einführung*. Hamburg: Buske.
- Burzio, Luigi
 1981 Intransitive verbs and italian auxiliaries. Ph.D. dissertation, Massachusetts Institute of Technology.
- Cristofaro, Sonia and Anna Giacalone Ramat
 2007 Relativization strategies in the languages of Europe. In: Paolo Ramat and Elisa Roma (eds.), *Europe and the Mediterranean as Linguistic Areas. Convergencies from a Historical and Typological Perspective*, 63–93. Amsterdam/Philadelphia: Benjamins.
- Coseriu, Eugenio
 1988 Der romanische Sprachtypus. Versuch einer neuen Typologisierung der romanischen Sprachen. In: Jörn Albrecht, Harald Thun and Jens Lüdtke (eds.), *Energiea und Ergon. Studia in Honorem Eugenio Coseriu*. Vol. I, 207–224. Tübingen: Narr.
- Fernández-Ordóñez, Inés
 2006 Del Cantábrico a Toledo: El ‘neutro de materia’ hispánico en un contexto románico y tipológico. *Revista de Historia de la Lengua Española* 1: 67–118.
- Fiorentino, Giuliana (ed.)
 2003 *Romance Objects. Transitivity in Romance Languages*. Berlin: Mouton de Gruyter.

- Gabriel, Christoph and Natascha Müller
2008 *Grundlagen der generativen Syntax. Französisch, Italienisch, Spanisch*. Tübingen: Niemeyer.
- Heinz, Matthias
2010 Zentrum und Peripherie in der Herausbildung romanischer Silbentypsysteme. Habilitation thesis, Universität Tübingen.
- Holtus, Günter, Michael Metzeltin and Christian Schmitt (eds.)
1988–2005 *Lexikon der Romanistischen Linguistik*. Tübingen: Niemeyer.
- Hopper, Paul and Sandra A. Thompson
1980 Transitivity in grammar and discourse. *Language* 56: 251–299.
- Jensen, Frede
1999 *A Comparative Study of Romance*. New York: Peter Lang.
- Kaiser, Georg
1992 *Die klitischen Personalpronomina im Französischen und Portugiesischen. Eine synchronische und diachronische Analyse*. Frankfurt/Main: Vervuert.
- Keenan, Edward and Bernard Comrie
1977 Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.
- Koch, Peter
2001 Lexical typology from a cognitive and linguistic point of view. In: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds.), *Language Typology and Language Universals. An International Handbook*, 1142–1178. Berlin/New York: de Gruyter.
- Koch, Peter and Wulf Oesterreicher
2001 Langage parlé et langage écrit. In: Günter Holtus, Michael Metzeltin and Christian Schmitt (eds.), *Lexikon der Romanistischen Linguistik*. vol. 1.2, 584–627. Tübingen: Niemeyer.
- Koerner, Karl-Hermann
1987 *Korrelative Sprachtypologie. Die zwei Typen romanischer Syntax*. Wiesbaden: Steiner.
- Ledgeway, Adam
2011 Syntactic and morphosyntactic typology and change in Latin and Romance. In: Martin Maiden, John Charles Smith and Adam Ledgeway (eds.), *The Cambridge History of the Romance Languages*, 382–471. Cambridge: Cambridge University Press.
- Lehmann, Christian
2005 Latin syllable structure in typological perspective. In: Gualtiero Calboli (ed.), *Latina lingua! Proceedings of the Twelfth International Colloquium on Latin Linguistics (Bologna, 9–14 June 2003)*, vol.1, 127–147. Roma: Herder.
- Mackenzie, Ian
2006 *Unaccusative Verbs in Romance Languages*. Houndmills/New York: Palgrave.
- Maiden, Martin, John Charles Smith and Adam Ledgeway
2011 *The Cambridge History of the Romance Languages*. Cambridge: Cambridge University Press.
- Pomino, Natascha and Elisabeth Stark
2009 Losing the “neuter”: The case of the Spanish demonstratives. *Probus* 21: 217–248.

Raible, Wolfgang

1992 *Junktion. Eine Dimension der Sprache und ihre Realisierungsformen zwischen Aggregation und Integration*. Heidelberg: Winter.

Tecavčić, Pavao

1972 *Grammatica storica dell'Italiano*. Vol II: *Morfosintassi*. Bologna: Il Mulino.

Ulrich, Miorița

1985 *Thetisch und kategorisch: Funktionen der Anordnung von Satzkonstituenten am Beispiel des Rumänischen und anderer Sprachen*. Tübingen: Niemeyer.

van der Auwera, Johan

2009 The Jespersen Cycles. In: van Gelderen, Elly (ed.), *Cyclical Change*, 35–71. Amsterdam/Philadelphia: Benjamins.

Weinrich, Harald

1962 Ist das Französische eine analytische oder eine synthetische Sprache? *Mitteilungsblatt des Allgemeinen Deutschen Neuphilologenverbandes* 15: 177–186.