

# 1

---

## Overview of the grammar

### 1.1 Introduction

#### 1.1.1 *The extent of Modern Spanish*

Modern Spanish is spoken by just under 300 million people worldwide, and is thus one of the three or four most widely spoken languages, after Mandarin Chinese, English and possibly Hindi.<sup>1</sup> Spanish is the primary or official language in numerous countries, including Spain and its dependencies, Equatorial Guinea, eighteen countries of Central and South America, and the US protectorate of Puerto Rico.<sup>2</sup> Spanish is robust as a first or second language in many areas of the southwestern United States, as well as in other agricultural areas of the US, and urban areas such as Miami and New York. According to the 1990 census, about 17.3 million people over the age of five speak Spanish at home in the US.

Many countries in which Spanish is the official or primary language are linguistically diverse, with bilingualism a common, but not universal, phenomenon. In the north of Spain, primary languages include Basque, Catalan and Galician.<sup>3</sup> In Latin America, many indigenous languages are used alongside Spanish. In Bolivia, for example, at least half the population speaks either Aymara or Quechua natively, and it is estimated that 40% of these speakers

<sup>1</sup> Mandarin has well over 700 million speakers, English over 400 million. Estimates for Spanish speakers range from 266 million (Bright 1992) to 290 million (Green 1992), and estimates for Hindi range from 182 million (Bright 1992) to 290 million (Décsy 1986).

<sup>2</sup> Spanish is the official language of most countries of Latin America. In Peru, both Spanish and Quechua are official languages. In Bolivia, Spanish, Quechua and Aymara are all official languages.

Although Spanish is the official language of Equatorial Guinea, it is estimated that only 4–5% of the population speaks Spanish (Kurian 1992:600).

<sup>3</sup> Galician or Gallego is considered more closely related to Portuguese than to Spanish. Catalan is more closely related to Occitan than to Spanish. Basque is a linguistic isolate.

## 2 The syntax of Spanish

do not speak Spanish (Grimes 1988:85–87; Kurian 1992:184). In Paraguay, Guaraní is spoken by over 3 million speakers, with a majority of rural speakers being monolingual (Grimes 1988:125). Relatively large populations of speakers of indigenous languages are also found in Peru (Ayacucho Quechua and Cuzco Quechua), Guatemala (Mayan languages) and Ecuador (Quichua). Many other indigenous languages are spoken, by populations numbering from dozens of speakers to tens of thousands. Relatively small populations speak Creole languages in Honduras, Nicaragua, Costa Rica, the Dominican Republic and Panama.<sup>4</sup> English is growing as a second language in some parts of the Caribbean, such as the Dominican Republic and Puerto Rico, in northern Mexico, and in urban areas elsewhere in Latin America.

Dialects of Modern Spanish on the Iberian peninsula include Castilian, the northern dialect families of Navarro-Aragonese, Leonese and Asturian, and the southern, Andaluz dialects.<sup>5,6</sup> *Ladino* or *Judeo-español* is a dialect of Spanish spoken by Sephardic Jews expelled from Spain at the end of the fifteenth century. It is a “fossil” dialect in that it retains characteristics of the pronunciation of that time. In Latin America, the problem of defining dialect boundaries is a complex one.<sup>7</sup> The grammar is differentiated along phonological, morphological, syntactic and lexical lines, but the degree of variation makes classifying “discrete” dialect boundaries extremely difficult. Latin America is more conveniently described in terms of dialect “areas” which are associated loosely with general linguistic patterns. These include such areas as the River Plate region of Uruguay and Argentina, the Andean highlands, and the Caribbean. Section 1.7 below summarizes general patterns of syntactic variation in these areas.

### 1.1.2 *The spread of the Castilian dialect*

Although Spanish is spoken over an extremely broad geographical expanse, it is nevertheless relatively uniform syntactically. This is due in part

<sup>4</sup> In Belize, 25–40% of the population is Spanish-speaking, and most of the population speaks an English-based Creole (Kriol). The official language of Belize is English. Statistics on the occurrence of Creoles are based on Grimes (1988) and Kurian (1992).

<sup>5</sup> For detailed discussion of Iberian dialects see Alvar (1996), Otero (1971).

<sup>6</sup> Among Andaluz dialects, which are characterized by weakening of word-final *-s*, there are areas in which final *-s* appears to be disappearing. This (eventually) may have syntactic consequences with respect to the “richness” of features for number and person, since *-s* distinguishes plurality in nominals and distinguishes 2nd person in verbal paradigms.

<sup>7</sup> For detailed discussion of the problem of classification of Latin American dialects see Lipski (1994).

to the early political unification of Spain, and to the spread of the Castilian dialect throughout the unified area. This unification was a consequence of the drive to re-conquer the peninsula after its occupation by the Moors in the early eighth century. The area from which the reconquest was launched was Castilla la Vieja (Old Castille). In the course of the centuries-long battle against the Moors, the Castilian dialect spread throughout much of modern Spain. Castilian thereby coexisted with other Spanish dialects that had evolved in various areas, and largely replaced them over the course of time.

Most of Iberia had been Romanized during the period of the expansion of spoken Latin.<sup>8</sup> With the decline of Rome, the peninsula was invaded by successive waves of Germanic tribes, and eventually came under the control of Visigothic kingdoms during the fifth to eighth centuries. This period marks a transition during which spoken Latin was initially similar enough to the written form of Classical Latin to remain viable for administrative purposes.<sup>9</sup> Meanwhile the increasing political weakness of the Visigothic kingdoms and the beginnings of feudalism accelerated the growth of local Romance varieties. This was especially characteristic of northern and northwestern Iberia, where Romanization was never extensive, urbanization was minimal, and Romance coexisted with Basque, and perhaps other indigenous languages.

With the Moorish conquest, Iberia was for a time severed from the rest of Europe, where emerging monasteries provided a linguistic and cultural counterweight to feudal isolation. Throughout much of Iberia, Mozárabe<sup>10</sup> became the standard form of Romance. The mountainous north, however, which the Moors never successfully settled, retained its dialect diversity (Alatorre 1989:108). As Moorish control of the peninsula receded, the north and northwest became Christian strongholds with renewed ties to the rest of Europe. Santiago de Compostela was an important destination for Christians from throughout Europe, and monasteries and cathedrals emerged. At the

<sup>8</sup> Although spoken Latin was in use and undergoing evolution from much earlier times, the period of its great geographic expansion might be taken to begin around 100 BC, when Latin replaced Oscan as the official language of central Italy, to AD 200, when the empire reached its broadest expanse. Although Romanization of the Hispanic peninsula began earlier with the Second Punic War, the legionnaires (and colonizers) of this period were perhaps not predominantly Latin speakers. Lapesa (1981:94–101) notes that significant numbers may have been speakers of the Oscan–Umbrian subfamily of Italic, which was spoken in southern regions of Italy.

<sup>9</sup> The question of whether speakers considered their spoken and written languages to be one and the same has been debated in recent studies. For discussion and references see Wright (1991).

<sup>10</sup> The term “Mozárabe” refers either to Christians who lived in Moorish-controlled Spain, or to the variety of Spanish spoken by Christians (and non-Christians). See Galmés de Fuentes (1996).

4 The syntax of Spanish

southern periphery of Asturias (the then kingdom of Oviedo), a relatively unpopulated area known previously as Bardulia (Alvar 1994:81) had been newly settled and fortified with *castiellas* against Moslem raids. By the ninth century the area was known as “the place of the castles,” or Castille. According to Lloyd (1987:177), Castille was populated by settlers from different areas, who abandoned peculiar features of pronunciation associated with their origins. Castille was also an area where Basque was spoken, and some features of Spanish, such as initial *f* > *h* have been attributed to Basque influence.

Over the subsequent centuries, Castille became a dominant power in the north, and was the center from which the reconquest of the peninsula was launched. Although Castilian was not a prestige dialect, it gradually spread southward and became dominant as Spain was politically unified and Christianized.<sup>11</sup> The religious zealotry which followed the reconquest included linguistic “purification,” as Arabic books were burned in Granada, and the use of Arabic (and even Arabic borrowings) was increasingly condemned throughout the sixteenth century. Between 1609 and 1614, as many as 300,000 *moriscos* (non-assimilated or partially assimilated Moors and their descendants) were expelled from Spain.

The form of the language that took root in Latin America was affected by a number of unifying influences. One of these was the social climate of conformity – including linguistic conformity – which held sway in Spain at the time of colonization. This tendency was made concrete policy with respect to colonization, as the monarchs prohibited emigration of Jews and Moors to the new world (Sánchez-Albornoz 1984:15). Another factor that minimized diversity during the era of colonization was the relatively short time frame during which much of the settlement occurred. Immigration was most extensive before 1650, and dropped off sharply by the 1700s.<sup>12</sup> Colonization also coincided with the introduction of the printing press, the first of which was brought to Mexico City by the 1530s (Alatorre 1989:138). Subsequent influences, such as ongoing commerce with Spain, the independence movements, bilingualism and the growth of mass media, have resulted in a rich range of

<sup>11</sup> Lloyd (1987:179–180) suggests that “reverse prestige” may have enhanced the spread of Castilian, given the role of Castille in the liberation of the peninsula from the Moors. An additional factor in the spread of Castilian was migration. An economic breakdown in the north triggered significant migration from northern Castille to the south during the sixteenth century, which reinforced the spread of Castilian.

<sup>12</sup> Sánchez-Albornoz (1984:15–16) estimates that from 200,000 to 243,000 people immigrated during the sixteenth century, and an almost equal number during the first half of the seventeenth century. The extent of immigration is small overall, compared with immigration to the United States from other countries.

phonological and morphological variations in the grammar, but less variation in the syntax.

### 1.1.3 *The evolution of Spanish syntax*

The evolution of spoken Latin into proto-Romance was characterized from early on by simplification of inflectional paradigms for nouns, adjectives and verbs, and emergence or broader use of periphrastic constructions which fulfilled some of the same grammatical functions. The nominal case paradigms were reduced to a Nominative/Accusative distinction, and prepositions emerged as markers of other cases. Definite and indefinite articles evolved (from Latin demonstrative *ille* “that” and the cardinal *unum* “one,” respectively). Periphrastic comparative forms of adjectives replaced synthetic forms. In the verbal paradigms, simplification of Classical inflections included the loss of the future tense, of synthetic passives, and of diverse non-finite forms. Many of these changes were incipient or well underway in spoken Latin, and some were accelerated as a result of phonological changes such as loss of many word-final consonants and loss of distinctive vowel quantity. The most stable inflectional features were person, number and masculine/feminine gender markers, and the [±PAST] inflection for verbs.

The “break-up” of proto-Romance into the early differentiated Romance languages is generally dated from the point at which written Latin was no longer comprehensible to the Romance speaker, roughly between the fifth and ninth centuries.<sup>13</sup> Characteristics of early Spanish are deduced from documents dating from the eleventh century. Grammatical changes during this period continued those trends described above: inflectional simplification and grammaticalization of functional and quasi-functional morphemes; in many instances these changes were common across languages. For example, nouns lost their Nominative/Accusative distinction. In western varieties of Romance, accusative plural *-s* was reanalyzed as a plural marker. Object pronouns were de-stressed and became clitics. Verbal auxiliaries evolved in passives, compound perfect, future and conditional tenses. The clitic *se* (Latin 3rd.sg./pl. Refl.) was grammaticalized, first as a detransitive (anti-causative) morpheme, then as a marker of middles, and (in Spanish) as a marker of passive voice (Hanssen 1945:230–231).

<sup>13</sup> Because classical Latin was used as a written form under the Visigothic administrations, it is more difficult to date the transition from proto-Romance to Romance in the Iberian peninsula than elsewhere. In France, by contrast, “translations” began to occur in 813 (cf. note 8; see also Palmer 1954:178–179). Only in the eleventh century did Carolingian writing replace the Visigothic system (Lapesa 1981:169).

6 The syntax of Spanish

One syntactic innovation from this period is the emergence in Spanish of the “personal *a*,” a marker of specific, human direct objects. Personal *a* occurred most consistently at first with proper names and pronouns, less consistently with common nouns (Lapesa 1981:213). Torrego (1998:42; citing Lapesa 1968) mentions an additional factor which governed the distribution of personal *a* around the thirteenth and fourteenth centuries. *A* appeared with the complements of verbs that denote an action that affects an individual physically or psychologically. Only later did it occur with non-affected animate direct objects.

The constituent order of Old Spanish differs from that of Modern Spanish in several respects. In Old Spanish, only phrases headed by closed class items (such as articles, complementizers and prepositions) were clearly head-initial. Lexical, or “open class,” heads of phrases (nouns, adjectives and verbs) allowed both complement–head and head–complement order. The basic order of the verb and its objects is analyzed as having switched from OV to VO order (Otero 1975; Saltarelli 1994). It is interesting to note that auxiliary–main verb complexes gradually evolved from verb–auxiliary to auxiliary–verb (Rivero 1993; Lapesa 1981:217; Hanssen 1945:249, 251). The constituents of clauses also patterned differently in Old Spanish. Fontana (1993) argues that Old Spanish is a V2 (verb second) language, not of the German type (which exhibits second-position verbs in main clauses only), but of the Icelandic type: with verbs occupying second position in subordinate clauses also. Fontana terms this “symmetric V2.”

Another difference between Old Spanish and Modern Spanish concerns the behavior and the placement of pronominal clitics. Modern Spanish clitics attach only to verbs, and either precede or follow the verb according to whether the verb is finite or non-finite. Old Spanish pronominal clitics occupied second position in the clause, and were phonologically dependent on the preceding constituent – whether that constituent was a verb or not. This is shown by the fact that they could not occur clause-initially following a pause.<sup>14</sup> In this respect, the pronominal clitics behaved like other atonic elements, including *non* “not,” conjunctions and some auxiliaries. Auxiliaries mostly lost this restriction during the period of Old Spanish (cf. Hanssen 1945:251–252).

Old Spanish displayed auxiliary switch, similar to that of Modern French and Italian (Vincent 1982). Auxiliary *ser* “to be” alternated with *aver* “to have” in the compound perfect tenses. In these tenses, *ser* was generally used with unaccusatives and “reflexive” (anticausative) intransitives, and *aver* with

<sup>14</sup> For detailed discussion of the syntax of Old Spanish clitics see Rivero (1986, 1991), Wanner (1987), Fontana (1993).

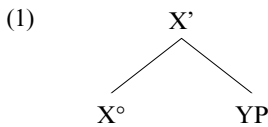
transitives (Lapesa 1981:212; Hanssen 1945:230–233). The compound perfect tense also displayed past participle agreement with the object. However, both auxiliary switch and past participle agreement were inconsistent.<sup>15</sup>

## 1.2 General characteristics of the syntax

Many characteristics of Spanish syntax are typical of the Indo-European family, including the relative richness of verbal morphology compared with nominal morphology, and the overt movement of interrogative phrases and of noun phrases (e.g., in passives). Other characteristics are prevalent within the Romance family. These include head-initial constituent order, pronominal clitics, negative concord, rich agreement morphology and null subject phenomena. Two characteristics of Spanish which are relatively isolated within Romance include the so-called “personal *a*” which precedes animate direct objects under certain conditions,<sup>16</sup> and clitic “doubling” of indirect objects (and dialectally, direct objects). This section summarizes features of Spanish syntax which place the language typologically, and which provide an introduction for subsequent discussion.

### 1.2.1 Constituent order

Modern Spanish is a head-initial language. As shown in (1), the construction of a phrasal head, or  $X^0$  with a complement, gives the order: head-complement. Thus, nouns, adjectives, verbs and prepositions precede their complements. Examples are in (2):



- (2) a. construyeron un puente [V<sup>0</sup> – NP]  
       built a bridge  
       “(they) built a bridge”

<sup>15</sup> Lapesa (1981:212) notes the inconsistent usage of *ser* and of past participle agreement, and notes that “contradictory uses” due to foreign influences were not uncommon.

<sup>16</sup> Lapesa (1981:94–101) observes that the use of “personal *a*” is one of several grammatical features which Iberian dialects share with Sicilian and other southern Italian varieties.

8 The syntax of Spanish

- |                         |           |
|-------------------------|-----------|
| b. con un martillo      | [P° – NP] |
| with a hammer           |           |
| c. estudiante de física | [N° – PP] |
| student of physics      |           |
| d. leal a los ideales   | [A° – PP] |
| loyal to the ideals     |           |

Functional categories also precede the lexical categories which they govern, for example determiners precede noun phrases, and complementizers precede clauses. Auxiliary verbs, which might be considered functional or quasi-functional items, also precede the main verb of the clause:

- (3) a. Habíamos hablado del problema.  
 had spoken of+the problem  
 “(We) had spoken about the problem.”  
 b. \*Hablado habíamos del problema.

The order of adjuncts, or optional modifying phrases, relative to the head varies according to several factors. All of the positions in (4) are possible with normal (unbroken) intonation:

- (4) [ (adjunct) head (adjunct) complement (adjunct) ]

Structurally complex adjuncts typically follow the head and its complements. Several factors condition the availability of pre-head adjuncts, including structural and lexical properties of the adjunct as well as the category of the head. Adjunct order is discussed in relation to the Noun Phrase (Chapter 2), the Verb Phrase (Chapter 4) and the clause (Chapter 5). The order of subjects is addressed below (1.3.) and in Chapter 5.

### 1.2.2 Case

Spanish has a Nominative/Accusative case system. Case is not manifested morphologically on lexical nouns or determiners; only personal pronouns and some relative pronouns retain vestiges of Latin case distinctions. The strong (i.e., tonic, or stressed) personal pronouns display morphologically distinct forms to the extent shown in (5), illustrated with the 1st person singular form:

- |                    |              |                    |
|--------------------|--------------|--------------------|
| (5) a. Nominative: | yo           | “I”                |
| b. Objective:      | mí           | “me”               |
| c. Genitive:       | mí(o/ a (s)) | “my” <sup>17</sup> |
|                    | (m./f.(pl.)) |                    |

<sup>17</sup> The strong forms of possessive pronouns agree in number and gender with the modified noun.

Objective Case in (5) is the form common to objects of prepositions. The weak pronouns (Section 1.2.4) may have different form and distribution depending on whether the object is direct or indirect. These differences lead to subclasses of Objective: (a) Accusative (direct object of V<sup>o</sup>), (b) Dative (indirect object of V<sup>o</sup>) and (c) Oblique (object of P<sup>o</sup>). The following discussion will briefly summarize the contexts for Nominative, Genitive and the three subcases of Objective case.

Nominative is the case of subjects of finite clauses, both indicative and subjunctive; of predicative NPs linked to the clausal subject; and of subjects of participial and infinitival adjunct clauses. The example in (6) illustrates that pronominal subjects of both indicative and subjunctive clauses appear in Nominative form:

- (6) Insisto                    **yo** en que lo hagas                    **tú**.  
 Insist-pr.ind.1st.sg. I on that it do-pr.subj.2nd.sg. you  
 “I insist that you do it.”

Predicative NPs with Nominative form are shown in (7):

- (7) a. El campeón eres **tú**.  
 “The champion is you(Nom.)”  
 b. Lo que encontraron era **yo**.  
 “What (they) found was I(Nom.)”

In (7), the verb agrees in person and number with the predicative pronoun (cf. English “It *is/am* I”).

Adjunct clauses with Nominative subjects are shown in (8):

- (8) a. [Llegada **ella**] empezó la fiesta.  
 arrived-f. she(Nom.) began the party  
 “(With) her arrived, the party began.”  
 b. [Habiendo llegado **ella**], empezó la fiesta.  
 have-prt. arrive-pprt. she(Nom.) begin-pret. the party  
 “With her having arrived, the party began.”  
 c. [Al cantar **tú**], empezó la fiesta  
 upon+the sing-inf+it you(Nom.) began the party  
 “Upon your singing it, the party began.”  
 d. [De ganar **ellos**] los felicitaremos.  
 of win-inf. they (Nom.) CL(DO) congratulate-fut.1st.pl.  
 “If they win, we will congratulate them.”

In the above constructions, the participle or infinitive must precede the subject, but some dialect variation occurs (see 1.7). The participial clause in (8a) shows number and gender agreement with the subject; the participial clause (8b) and infinitives (8c), (8d) are non-agreeing forms.<sup>18</sup>

<sup>18</sup> Rigau (1992) shows that constructions like (8c), which appear to be nominalized, are in fact clausal.

## 10 The syntax of Spanish

Genitive is the case assumed by the subject of a noun phrase, and is marked either by the preposition *de* with a non-pronominal, as in (9), or by the Genitive form of a pronominal, as in (10). Genitive pronominals have both weak (pre-nominal) and strong (post-nominal) forms, illustrated in (10a) and (10b) respectively:

- (9) el retrato de Josefina  
 the portrait of J.  
 “Josefina’s portrait”
- (10) a. mis libros  
 my-pl. book-m.pl.  
 “my books”  
 b. los libros míos  
 the-m.pl. book-m.pl. my-m.pl.  
 “my books”

In (9), the *de*-phrase is ambiguous between possessor, agent, and subject of the portrait. This illustrates that Genitives are not necessarily possessors, and also that *de* is not exclusively Genitive. The examples in (10) illustrate that Genitive pronominals agree in number (and gender) with the possessed noun. In contrast with Italian, determiners do not co-occur with a pre-nominal possessive (*\*los míos libros* “the my books”) in most dialects of Spanish. In contrast with English, “double genitives” of the form “a book of his” (*\*un libro de suyo*) do not occur. Post-nominal genitives show either *de*, as in (9), or genitive morphology, as in (10b).

Relative pronouns display a distinguishable Genitive form, although interrogatives do not. This is illustrated by the contrast between the relative pronoun in (11a) and the interrogatives in (11b, c):<sup>19</sup>

<sup>19</sup> Interrogative forms do not show case distinctions in general. *Qué* “what,” and *quién* “who,” for example, serve as both Nominative and Accusative arguments:

- (i) a. ¿Qué pesa 7 kilos?  
 What(Nom.) weigh-pr.3rd.sg. 7 kilos  
 “What weighs 7 kilos?”  
 b. ¿Qué dijo Susana?  
 what(Acc.) said Susana?  
 “What did Susana say?”
- (ii) a. ¿Quién trabaja aquí?  
 “Who works here?”  
 b. ¿(A) quién buscan?  
 PA who(Acc.) look-for  
 “Who are they looking for?”

The case of non-Nominative interrogatives is marked by prepositions, including personal *a*, as in (iib).