By contrast with phonological, morphological and syntactical systems, the lexical systems of language are inherently open-ended, and it is not intended here to attempt an exhaustive treatment of the Spanish lexis (such a treatment is, in any case, by definition impossible) or to account in detail for the cultural conditions which were responsible for borrowings from various sources. Partial discussion of these matters will be found in the various sections of the Introduction (1.1–5); in the sections that follow, the intention is to examine in outline the main sources of the vocabulary of Spanish, using a minimum of exemplification.

4.1 Vocabulary inherited from Latin

The core vocabulary of Spanish, including many hundreds of the most frequent words, as well as many less frequent items, has descended from spoken Latin, passed on orally in unbroken succession from generation to generation, and undergoing the various phonological changes detailed in chapter 2. Such words have already been defined (see 2.2.1) as popular words.

However, it should be noted that semi-learned words (see 2.2.3), over which there has been much controversy, are here regarded also as orally inherited, differing from popular words only to the extent that semi-learned words have undergone one or more modifications of form due to the influence exerted upon them by the pronunciation with which related Latin words were read aloud at various periods, either as part of the offices of the Church or in legal/administrative circles.

4.2 Words of pre-Roman origin

As the use of Latin spread across the Peninsula, in the centuries following the first Roman involvement in Spain (218 bc), Latin came to be used bilingually with other, pre-existing, languages. Since the linguistic map of pre-Roman Spain was complex, the nature of this bilingualism differed from area to area, but it is evident that the conditions existed for the borrowing of lexical items by the Latin of Spain from a variety of other languages. Such borrowing does
not appear to have been extensive, since the languages spoken bilingually with Latin mostly enjoyed low prestige and no doubt exhibited the more restricted vocabulary associated with the less-developed cultures they served. However, in some instances there would be no Latin term to express some concept (usually related to local flora or fauna, to local life-styles or techniques) and the remedy was to borrow the local term for the concept. On other occasions, despite the prior existence of a Latin word, the borrowing of a local word (to replace its Latin counterpart) could remedy some Latin lexical defect. Such a defect might be that the Latin term had acquired in some contexts a negative value (e.g. sinister ‘sinister’); borrowing of a non-Latin word (in this case, the Basque word which gives rise to izquierdo), to replace the Latin term in its primary sense (‘left’), solved a problem of potential ambiguity.

Celtic was widely spoken in central and western parts of the Peninsula in pre-Roman and Roman times and has provided a number of loans to the Latin of Spain. The following can be included among such borrowings, with varying degrees of certainty: álamo ‘poplar’, berro ‘watercress’, bota ‘leather wine-bottle’, brezo ‘heather’, brío ‘verve’, engorar ‘to addle’, gancho ‘hook’, greña ‘(greasy) lock of hair’, lama ‘silt’, légamo ‘slime’, losa ‘flagstone’, serna ‘ploughed field’. However, rather more words of Celtic origin were borrowed outside the Peninsula (mostly from the Gaulish speech of France) and became part of the word-stock of popular Latin wherever it was used, including Spain. Spanish words arguably inherited in this way from Celtic have cognates in other Romance languages and include abedul ‘birch’, alondra ‘lark’, arpende ‘unit of land-measurement’, braga ‘breeches’, cabaña ‘shack’, camino ‘road’, path, camisa ‘shirt’, carpintero ‘carpenter’, carro ‘cart’, cereza ‘beer’, legua ‘league’, saya ‘skirt’, vasallo ‘vassal’.

Basque provides a number of borrowings in Spanish, many of which were no doubt introduced into the Latin of Spain in the period following the Roman conquest of the northern Peninsula. However, since Basque, alone among the pre-Roman languages of Spain, has continued to be spoken down to the present, this language has at all times constituted a potential source of borrowing. Especially frequent are personal names (e.g. García, Íñigo, Javier, Gimeno, Sancho, although the last may be from Lat. sanctus; see Lapesa 1988), but a certain number of other words (especially nouns) have often been claimed to have Basque origin, including: aquellarre ‘witches’ sabbath’, boina ‘beret’, (caer) de bruces ‘(to fall) headlong’, cachorro ‘whelp’, cencerro ‘animal bell’, chaparro ‘dwarf oak’, izquierdo ‘left’, laya ‘spade’, legaña ‘rheum’, narria ‘sledge’, pizarra ‘slate’, socarrar ‘to scorch’, urraca ‘magpie’, zuro ‘left-handed’. However, it should be noted that R. L. Trask (1997a: 415–21) rejects this origin for all these (and other) words, except izquierdo.

It is probably therefore wise to include most of the words just mentioned among a somewhat larger group of Spanish terms which in all probability were borrowed from unidentified pre-Roman sources (in some cases, perhaps Basque). Cognates of these words are sometimes to be found in Portuguese and/or Gascon, but they do not appear to be related to any known Latin or Celtic word. The following may be included (with varying degrees of certainty) in this portion of the Spanish vocabulary: abarca ‘sandal’, aliso ‘alder’, alud ‘avalanche’, arroyo ‘stream’, ascua ‘live coal’, balza ‘pool’, barro ‘clay’, mud’, becerro ‘calf’, bruja ‘witch’, camin ‘bed’, chamorro ‘close-cropped’, charco ‘puddle’, garganta ‘tickle’, gazo ‘young rabbit’, gusano ‘maggot’, caterpillar, madroño ‘strawberry tree’, manteca ‘lard’, navia ‘marshy valley’, páramo ‘moor’, peseta ‘eyelash’, sapo ‘toad’, sarna ‘scabies’, sarro ‘plaque (on teeth)’, vega ‘river-plain’, zarza ‘bramble’.

4.3 Latinisms

From the Middle Ages onwards, Spanish has made lexical borrowings from Latin; this element of the Spanish vocabulary is often referred to as that of ‘learned words’ (see 2.2.2), although some scholars do not equate the two terms, and define ‘learned’ in terms of stylistic value, whatever the source of the words concerned. Latinisms have been transferred to Spanish from written Latin sources, and with minimal change (usually limited to the final syllable, to fit the Latinism to the morphological patterns of Spanish). The requirement which Spanish has continually had for new vocabulary (chiefly, but by no means exclusively, relating to non-material aspects of life) could often be met by borrowing from Latin, whether from the Latin of the Church, the law or administration, or, especially from the Renaissance onwards, from Classical Latin sources. Because of the prestige associated with the Latin language, over the centuries and still today, the Latin lexicon has usually been the first source to which Spanish speakers and writers have turned to provide labels for new concepts.

There is no period in which Latinisms have not been introduced abundantly into written Spanish, and although a proportion of those introduced were subsequently abandoned, it has been calculated (Alvar and Mariner 1967: 21–2) that such loans comprise between 20 and 30 per cent of the vocabulary of Modern Spanish (although, if the calculation is made on the basis of the frequency of words in Spanish, the proportion is much smaller).

Before the development of reasonably consistent vernacular spelling (in the later twelfth century), it is impossible to distinguish Latinisms from archaically spelt popular or semi-learned words. Thus, an early spelling like desiderio (Silos gloss 132: see e.g. Menéndez Pidal 1964a: 16) may be interpreted as a rendering of a borrowing from Latin pronounced /desiderio/ (which failed to become established or later fell out of use) or (more probably) of a noun pronounced desidéi (derived from the verb desear < DESIDIARE, for DESIDERARE). However, from the early thirteenth century onwards, following the adoption of
spelling norms based upon phonological principles (Wright 1982), latinisms can be clearly distinguished from inherited vocabulary, so that, for example, we are safe in interpreting vision (Disputa del alma y el cuerpo, 4) as a learned borrowing pronounced [biˈz̞oːn] (< vīsīō, vīsīōnis), later pronounced [biˈson]. And from this point on, borrowings from Latin can be identified in every period, in greater or lesser quantities according to the cultural climate. In periods of literary development in which translation from Latin or adherence to Latin literary models is frequent (as in the later thirteenth century, the fifteenth century, much of the Golden Age and the eighteenth century), latinisms are numerous. Similarly, borrowing from Latin has satisfied part of the requirement for new scientific and commercial terminology, especially in the nineteenth and twentieth centuries, although a substantial proportion of the latter borrowings were probably not made directly from Latin, but via other modern European languages.

A selection of Latin borrowings made by Spanish at various periods can be made as follows:


In the case of the latinisms of the nineteenth and twentieth centuries, it is often the case that the words concerned entered Spanish through the medium of some other modern language, at first usually French, now usually English. Such words should therefore be regarded strictly as gallicisms or anglicisms. At all events, we are dealing here with forms it would be pointless to list, since they have cognate forms (usually with identical meaning) in a variety of other languages.

For the problem of phonological adaptation posed by some latinisms, and the resolution of this problem in the Golden Age, see 2.6.5. For the definition and general discussion of latinisms in Spanish, see Clavería Nadal (1991), and the introduction to Castro (1936). For latinisms introduced up to and including the sixteenth century, see Herrero Ingelmo (1994–5).

### 4.4 Hellenisms

Apart from a handful of place-names, the language spoken in the Greek settlements apart the east coast of the Peninsula in pre-Roman (and possibly in Roman) times has left no legacy in Spanish. All the hellenisms in Spanish found their way into the language as a result either of being first incorporated into the Latin of Rome or of being borrowed directly from literary Greek. Three separate strands of such Greek borrowings can be recognized in Spanish.
Firstly, popular Latin borrowed a considerable number of words from Greek, owing to contact between Greek speakers and Latin speakers at all levels of Roman society and over a period of several centuries. Such borrowings became part of the popular Latin word-stock and in many cases were inherited orally by Spanish (that is to say that, once borrowed by Latin, such words are treated identically to words of native Latin origin and in the process of their inheritance they undergo all the changes which affect orally inherited Latin words, although a number have been subjected to semi-learned remodelling, again in the same way as many inherited Latin words). The hellenisms of this first stratum typically refer to aspects of everyday life, and include names of tools and domestic items (ampolla ‘flask’, ancla ‘anchor’, baño ‘bath’, cesta ‘basket’, cuévano ‘basket (carried on the back)’, cuchar’u ‘spoon’, cuerda ‘rope’, espada ‘sword’, espuela ‘(non-rigid) basket’, estopa ‘tow’, lámpara ‘lamp’, linterna ‘lantern’, sábana ‘sheet’, saco ‘bag’), items referring to the house and to building (bodega ‘wine-cellar’, cal ‘lime’, cámara ‘chamber’, mármol ‘marble’, piedra ‘stone’, plaza ‘square’, torre ‘tower’, yeso ‘plaster’), terms relating to the land (greda ‘chalk’, yermo ‘wasteland’), botanical names (ajenjo ‘absinth’, caña ‘cane’, cánamo ‘hemp’, cereza ‘cherry’, cima ‘tree’ (top), cizaña ‘darnel’, codeso ‘laburnum’, espárrago ‘asparagus’, esparto ‘esparto’, olivo ‘olive tree’, rábano ‘radish’, regaliz ‘licorice’, tallo ‘stem (of plant, etc.)’, animal names (concha ‘shell’, esponja ‘sponge’, morena ‘moray (eel)’, ostra ‘oyster’, perdiz ‘partridge’, pulpo ‘octopus’, púrpura ‘purple (originally extracted from the murex)’), words related to man (golpe ‘blow’, huérfano ‘orphan’, lágrima ‘tear’, pena ‘grief’, talento ‘talent’, tío ‘uncle, aunt’), as well as a small number of grammatical particles (cada ‘each’ (see 3.5.5), OSp. maguer ‘although’ (see 3.8.2)).

Secondly, since the language of the Church, even in the Western Empire, was at first Greek, ecclesiastical Latin was full of hellenisms, many of which have passed into Spanish. However, since ecclesiastical Latin (with its Greek lexical component) continued to be heard in church down to the twentieth century, vernacular Spanish words of Greek origin relating to Christianity and the Church (like their purely Latin-derived counterparts) were especially open to latinizing influences, so that, on the one hand, much of the vocabulary concerned reveals semi-learned transmission, while, on the other, the other, ecclesiastical hellenisms are fully learned. Words of this group include: abismo ‘abyss’, bautismo ‘baptism’, bautizar ‘baptize’, biblia ‘bible’, blasfemar ‘blaspheme’ (and, via a modified Latin form of the same hellenism, lastimar ‘harm, pity’), canónico ‘canon’, cátedra (‘bishop’s) chair’, later also ‘university chair’ (contrast the early borrowing cadera ‘hip’, from the same etymon; see 5.2.2.), catedral ‘cathedral’, católico ‘catholic’, celo ‘zeal’, cementerio ‘cemetery’, cisma ‘schism’, clérigo ‘cleric’, coro ‘choir’, chorus, diablo ‘devil’, diácono ‘deacon’, ermita ‘hermitage’, himno ‘hymn’, iglesia ‘church’, lego ‘layman’, limosna ‘alms’, mártir ‘martyr’, misterio ‘mystery’, monaguillo ‘altar boy’, monasterio ‘monastery’, palabra originally ‘parable’, then ‘word’, papa ‘pope’, paraíso ‘paradise’, parroquia ‘parish (church)’, Pascua ‘Easter’, Christmas, patriarca ‘patriarch’, profeta ‘prophet’, salmo ‘psalm’.

Thirdly, Greek has served, throughout the history of Spanish, as a source of technical and scientific vocabulary. Most such words are ones which have first passed into Latin, as Latin expanded its own lexical resources, and should therefore be regarded as a subset of the learned words discussed in 4.3. However, a number of words from this stratum (especially medical terms) passed sufficiently early into everyday speech to undergo the phonological changes typical of the vernacular and must therefore be regarded as popular words, while other hellenisms of this stratum which passed into the vernacular were subject to partial remodelling (particularly from the subject of the humanists onwards) and therefore constitute examples of semi-learned transmission.

As in the case of latinisms, a small selection of technical and scientific borrowings from Greek is given, organized by approximate period of adoption.


The words borrowed from Greek during the late sixteenth and seventeenth centuries are almost all international words (and in many cases probably reached Latin from other modern European languages, rather than directly from Greek or Latin). They include a number of words which combine Greek lexemes in ways not observed in Greek. Only a small selection of the many recent hellenisms is being used: anemia ‘anemia’, anestesia ‘anaesthesia’, clínico ‘clinical’, neumónia ‘pneumonia’, psiquiatría ‘psychiatry’, quirofano ‘operating theatre’, quiste ‘cyst’, raquitismo ‘rickets’; autógrafo ‘autograph manuscript’, biografía ‘biography’, fonética ‘phonetics’, taquigrafo ‘stenographer’; asteroide ‘asteroid’, cosmos ‘cosmos’, cráter ‘crater’, sísmo ‘earthquake’; arcaico ‘archaic’, arqueología ‘archaeology’, laico (beside lego, see above) ‘lay’, programa ‘program’.

For Greek words which passed to Spanish via Arabic, see 4.6, and, for further discussion of hellenisms in general, see Fernández Galiano (1967) and Eseverri Hualde (1945).
helmet", *hrapôn > rapar ‘to crop (hair, etc.)’, *spaiha > espía ‘spy’. Similarly, since spoken Latin had no syllable-initial [w] (early [w], spell v, had rapidly become [β] or [v]; see 2.5.3.1), Germanic /w/ was replaced by [gw] (which was familiar in traditional Latin words like lingua, by this stage pronounced [ˈlɛŋɡwa]). This adapted pronunciation has survived in Spanish where the following vowel is /a/, but where a front vowel followed, [gw] was later reduced to /j/, although the same spelling (gu-) continues to be used with both values: waitheir (+ suffix) > guadaña ‘scythe’, vardón > guardar ‘to guard, put away’; werra > guerra ‘war’, visa > Osprey guisa ‘manner, way’. Latin and its successors had no /θ/ (Spanish /θ/ arises only in the seventeenth century; see 2.6.2), and in Germanic words containing this phoneme, /θ/ is replaced by /h/: thriska > triscar ‘to gambol’, *θaðhys > late Latin taxō, *onis > tejón ‘badger’. Similarly, it seems likely that the Germanic intervocalic voiceless plosives differed from their nearest Latin counterparts, perhaps in that the Germanic phonemes were aspirated (like English initial /p/, /t/, /k/: [pʰ], [tʰ], [kʰ]), and were sometimes equated with Latin /pp/, /tt/, /kk/, rather than with /p/, /t/, /k/. As a consequence (see 2.5.3.2), Germanic intervocalic /p/, /t/, /k/ often appear in Spanish as /p/, /t/, /k/, rather than as /β/, /v/, /β/, *raupa >ropa ‘clothes’, spitu > espetu ‘(roasting) spit’, reiks (whence *ricus) > rico ‘rich’.

4.6 Arabisms

From the eighth to the fifteenth century, Arabic was the official language of a considerable (but eventually shrinking) portion of the Peninsula, and within this area (i.e. within Al-Andalus) all inhabitants would have at least some familiarity with that language, whether as native speakers or as second-language users. However, the multilingual nature of Al-Andalus cannot by itself explain the host of arabisms incorporated into the Spanish lexicon, since the forms of Romance (the Mozarabic dialects) spoken bilingually with Arabic in Islamic Spain were, of course, not forms of Castilian, but independent descendants of spoken Latin, which became extinct in the later Middle Ages. The Castilian dialect, from which standard Spanish descends, originated outside Al-Andalus, so that the loans it made from Arabic (the large majority of which were made in the period up to the tenth century, before the substantial expansion of Castile and its dialect into Arabic-speaking territory) were loans made from a neighbour rather than from a language spoken in the same territory. The reasons for the heavy borrowing of Arabic words by Castilian must therefore be sought in factors other than extensive bilingualism, and can probably be reduced to two: firstly, the need for names applicable to the many new concepts (both material and non-material) which reached Castile from Al-Andalus, and which it was most convenient to name by means of words borrowed from the dominant
language used in that area, and secondly, the very high prestige associated with Arabic in the early Middle Ages, owing to the fact that Arabic was the vehicle of a culture which was considerably more advanced than that of Christian Spain, and indeed than that of the rest of Christian Europe. The first of these factors is responsible for the numerous additions to the Spanish vocabulary from Arabic sources, while the second factor is responsible for the less frequent replacement of existing Castilian words by Arabic synonyms. However, other factors must be borne in mind in conjunction with those just mentioned. On the one hand, even in the period up to the tenth century, there was a certain influx into Castile (as into other Christian territories) of southern Christians (Mozarabs) already familiar with Arabic and perhaps speakers of it. On the other hand, from the tenth century on, through the southward expansion of Castile and the movement of Castilian-speaking population into newly reconquered Arabic-speaking territory, there would have come into existence for the first time a certain body of bilingual Castilian–Arabic speakers, who would have been responsible for the borrowing of at least some Arabic lexical items.

A very high proportion of the arabisms in Spanish are nouns, an even higher proportion than that observable in other cases of heavy inter-language borrowing, where nouns normally predominate. These loans very frequently begin with the syllable al-, owing to the fact that the Arabic definite article al- (which was invariable for gender and number) was interpreted by speakers of Romance (where the definite article varied in form) as an integral part of the noun and therefore borrowed together with the noun it accompanied. The Spanish vocabulary contains many hundreds of arabisms, some of which are among the most frequently used words of the language, spread across most semantic fields. The semantic fields selected for discussion here are those in which borrowing from Arabic was particularly frequent, usually because of the large number of new concepts, introduced to Spanish speakers, which belonged to those fields.


Civil life too was affected by new arrangements, borrowed from Moorish Spain together with the necessary vocabulary: alcáde ‘mayor’, aldea ‘village’, algucil ‘bailiff’, almácén ‘warehouse, department store’, arrabal ‘suburb’, barrio ‘district (of town)’.


Words referring to the natural world which were borrowed by Castilian from Arabic are not numerous, but a small number of such terms has remained frequent: alacrán ‘scorpion’, alcaraván ‘stone curlew’, alcatraz ‘gannet’, bellota ‘acorn’, garra ‘claw’, jabali ‘wild boar’.
A number of new foods found their way into the Spanish diet, bringing their Arabic names: álbanígda 'meat ball', alfeñique 'sugar paste', alibmar 'syrup', fídeos 'vermicelli', jarabé 'syrup', mazapán 'marzipan'.

Because Arabic science, during much of the Middle Ages, was considerably more advanced than that of Christian Spain (and indeed of Europe generally), Castilian speakers acquired almost the whole of their scientific vocabulary from Arabic. A good number of these medieval scientific terms persist in Spanish: alambeique 'retort', alcantar 'camphor', alcoho 'alcohol', algeba 'algebra', almanaque 'almanac', alquímia 'alchemy', azogue 'quick-silver', cénit 'zenith', cero 'zero' (via Italian; see 4.1.3), ciftra, at first 'zero', now 'figure', nadir 'nadir'.

It has already been said that arabisms have penetrated almost all fields of the Spanish lexicon. In addition to those considered above, the following terms may serve to illustrate the degree of penetration of the Spanish vocabulary by Arabic and the persistence there of the Arabic-derived words: ajedrez 'chess' (and the related alfíl 'bishop'), albornoz 'bathrobe', alcurnia 'ancestry', alhaja 'jewel', alquitrán 'tar', anil 'indigo', asesino 'murderer', ataúd 'coffin', azafata 'stewardess', azul 'blue', carmesí 'crimson', dado (?) 'dice', fonda (?) 'inn', fulano 'what's-his-name', gandul 'good-for-nothing', hazaña 'feat', horobo 'hunchback', marfil 'ivory', melena 'mop (of hair)', mezquino 'wretched; pauper', mengano 'what's-his-name', mezquita 'mosque' (together with alminar 'minaret', almúdejano 'muezzin'), nucia 'nape', ola 'wave', ojáli 'if only!', recamar 'to embroder (in relief)', tabaco (?) 'tobacco', taza 'task', zagal 'youth'.

With the decline in prestige of Arabic culture in the late Middle Ages and in the Spanish Golden Age, some losses did occur from among the arabisms of Castilian. In a number of cases, an arabism was replaced by a borrowing from a source at that stage perceived as more prestigious, or by a term created by means of the derivational resources of Spanish. Thus albeitear gave way to veterinario (from Latin), alfajeme 'barber' disappeared in the face of barbero (a derivative of pre-existing barba 'beard'), alfayate 'tailor' was replaced by sastre (from Occitan), alarife was displaced by arquitecto (from Greek, via Latin).

Not all of the arabisms which appear in Spanish are borrowings of words which belong to the native Arabic word-stock. Because Arabic, after the seventh century, was the language of a cultural zone which stretched from India to the Atlantic, it was in contact with many other languages, some of which had great prestige, and from which Arabic consequently borrowed numbers of words. Such words were in some cases passed on to Spanish (and to other European languages). Examples of such complex transmission include Sp. ajedrez and alcantar (from Sanskrit), alfalfa, alfeñique, almibar, aníl, azul, jazmín, naranja (from Persian), and acelga, adarme, alambeique, alquímia, arroz (from Greek). In addition, before its expansion out of the Arabian peninsula, Arabic came into contact with Latin, sometimes directly, sometimes via Greek, and borrowed from that source a number of words which were subsequently passed to Spanish. These include albaricoque (Lat. praeccou 'early peach'), alberchigo (Lat. persicu, whence also Sp. prisco 'apricot', by normal descent, alcázar (Lat. castri, from a diminutive form of which, castelli, descends castillo, by direct transmission), almad (Lat. modiu) 'measure for grain'.

The incorporation of arabisms into the Spanish vocabulary posed considerable problems of phonemic adaptation (unlike the incorporation of Germanic loans; see 4.5). At the time of intense borrowing from Arabic (eighth to tenth centuries), and later, Arabic contained a fair number of phonemes with no near equivalent in Romance. These phonemes were for the most part velar and laryngeal consonants, but the Arabic dental fricatives and Arabic /l/ were also problematical for Romance speakers. In addition, there were problems of distribution (phonemes similar to those of Romance but in unfamiliar positions in the word) to be overcome.

The Arabic velars and laryngeals were occasionally replaced by Castilian /h/ (spelt j), as in hinna > alfeña (later almíña) 'privet, henna', hanbul > OSp. alfamor, later almamor 'carpet'. However, the Spanish velars /h/ and /l/ were also used as replacements for the 'difficult' Arabic phonemes: arabiya > algarabia 'hubbub', mandh > almanaque 'almanac', sáix > OSp. seque, MSp. jque 'sheikh', harsfaja > alcahófa 'artihoche', harríba > algarroba 'carob bean'. A further solution was to omit the Arabic phoneme: áqar > alcarán 'scorpion', arif > alarife 'architect', ard > alarde 'display', ázár > alazán 'chestnut-(coloured)', húyal > ola 'wave', zára > tara 'task'. Where /l/ appears in Modern Spanish in correspondence with an Arabic velar or laryngeal (e.g. horz > alforja 'saddlebag'), it is probably the case that the Arabic word was first borrowed by a non-Castilian variety of Romance (where no /h/ was available as a replacement for the problematic phoneme) and that the word was then passed to Castilian at a relatively late date (see Penny 1990b).

The Arabic dental fricatives represented by s (sin) and z (zayn) (both unvelarized) and by š (sad) (velarized) did not correspond closely to Romance /s/ and /z/, which were apico-alveolar. The Arabic phonemes were consequently replaced by the nearest Romance dental phonemes: voiceless š and z (by the Old Spanish affricate /h/) and voiced z by the affricate /d/. These phonemes were spelled še- and z respectively in the medieval period, and merged to provide /th/ in the modern language (see 2.6.2): seňka > ceca 'mint', šftr > ciftra 'figure', sufanariya > OSp. çahanoria > MSp. zanahoria 'carrot', saça > zaga 'rearguard'. A special problem arose in the case of the Arabic group /g/ for the sibilant was first replaced, as we have just seen, by Romance /h/, and then the resulting group /th/ was simplified to /t/ (by since the seventeenth century): ustowán > OSp. açagudn > MSp. zaguán 'vestibule', mustarib > moçarabe > mozarabe 'Mozarab'.

4.6 Arabisms

269
4.7 Mozarabisms

Castilian displays a number of borrowings from Mozarabic, the vernacular speech of Christians (but also of many Muslims and Jews) in Al-Andalus, that is, in those areas of medieval Spain which were under Islamic rule. With reference to language history, the term ‘Mozarabic’ indicates a series of descendents of Latin, spoken until at least the thirteenth century (and perhaps as late as the fifteenth in Andalusia) in the southern two-thirds of the Peninsula. These southern Hispano-Romance varieties were eventually replaced by Catalan, Castilian and Portuguese, as the latter, northern, varieties of Hispano-Romance expanded in the wake of the Christian Reconquest of Islamic Spain. In New Castile, Murcia and Andalusia, Castilian was spoken alongside Mozarabic, no doubt often by the same individuals, for considerable periods of time after the conquest of each city. It is known, for example, that in the case of Toledo, reconquered in 1085, Mozarabic still enjoyed some use in the early thirteenth century (see González Palencia 1926–30, Galmés 1983), in part because its speakers were often more cultured than the incoming Castilian speakers, owing to their ancestors’ participation in the culture of Al-Andalus, which until the eleventh century was more culturally and economically developed than the Christian north. Until the creation of literary Castilian (beginning in the late twelfth century) and until Castilian came to be used as a national language of administration (in the later thirteenth century), Mozarabic no doubt enjoyed considerable social prestige in reconquered areas and was therefore in a position to exert influence upon Castilian. This influence, as we have seen (4.6), often resulted in the transmission to Castilian of arabsisms previously adopted by Mozarabic, but also accounts for the borrowing by Castilian of certain Mozarabic words of Latin origin (in some cases, perhaps, replacing traditional Castilian forms). Among such borrowings, words referring to agriculture and the living world were particularly frequent, as can be seen from the following selection of probable mozarabisms in Spanish: cagarruta (animal dropping), campiña (area of cultivated land), camélido (jag, bucked of chain-pump, etc.), capacho (shopping basket), capuz (hood, hooded cloak), corcho (cork), chiclara (pea) (in Andalusia, Galicia, Cuba,
4.8 Gallicisms and ocismantias

Although there are arguably a few gallicisms which passed into Spanish in the early Middle Ages, almost all have entered the language since the eleventh century. The reasons for the very frequent medieval borrowing from French and Occitan lie, naturally, in the cultural importance of northern and southern France in the later Middle Ages, and can be seen to be related, in particular, to four spheres of life, political, religious, literary, and commercial.

French involvement in the Reconquest and in the settlement of reconquered areas is well known, as is their role (particularly that of the monks of Cluny and Citeaux) in monastic and other religious reform and in the pilgrimage to Santiago (the majority of pilgrims were French and many French men and women were permanently established along the pilgrim routes). The indebtedness of Medieval Spanish literature to French and Occitan writing is well established, as is the growing commercial importance of France from the late Middle Ages. Many of these general motives for borrowing have persisted into the modern period (although religious influence has perhaps been slight in recent centuries).

In some cases, it is difficult to establish whether a borrowing has been taken from French or from Occitan, and this will not be attempted here. However it should be noted that ocismantias are almost entirely restricted to the medieval period, since before the end of that period Occitan culture was in severe decline, dominated by that of northern France. For Germanic words passed to Spanish via French, see 4.5.

In the following discussion, mention will be made only of words which have survived to the present (sometimes with changed meaning); we shall pass over the many borrowings, some of them very frequent in their day, which have become obsolete.

Contacts with territories north of the Pyrenees became increasingly important from the eleventh century, owing to the greater political stability of the Christian kingdoms of Spain. This increased stability attracted French and Occitan immigrants both of a temporary kind, such as pilgrims, and of a more permanent kind, like monastic reformers and settlers of newly reconquered territories. In this first period of frequent borrowings from French and Occitan (eleventh to thirteenth centuries), we observe the most remarkable loan of all, español 'Spanish', replacing native español. Borrowed military terms, reflecting French participation in the Reconquest, include aliar 'to ally', blandir 'to brandish', corcel 'steed', dardo 'sword', esgrimir 'to wield', estandarte 'banner', flecha 'arrow', galopar 'to gallop', maestre 'master (originally of a chivalric order)', mallat 'chainmail', trotar 'to trot', together with the more general emplear 'to use'. Religious terminology is also well represented: capellán 'chaplain', capitel 'capital (of a column)', deán 'dean', fraile 'frar, monk', hereje 'heretic', hostal 'hostelry (originally a religious foundation)', preste 'priest' (archaic). Terms related to the feudal system and to a leisured lifestyle are relatively frequent: bachiiler 'holder of certificate of secondary education' (originally 'young knight'), doncel 'squire', doncella 'maiden', duque 'duke', homenaje 'homage', linaje 'lineage'; bailar 'to dance', danzar 'to perform a dance', rima 'rhyme', trobador 'poet, troubadour', vihuela 'viol', deleíte 'pleasure', vergel 'orchard'; joya 'jewel', granate 'garnet', estuche 'small box, case', cuscabel 'small bell (on clothing, a hawk's leg, etc.)', polaina 'gaiter', palafreñ 'palfrey'. A number of words refer to the household and to food: arenuque 'herring', jamón 'ham', jengibre 'ginger', manjar 'food', vianda 'food'; antorcha 'torch', chimenea 'chimney', jaula 'cage', mecha 'wick'. A few terms belong to the natural world: bayá 'berry', laurel 'laurel', papagayo 'parrot', raiseñor 'nightingale'. Others, including abstracts, cover a wide field: desdén 'contempt', desnayar 'to faint', enojar 'to annoy', esquina 'handbell, cattlebell', gris 'grey', jornada 'period of a day', jornal 'day's wage', ligero 'light (in weight)', mensage 'message', tacha 'blemish'.

In the fourteenth and fifteenth centuries, borrowing from French and Occitan had passed its apogee but was still reasonably frequent. In this period we continue to find military and related (including naval) terms: baluarte 'bastion', botín 'booty', heraldo 'herald', pabellón 'pavilion; flag'; amarrar 'to moor', cable 'cable', quilla 'keel'. The courtly life continues to be reflected, for example in: dama 'lady', paje 'pageboy', gala 'elegant dress; elegance', galán 'gallant', jardín 'garden', patio 'courtyard'; balada 'ballade', chimiría 'shawm', flauta 'flute', refrán 'refrain; proverb'. The material and natural worlds are reflected in: cordel 'cord, rope', correo 'courier; post', despachar 'to dispatch, settle', forjar 'to forge', maleta 'suitcase', perfíl 'outline, profile', pinzas 'pincers', trinchar 'to carve'; avestruz 'ostrich', fajáin 'pheasant', salvaje 'wild'. Other borrowings of this period include: adite 'furthing', bundel 'brothel', desastre 'disaster', embajada 'mission; embassy', jerigónza 'gibberish', lisonja 'flattery', parlar 'to chat'.

In the Golden Age, the conflict between France and Spain led to many borrowings in the military and naval spheres, including: arcabuz 'arquebus', asambleta 'assembly', barricada 'barricade', batallón 'battalion', batería 'battery', bayoneta 'bayonet', brecha 'breach', calibre 'calibre', carabina 'carbine', cartucho 'cartridge', coronel 'colonel', jefe 'head, leader', marchar 'to march', piquete


For further details of borrowing from French and Occitan, see Colón (1967a), Lapesa (1980), Pottier (1967).

### 4.9 Amerindianisms

The first European contact with the New World (as a result of Columbus’s voyages of discovery) took place in the West Indies and the major West Indian islands became an indispensable staging-post for the later conquest of the northern and southern continents. The first contact of Spanish with Amerindian languages was thus with the languages of the Caribbean (Carib and Arawak, the latter including Taíno, the variety of Arawak spoken in the major Caribbean islands), followed by contact with the main language of Mexico (Nahuatl) and
that of the Inca Empire (Quechua). The majority of loans to Spanish were made from these languages, although numbers of borrowings were made (and continue to be made) from other Amerindian sources, including Maya (southern Mexico and the northern isthmus), Chibcha (Ecuador, Colombia and the southern isthmus), Tupi-Guaraní (the major river-basins of the southern continent, including Paraguay), Araucanian or Mapuche (central Chile and the Pampa area of Argentina). However, very few words borrowed from languages other than Carib, Arawak, Nahuatl and Quechua have become universal in Spanish or have even become general in American Spanish.

There follows a selection of words which Spanish has borrowed from a variety of Amerindian sources. The words selected are those which show a significant geographical spread, having become universal in the Spanish-speaking world, or being used throughout Spanish America or in several major Spanish American countries. The origin of the words indicated with (?) is still the subject of dispute, either as to the precise Amerindian source-language or as to whether they are Amerindian borrowings at all.

Borrowings from Arawak (including Taño) comprise, among others, ají 'chilli', batata 'sweet potato', bejuco 'liana, rattan', bobó 'cabin, hut', cacique 'Indian chief, local political boss', canoa 'canoe', caoba 'mahogany', cayo 'low island, key', comején 'termite', enaguas 'petticoat', guacamayo 'macaw', hamaca 'hammock', huracán 'hurricane', iguana 'iguana', maguey 'agave', maíz 'maize', maní 'peanut', sabana 'savannah', tana 'prickly pear', yuca 'cassava'.

Loans from Carib include batea (?) 'flat pan for separating gold from sand', butaca 'easy chair', canibal 'cannibal', curare 'curare', loro 'parrot', mico 'monkey', piraquá 'wooden canoe'.

A few amerindianisms borrowed in the Caribbean area may be from Arawak or from Carib: ají 'kind of yam', guasa 'joke', guateque 'party', guayaba 'guava'.

Among the many loans made to Spanish by Nahuatl are aguacate 'avocado', cacahueta 'peanut', cacao 'cocoa', coyote 'prairie wolf', chicle 'chewing-gum', chile 'chilli', chocolate 'chocolate', galpón 'shed', guajolote 'turkey', hule 'rubber', jicara 'cup (for chocolate)', nopali 'prickly pear', ocelote 'ocelot', petaca 'tobacco pouch, cigarette case', petate 'palm matting', sionsote 'mockingbird', tiza 'chalk', tocayo (?) 'namesake', tomate 'tomato', zapolote 'buzzard'.

Loans from Quechua include alpaca 'alpaca', cancha 'open space; (tennis) court, etc.', coca 'coca', cóndor 'condor', guanaco 'guanaco', guano 'guano', llama 'llama', mate 'Paraguayan tea', palta 'avocado', pampa 'pampas', papo 'potato', puma 'puma', puna 'high plateau, altitude sickness', soroche 'altitude sickness', vicuña 'vicuña'.

---

4.10 Anglicisms

From Tupi-Guaraní have been borrowed ananá(s) 'pineapple', caboya, -o (?) 'guinea-pig', jaguar 'jaguar', mandioca 'manioc', ñandú 'rhea, American ostrich', petunia 'petunia', tapioca 'tapioca', tapir 'tapir', tiburon 'shark', tucán 'toucan', zarigüeya 'opossum'.

In addition, there are a number of frequent words, like cauchó 'rubber', whose origin is Amerindian but whose precise linguistic source is unknown.

When speakers of Spanish travelled or settled in the New World, their new experiences were often interpreted as we have just seen, by means of words borrowed from native languages. The earliest of these (canoa, cacique and the disputed borrowing niamés 'yams') are recorded in Columbus's shipboard log of 1492–93. However, it should not be forgotten that borrowing is not the only means of giving linguistic labels to new concepts; a pre-existing word may extend its sense to encompass the new experience. Thus, the jaguar, the puma, and the pineapple were at first named by means of the Old World terms tigre, león and piña, words which eventually came to have competitors of Amerindian origin (jaguar, puma, ananá(s)), and this competition may survive through the centuries. In the cases mentioned, jaguar and puma became part of the Old World Spanish, while tigre and león continue in use (at least in popular speech) in much of the area where these animals have their habitat; on the other hand, the borrowing ananá(s) is not used in Peninsular Spanish or in large regions of Spanish America (where only piña occurs).

On other occasions, a New World concept may be labelled by competing amerindianisms. A well-known case is that of the avocado; the term aguacate (from Nahuatl, as noted above) is used in Spain, in North and Central America and in the northwestern region of the southern continent (Colombia and Venezuela), while in territories further south (Ecuador, Peru, Bolivia) the Quechua loan palta is used.

For further discussion of Amerindian loans, including those whose distribution is limited to part of the American continent, see Buesa (1967).

---

4.10 Anglicisms

Until the middle of the twentieth century, almost all English words borrowed by Spanish were of British English origin, and were usually transmitted through writing, often via French. From the 1950s onwards, the main source of such borrowing has been American English, and anglicisms have been transmitted partly through written media (especially newspapers, translation of scientific works, etc.), but increasingly through the oral media (dubbing of American films, TV programmes, etc.).

The term 'anglicism', like others referring to interlanguage borrowing, is poorly defined. The definition adopted here is that of Pratt (1980) (from whom
many examples are taken); ‘anglicisms’ are loans whose immediate etymon is an English word or expression, irrespective of the source of the English word (whose ultimate etymon is often a word or expression of a third language). By this criterion, it is appropriate to include among the anglicisms of Spanish such items as the following, whose ultimate etymology is indicated in parentheses: anorak, kayak (Eskimo), kindergarten (German), kimono→quimono, judo, karate (Japanese), géiser (Icelandic), gongo (Malay), caqui, pijama (Persian). Similarly, it will be necessary to exclude from the list of anglicisms in Spanish such terms as the following, whose ultimate etymon is an English word, but which have reached Spanish through the medium of another language (usually French): auto-stop ‘hitch-hiking’, camping ‘campsite’, dancine ‘dance-hall’, footing ‘jogging’, parking ‘car-park’, récordsman ‘record-holder’, (espejo) retrorvisor ‘rear-view (mirror)’, smoking ‘dinner jacket’, en directo ‘live (e.g. TV programme)’. The items quoted should be considered to be gallicisms, since they are attested earlier in French than in Spanish and often reveal French processes of semantic adaptation.

The most frequent (but by no means the only) manifestation of the influence of English upon Spanish is lexical loan. Anglicisms of this kind generally reflect the need to label new concepts (although some are introduced for non-linguistic reasons such as snobbery) and affect a very wide range of semantic fields. The language of almost every aspect of urban, sophisticated life reflects borrowings from English, but the language of the media, fashion, business, science and sport are particularly affected. In the vocabulary of the communications media, which naturally intersects with the terminology of the technological world, we find: bestseller, cámara, camaraman, cassette, cinematoscope, clip, copyright, disc-jockey, folding, film(e), flash(h), hit, interviewar→interviar, intervii, LP→elpé, mass-media, monitor, offset, off (e.g. una vez en off), pick-up, playback, pop, póster, rol, scriptgirl, show, sketch, speaker→espiquer ‘news-reader’, spot ‘advertising spot’, suspense, trailer, transistor, video, videogatape. Lexical borrowings from English in the field of fashion and cosmetics are also common, such as: anorak, bikini, coldcream, cosmético, champú, cheveut, eslip→slip, jersey, jumper, kilt, kimono→quimono, loción, minifalda, nylon→nailon, overol, panty, pijama, pullover, rasplán, rimel, retaz, shorts, sueter→sweater, tweed. The language of commerce and finance provides examples such as: actuario, boom, boutique, broker, cartel, charter, delegación, devaluación, dumping, factoring, holding, inflación, leasing, manager, marketing, self-service), stock, témida, turismo. In the world of science and technology (including medicine) we note many anglicisms, including: acrilo, aeropuerto, aerosol, ameba, analgesia, baquelita, cibernética, ciclámato, colesterol, coma, contáiner, cracking, detergente, ecología, esquizofrenia, fobia, fuel-oil, gasoil, polución, quántum, quark, radar, robot, síndrome, spray ‘aerosol’, stress→estres, trobleg. Sporting language has long shown a particular openness to anglicisms, which can be exemplified by: bantam, bíxibol, bob, bobseigh, boxes, bunker, cadie, córner, crawl→crol, croquet, cross country, chutar, doping, dribbling, fútbol, gol, golf, groggy→grogui, handicap, hockey→jóquey, jockey, judo, júnior, karate, karrting, kayak, knock-out→nocaout, lob, match, off-side, par, penalty, ping-pong, pony, record, ring, round, rugby, set, slam, smash, sparring, sprint, tándem, tenis, volleyboll, wéltler.

English can also be seen to have affected the morphology and syntax of Spanish (see Pratt 1980). The semantic influence of English upon Spanish, extending the sense of pre-existing Spanish words, has also been considerable; this latter phenomenon is examined and exemplified in 5.1.5.

4.11 Catalanisms

Considerable numbers of words have passed from Catalan to Castilian, especially in a number of fields in which the inhabitants of Catalonia, Valencia and the Balearic Islands were perceived by Castilian speakers to be pre-eminent. In some cases, these words were earlier borrowed from other sources, including Occitan, Old French, Italian and Arabic.

Food is one semantic field in which such borrowing is common. We find names of fish, such as anguilla ‘eel’, calamar ‘squid’, jurel ‘horse mackerel’, maja ‘mullet’, rape ‘monkfish’, together with many other culinary terms including anís ‘aniseed’, batifarra ‘Catalan sausage’, entremés ‘hors d’oeuvre: short comedy’, escalaf ‘to poach (eggs)’, escarola ‘endive’, horchata (?) ‘drink made of almonds or chufas’, paella ‘id.’, sémola ‘semolina’, vinagre ‘vinegar’, vinagreta ‘vinaigrette’.


The Catalans were the predominant sea-faring group in the Peninsula until the end of the Middle Ages and this predominance is reflected in the many Catalan borrowings belonging to this semantic field, such as aferir ‘to grapple, anchor, grasp’, betún ‘bitumen’, buque ‘ship’, calafatear (?) ‘to caulk a ship’, esquele ‘skiff’, galera ‘galley’, gobernalle ‘helm’, golfo ‘gulf’, muelle ‘dock, pier’, naut ‘sailing ship’, socaire ‘lee’, surgir ‘to anchor’, timonel ‘helmmsman’. Closely related are catalanisms relating to trade, which include a granell ‘in bulk’, mercader ‘merchant’, oferta ‘offer, tarifa (from Ar.) ‘tariff’.

4.13 Italianisms

The earliest borrowings made by Spanish from Italian are found towards the end of the Middle Ages, but it is in the sixteenth and seventeenth centuries that such loans reach their apogee, owing principally to the prestige of concepts emanating from Renaissance Italy and to the military involvement of Spain in the Italian Peninsula. Borrowing has continued, particularly in the field of music, down to the present, but became relatively infrequent after the eighteenth century. It should be noted that not all Italianisms originate in Tuscan; some (particularly maritime terms) are borrowings from Genoese, Venetian, Milanese, etc., or from southern Italian or Sicilian varieties.

Adaptation of italianisms to Spanish is, as might be expected, generally straightforward. However, there is a small number of words which in the fifteenth century and in the Golden Age appear in Spanish with final /e/ following an ungrouped dental or alveolar consonant (contrary to the normal Spanish development process, for which see 2.4.3.2.), and which are usually described as latinisms, but which may in fact more reasonably be described as italianisms: felice ‘happy’, infeliz ‘unhappy’, interes ‘interest’. This /e/ is later lost, bringing these words into line with the regular pattern of Spanish words: feliz, infeliz, interés, etc.


4.14 Word-formation

The vocabulary of Spanish can be said to consist of three components. In addition to words inherited from Latin (popular and semi-learned words, see 4.1), and to words borrowed from other languages (4.2–13), the Spanish lexicon includes items which have been created, through word-formation, by means of the language’s internal resources. The term ‘word-formation’ includes reference to prefixation (4.14.1), derivation (4.14.2) and composition (4.14.3).

4.14.1 Prefixation

Latin prefixes were in close relationship with the prepositions of the language (see the discussion at 3.8.1), in that many particles which functioned as prepositions also functioned as prefixes, usually with similar sense. However, not all the particles that have survived as prepositions have also survived as prefixes.

Many prefixed Latin words ceased at an early stage to be analysed by speakers as consisting of {prefix + base morpheme} and developed thereafter as unitary words. This is exemplified by cases like PROFECTU > provecho ‘advantage’, where Latin /f/ has been given the treatment normally accorded to this phoneme in intervocalic position (i.e. 1-E → OSP. /f/, MSPh. /f/; see 2.5.3.2.2). Similarly, DECOLLARE > degollar ‘to cut the throat’, where /k/ is given intervocalic treatment (> /g/). Where speakers continued to perceive the morphological complexity of a word, any consonant or consonant group which follows the prefix (and therefore stands at the beginning of the base morpheme) is treated in the same way as if it were in word-initial position. Thus, DEFENSA – forbidden (land) – must for some centuries have continued to be perceived as a complex word, since it develops to dehesa ‘enclosed pasture’, showing the Latin /h/ is here given word-initial treatment (becoming /h/; later /h/, as in FOMU > humo ‘smoke’, etc.; see 2.5.6, 2.6.4). In a small but crucial number of cases, the complex structure of the prefixed Latin word has continued to be perceived as such throughout its history, as in APREHENDERE ‘to seize’ (in relation with PREHENDERE ‘id.’) > APRENDER ‘to learn’ (still in relationship with PRENDER ‘to seize’).

It is cases like APREHENDERE > APRENDER which provided the model (in spoken Latin or at any later stage including the present) for the creation of new words by the addition of a prefix to a pre-existing verb, noun, adjective or adverb (examples of prefixed adverbs can be seen at 3.4). The prefixes which have a continuous history from Latin to Spanish (and which served and serve to create new words) are discussed below. No distinction is made, in the examples, between simple prefixation and parasyntasis, in which a prefix and a suffix are simultaneously added to a root (as in des- + alm - + ado → desalmado ‘heartless’) and where the parasynthetic nature of the structure is recognizable by the absence from Spanish of words consisting of the same prefix and the same stem, or of the same stem and the same suffix (in this case, by the absence of such words as *desalma or *almado). It should also be noted that, whereas in Classical Latin sequences of two or more prefixes were rare (e.g. COMPROMITTO), spoken Latin and its successors (including Spanish) expand this possibility considerably.

AD- The Latin meaning (‘to, towards’, etc.) has been almost entirely lost from this prefix, and its descendant a- is generally void of sense. It is used above all to create verbs from nouns or adjectives (e.g. agrupar ‘to group’, amontonar ‘to pile up’, apaciguar ‘to pacify’, atormentar ‘to torment’, agravar ‘to aggravate’, amorecer ‘to deaden’), but may appear in other formations (e.g. adios ‘goodbye’). The form ad- appears in a number of prefixed forms borrowed from Latin: adaptar ‘to adapt’, adherir ‘to adhere’, admirar ‘to admire, wonder at’, etc.

DIS- The notion of ‘separation’ inherent in the Latin prefix was later expanded to include ‘away from’ (absorbing DE-) and ‘out of’, so that it came into competition with EX-, and in particular with the compound prefix DE + EX-, which (at least before stems beginning with a consonant) shared the same phonetic outcome as DIS-, namely des-. The competition with ex-, whereby Spanish came to have two prefixes, des- and es-, of identical meaning and similar structure, explicates the frequent alternation in Old Spanish and modern non-standard speech between these two prefixes (e.g. destender–estender ‘to spread, extend’). However, the modern
standard language has resolved these cases of alternation, usually on etymological grounds (so that *estender* is preferred, and is erroneously respelled *extend-, on the basis of *extendere*). Examples of *des-* include: *desconocer* 'to distrust', *deseo* 'to unsew', *desescribir* 'to retract', *deseschar* 'to discard', *deshacer* 'to undo', *deshonrar* 'to dishonour', *desmenzar* 'to deny', *desviar* 'to deflect'.

Forms in *dis-, *di-, or *de-* betray their learned origins: *discernir* 'to discern', *disformar* 'deformed', *divertir* 'to amuse', *denegar* 'to refuse'.

**IN-** This prefix retains the Latin sense 'in(to), on', and its form in Old Spanish is generally *en-,* even before *l/ and *h/* (reflecting the neutralization of *ml/ nl/ and *lh/* in syllable-final position), although modern spelling distinguishes *en- from *em-,* in imitation of the Latin alternation between *in-* and *im-.* *En-** is used to create verbs from nouns or adjectives: *embarrar* 'to cover with mud', *emborrachar* 'to make drunk', *empapelar* 'to paper', *empañar* 'to pledge, pawn', *encabecer* 'to lead', *enganchar* 'to hook'; *engordar* 'to fatten', *enlazocer* 'to go or drive mad', *enrasar* 'to make level'. Occasionally it is used to create new verbs from existing verbs: *embeber* 'to soak up', *encoger* 'to shrink', etc.

In a few instances, *in-* was compounded with *ex-,* producing the sequences *ens- or OSp. *en-*; *epend-: *ensalar* ('< *IN EX ALTARE* 'to exalt', *ensanchar* 'to widen', *enjalbegar* ('< *IN EX ALBICARE* 'to whitewash', *enjuagar* ('< *IN EX AQUARE* 'to rinse, *enjugar* ('< *IN EX ACIARE* 'to dry'.

The learned descendant of *in-,* Spanish *in-* has two values, the first of which is indistinguishable from that of popular *en-:* *immiscuir* 'to mix', *innato* 'inborn', *inspirar* 'to inspire'. The second, negative, value of Latin *in-* is also widely imitated in Spanish, not only under the form *in-**im-* but under forms which imitate the Latin assimilation of *n* before *l* and *r: *insenso* 'senseless', *impiedad* 'impity', *impopolar* 'unpopular', *ilegítimo* 'illegitimate', *ilimitado* 'unlimited', *irreal* 'unreal', *irrespetuoso* 'disrespectful'.

**EX-** The popular descendant of *ex-* is OSp. *ex-* (sometimes now respelled *ex-* but without phonological change). It has been seen (dis-, above) that *es-* has since Latin times been in competition with *des-* in the sense 'out of'. Forms with this prefix include *escazar* 'to escape', *escardar* 'to weed', *escoger* 'to choose', *extendir* 'to spread'.

Learned use of this prefix, in the form *ex-,* extends from borrowings of forms which already displayed this prefix in Latin (in which case we also find examples of *e-:* *emanar* 'to emanate', *exhibir* 'to exhibit', *extirpar* 'to eradicate'), to the addition of the prefix to stems not found in Latin with *ex-* (excentrico* 'eccentric', *excarrélar* 'to release from prison'), and its semi-independent use before certain nouns (ex primer ministro 'ex-prime minister', *ex presidente* 'ex-president').

**INTER-** The popular descendant of this prefix, *entre-,* appears in relatively few forms, but shows considerable development of sense from the Latin 'between' (as in *entrecomillar* 'to place in inverted commas') to 'partially' (*entreambir* 'to half open', *entresacar* 'to thin (the hair, etc.)'), 'reciprocally' (*entreayudarse* 'to help one another', *entreocurar* 'to inter-twine'), or 'intermediate' (*entrecano* 'greying (hair)', *entrefino* 'of medium quality').

Words which display the learned form of this prefix, *inter-,* are most frequently borrowings or calques from other languages, typically French or English (e.g. *interferir* 'to interfere', *intermuscular* 'intermuscular'), although some are borrowings made directly from Latin by Spanish (e.g. *interrumpir* 'to interrupt').

**TRANS-** The Latin sense 'across' rarely appears in the popular reflex of this prefix, *trans-,* (e.g. *travalar* 'to fly across'). Instead we find other meanings such as 'behind' (* traslapar* 'to overlap', *trastendea* 'back room (behind a shop)'), 'during' (*transnachar* 'to stay up late'), or 'excessively' (*transalar* 'to disrupt').

No doubt because of the frequent reduction (at least in conversational style) of syllable-final *ls/ to l/s* (see Navarro 1961: 112), the learned form of this prefix, *trans-,* is in many cases interchangeable with the popular form, both forms appearing with the same stem (e.g. *transmitir*—*transmitir* 'to transmit').

**SUB-** Although the popular reflex of the preposition *sub* has now all but disappeared, that of the prefix *sub-* is, reasonably well represented in Spanish. Most frequently it retains the Latin sense 'under': *sobarba* 'double chin', *socavar* 'to undermine', *solapar* 'to overlap', *solomillo* 'sirloin', *someter* 'to subdue', *soterrar* 'to bury'. Occasionally, this prefix is used to attenuate the action indicated by the stem: *soasar* 'to roof lightly', *sofreir* 'to fry lightly'.

The learned form of the same prefix, *sub-,* may indicate location, as in *subseuio* 'subsoil', *submarino* 'submarine', *subrayado* 'underlining', but more often has the metaphorical value of 'less than' the notion expressed by the stem, e.g. *subdesarrollado* 'underdeveloped', *suhvalorar* 'to underrate'.

**SUPER-** The relatively frequent popular descendant of this prefix, *sobre-,* appears with the senses 'above' (e.g. *sobrecama* 'bedspread', *sobrenadar* 'to float'), 'after' (e.g. *sobremesa* 'after-dinner chat',
sobrevivir ‘to survive’), and ‘in excess’ (e.g. sobrecargar ‘to overload’, sobremanera ‘exceedingly’, sobresueldo ‘bonus’).

Probably under the influence of English, the learned form of the prefix, super-, has enjoyed an enormous productivity in recent decades. It shows the same range of meanings as the popular form, but with the sense of ‘excess’ being by far the most frequent: superestructura ‘superstructure’, supervivencia ‘survival’, supercomisión ‘super-charging’, superpoblación ‘over population’.

RE- In the case of this prefix, it is impossible to distinguish popular and learned reflexes, on the basis of form; both types of transmission of course produce re-. However, on the basis of meaning it may be possible to make some distinction, since it seems likely that the repetitive sense now sometimes associated with re- is a recent development, due to influence from Latin or from other modern languages which commonly display the repetitive meaning of this prefix (e.g. French or English). Examples of this development include reanudar ‘to renew’, reaparecer ‘to reappear’, rehacer ‘to remake’. Traditional senses include reference to place (e.g. recámara ‘dressing room’ (i.e. ‘a place beyond a bedroom’), recocina ‘scullery’, rebotica ‘room behind a chemist’s shop’), and occasionally to time (e.g. redolor ‘discomfort remaining after an accident’), but most usually the prefix simply emphasizes the notion expressed by the stem: rehién ‘very well indeed’, rebuscar ‘to search thoroughly’, recalentar ‘to overheat’, remoler ‘to grind up’, repudiar ‘to rot completely’, retomar ‘to shake (violently)’.

Other prefixes derived from Latin are almost exclusively learned. Per- and pro-, whose prepositional correlates merge as por (see 3.8.1), only in one instance behave in parallel fashion, in the development perfectis > porfia ‘stubbornness’. Per- is found, with intensifying value, in the language of the rustic characters of Renaissance drama (probably reflecting the rural speech of the Salamanca area; see Penny 1990a), and with similar value in the rural speech of present-day Asturias, but is absent from standard Castilian except in learned formations like perdurar ‘to last’, perjurar ‘to commit perjury’, but also ‘to curse repeatedly’. Pro- is always learned (e.g. promedio ‘average’, prometer ‘to promise’).

All of the following, ultimately of Latin origin, have entered Spanish through writing, their frequency increased today by the many anglicisms which display them: post(e) (-< post-), ante(-<ante-), pre- (< prae-), com-<con~co- (< com-, the correlate of the preposition cum), retro- (< retrō), contr- (< contrā), extra- (< extra-), intra- (< intrā), infra- (< infra-), supra- (< suprā), circum-<circum- (< circum-), ultra- (< ultra-); e.g. posguerra ‘post-war period’, antepen ‘to antepose’, preindustrial


The stock of Spanish prefixes has similarly been enriched by the introduction of a number of forms whose origin is ultimately Greek but whose current frequency is again probably due to the influence of English, where these prefixes are especially numerous in scientific and journalistic language (see Pratt 1980: 185–91). The prefixes concerned are exemplified by the following selection of words: antioxidante ‘antioxidant’, autopromoción ‘self-promotion’, hipertensión ‘hypertension’, macroeconómico ‘macro-economic’, microorganismo ‘micro-organism’, pericráneo ‘pericranium’, polivalencia ‘polyvalence’, protohistoria ‘protohistory’ (see also 4.14.3).

4.14.2 Derivation

The addition of suffixes to pre-existing stems in Spanish serves two contrasting purposes: first, the creation of a word which refers to a different concept (albeit a related one) to that referred to by the original word, and second, the addition of a nuance which reveals the speaker’s attitude to the concept concerned. In the latter case, no new concept is involved; the original word and the derived word refer to the same concept. To exemplify the two processes we may consider, first, the relationship between vaca ‘cow’ and vacada ‘herd of cows’. It is evident that the derived form vacada indicates a concept which is different from (but related to) the concept indicated by the base-word. However, if one compares, secondly, the word gordo ‘fat’ with its derivative gordito ‘nice and fat, chubby’, it is clear that the concept referred to is essentially the same in each case, but that in the case of gordito the speaker’s attitude to the concept is approving and affectionate.

However, these two derivational processes are not always as sharply distinguished as this discussion suggests: they overlap in two ways. On the one hand, the ‘same’ suffix may serve both derivational purposes. The ending -ito which conveys overtones of affection in gordito expresses no such nuance in carrito; in this case, the concept referred to (‘trolley’) is a different one from that indicated by carro (namely ‘cart’, ‘car’, etc.). Similarly, the -azo which appears in vinazo ‘coarse wine’ (or, sometimes, ‘magnificent wine’) clearly carries emotive value but does not alter the basic reference (to ‘wine’), while the ‘same’
element in cabezazo invokes reference to a different concept ('head-but') from that of the base-word cabeza ('head'). On the other hand, the emotive nuance added to a concept by the addition of a suffix may be closely related to an element of objective meaning. Thus, there is often a correlation between the affectionate response of the speaker and relative smallness of the referent (e.g. gatito 'kitten'), or between the repugnance expressed by the speaker and the relative bigness or coarseness of the referent (e.g. novelón ‘long boring novel’).

Despite the difficulties posed by the overlaps between these two derivational processes, it is useful to keep them separate in discussion. Further, the standpoint taken here is that where emotive nuances are associated with a given suffix, such nuances predominate over any element of objective meaning which may also be present. This view is in accordance with that expressed in a fundamental study of the suffixes concerned (Alonso 1935) and can further be justified by reference to Gooch (1970). In what follows, we shall distinguish the two types of derivation by using the label lexical for the type which produces new names for new concepts, and affective for the type which expresses the speaker's attitude towards the concept evoked.

4.14.2.1 Lexical derivation

Suffixed may be used to create new nouns, adjectives and verbs. Each of these categories will be considered in turn. A derived noun may have as its base another noun, an adjective or a verb. In table 4.1 are listed the main suffixes (i.e. those which remain productive or which at some stage have been so) which may be applied to each type of base, together with the source of each suffix and representative examples of each combination.

It can be seen that in a number of cases, the form of the suffix has not been arrived at through regular phonological change. That is, a number of suffixes have passed into Spanish as a result of the borrowing of Latin words which contained the suffixes concerned, after which the suffix could become available for application to inherited stems. The learned suffixes concerned are: -ia, -ismo, -ista, -ante, -(i)ente, -(i)ancia, -(i)encia, -oro, -ento, -ción. In a few cases, such learned forms compete with inherited forms of the same Latin suffix; this can be seen in the case of -anza-ancia, -ero-orio, -mental-mentalismo-zón-ción. In some of these cases, the learned form (e.g. -ancia, -ción) is today more productive than the popular form; in the remaining cases (-ero, -mentismo) the reverse is true. Similarly, -afe results from the borrowing of French words from the Middle Ages onwards, and coexists with inherited -adgo, later -ago, the latter now unproductive.

Certain of the suffixes used to derive nouns from verbs, namely -ero, -or, -ura, descend from Latin suffixes which were applied to participles. This structural feature continues to characterize Spanish, so that in the cases concerned we find participial -ad-, -ed-, or -id- between stem and suffix.
4.14 Word-formation

Table 4.2 Derived adjectives

<table>
<thead>
<tr>
<th>Base</th>
<th>Suffix</th>
<th>Source</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjective</td>
<td>-ado</td>
<td>-TU</td>
<td>azulado 'blush'</td>
</tr>
<tr>
<td></td>
<td>-enco</td>
<td>?</td>
<td>azulenoco 'blush'</td>
</tr>
<tr>
<td></td>
<td>-ento</td>
<td>-ENTU</td>
<td>avarento 'misery'</td>
</tr>
<tr>
<td></td>
<td>-ino</td>
<td>-INU</td>
<td>blanquecino 'whitish'</td>
</tr>
<tr>
<td></td>
<td>-ista</td>
<td>-ista (&lt; -ISTA)</td>
<td>socialista 'socialist'</td>
</tr>
<tr>
<td></td>
<td>-izo</td>
<td>-ICU</td>
<td>rojozo 'reddish'</td>
</tr>
<tr>
<td></td>
<td>-oid</td>
<td>GK. - oCIDe</td>
<td>negroido 'Negroid'</td>
</tr>
<tr>
<td></td>
<td>-iso</td>
<td>-ISU</td>
<td>verdeiso 'greenish'</td>
</tr>
<tr>
<td></td>
<td>-usc</td>
<td>-USCU</td>
<td>purdusco 'greyish'</td>
</tr>
<tr>
<td></td>
<td>-ucce</td>
<td>-USCU</td>
<td>blanucce 'whitish'</td>
</tr>
<tr>
<td>Noun</td>
<td>-al→ar</td>
<td>-AL→ARE</td>
<td>invernal 'wintry'</td>
</tr>
<tr>
<td></td>
<td>-ano</td>
<td>-ANU</td>
<td>seglar 'secular, lay'</td>
</tr>
<tr>
<td></td>
<td>-ense</td>
<td>-ENSE</td>
<td>ateniense 'Athenian'</td>
</tr>
<tr>
<td></td>
<td>-ello</td>
<td>-ENEU</td>
<td>panameño 'Panamanian'</td>
</tr>
<tr>
<td></td>
<td>-ero</td>
<td>-ERIU</td>
<td>playero 'beach'</td>
</tr>
<tr>
<td></td>
<td>-és</td>
<td>-ENSE</td>
<td>montés 'wild'</td>
</tr>
<tr>
<td></td>
<td>-esco</td>
<td>-ISCU</td>
<td>gigantesco 'gigantic'</td>
</tr>
<tr>
<td></td>
<td>-ico</td>
<td>-ICU</td>
<td>bordónico 'Bourbon'</td>
</tr>
<tr>
<td></td>
<td>-ll</td>
<td>Ar. -l</td>
<td>irán 'Iranian'</td>
</tr>
<tr>
<td></td>
<td>-llo</td>
<td>-LE</td>
<td>estudianllo 'student'</td>
</tr>
<tr>
<td></td>
<td>-ino</td>
<td>-INU</td>
<td>cristalino 'crystalline'</td>
</tr>
<tr>
<td></td>
<td>-ón</td>
<td>-ON</td>
<td>narizón 'long-nosed'</td>
</tr>
<tr>
<td></td>
<td>-oso</td>
<td>-OSU</td>
<td>miedooso 'fearful'</td>
</tr>
<tr>
<td></td>
<td>-udo</td>
<td>-UTU</td>
<td>orejudo 'big-eared'</td>
</tr>
<tr>
<td>Verb</td>
<td>-able, -ible</td>
<td>-ABLE, -IBLE</td>
<td>inoxidable 'stainless (steel)'</td>
</tr>
<tr>
<td></td>
<td>-ado, -ido</td>
<td>-ITU</td>
<td>movible 'moveable'</td>
</tr>
<tr>
<td></td>
<td>-ante, -(i)ente</td>
<td>-ANTE, -(i)ENTE</td>
<td>agupado 'dull, weak'</td>
</tr>
<tr>
<td></td>
<td>-ero</td>
<td>-ERIU</td>
<td>aburrido 'boring'</td>
</tr>
<tr>
<td></td>
<td>-ivo</td>
<td>-IVU</td>
<td>impulso 'impulsive'</td>
</tr>
<tr>
<td></td>
<td>-izo</td>
<td>-ICU</td>
<td>olvidadizo 'forgetful'</td>
</tr>
<tr>
<td></td>
<td>-ón</td>
<td>-ONE</td>
<td>mirón 'nosey'</td>
</tr>
<tr>
<td></td>
<td>-or</td>
<td>-ORE</td>
<td>embriagador 'intoxicating'</td>
</tr>
</tbody>
</table>

Derivation of nouns in -o from others in -a, and vice versa, is a relatively unproductive process, but accounts historically for such pairs in Spanish as manzano 'apple'; manzana 'apple', cesto '(small) basket'; cesta '(large) basket', etc. On the other hand, deverbal nouns in -e, -o, -a are quite frequent and the process continues to be productive.

Derived adjectives may be based, similarly, on nouns, verbs, or on other adjectives, as listed in table 4.2.

It will be seen that adjectives derived from other adjectives for the most part denote colours related to those denoted by the base. Adjectives in -ista function also as nouns.

As in the case of derived nouns, we find cases of competition between popular and learned descendants of the same Latin suffix; thus popular -és coexists with learned -ense. Again as in the case of derived nouns, some deverbal adjectives have a participle as their base; this is generally so in the case of -ero, -izo and -or.

Derived verbs are again based upon nouns, adjectives or verbs. The principal productive types of derivatives are listed in table 4.3.

It can be seen that both the suffixes -ear and -izar descend ultimately from the same Greek causative suffix -λειν. The latter was adapted, as it passed into early spoken Latin by oral transmission, to -IDARE whose later development to Spanish -ear is regular (see 2.5.2.2.4(4)). The suffix -ear is applied straightforwardly to nouns and adjectives, but to verbs usually by means of an infix (e.g. toqu + et + ear, freg + et + ear, gim + et + ear) which generally has 'repetitive' value.

Gk -ξειν was borrowed a second time by (later literary) Latin, in the form -IZARE, which was later taken through writing into Medieval Spanish as -izar. Perhaps under the influence of other modern languages, where the corresponding suffix (-ise, -ize) is very productive, -izar has become one of the most frequent means of creating new verbs.

Latin -IFICARE also shows dual development. Via popular transmission, it shows the expected regular changes i > ie, voicing of intervocalic -f- and -c- to [β] and [v] respectively, and loss of pretonic i: *-eβiyar. At this stage,
/β/ was modified to a closing glide (as in the case of OSp. /βd/; 2.5.5 end), but then metathesized with the following consonant: [−eβˈɣar] > [−eʊˈɣar] > [−eˈɣar]. Pretonic /e/ was then raised to /i/ by assimilation to the following glide (cf. AEQUALE > OSp.  egal >  igual ‘equal’): -iguar. When transmitted through writing, -ificare shows the usual minimal change to -ificar.

It will be noted that all the productive verbal suffixes belong to the -ar class, with the exception of -ecer, which usually appears in parasythetic derivatives. In the Middle Ages, this type often competed with non-derived verbs in -ir (escarnir~escarner, gradir~agradecer, guarnir~guarnerce, resplandir ~resplandecer, etc.), eventually ousting the simple verb in each case.

4.14.2.2 Affective derivation

In 4.14.2 we defined affective derivation as the process of adding suffixes which reveal the speaker’s attitude towards the concept denoted by the base. The view was also expressed that the affective content of the suffixes concerned was more salient than any objective meaning (such as ‘smallness’, ‘largeness’, ‘coarseness’, etc.) the suffixes might also convey. Each of the main affective suffixes of Spanish will be discussed here, with comments on emotive and objective content, whether the suffix lends itself to lexicalization (the denotation of a concept different from that of the base and with loss of affective value), and with consideration of its origin. The suffixes may be applied to nouns, adjectives and participles, or adverbs.

-ito denotes approval/affectation and is diminutive in value: osito ‘teddy bear’, librito ‘nice (little) book’, crecitito ‘nice and tall (of child)’, bajito ‘small but nice’. Its Latin origins are hazy; it may have been extracted from certain personal names (IULITTA, BONITTA, SALVITTA), but must have acquired frequent use in spoken Latin, since it is well represented in Romance (Fr.-et, -ette, It.-etto, -etta, Cat.-et-eta, etc.). The form of the Latin suffix appears to have alternated between *-ittu, whence Sp. -ito, and *-ittu, from which the remaining Romance forms descend, including Fr., Occ., Cat.-et, borrowed by Castilian as -ete (see below). The suffix -ito is infrequently represented in thirteenth- and fourteenth-century texts (there is only one example in Berceo, who makes frequent use of other affective suffixes, and one in Juan Manuel) but gains in frequency in those fifteenth-century writers who most closely reflect the spoken language (the Archpriest of Talavera, Fernando de Rojas, etc.). Its frequency grows among similar writers of the Golden Age (especially Santa Teresa), whereafter it gradually displaces competitors (chiefly -illo and -ico) until it achieves its present dominance. Only in Andalusia does -illo retain its status as the preferred affectation suffix.

-ico is also affectionate and also has diminutive value: besico ‘sweet little kiss’, malico ‘poorly’, un tantico ‘a wee bit’. Its origins are unknown, it has few Romance cognates, and its history until the Golden Age closely parallels that of -ito. But since that period, -ico has retreated from Castilian (being used now only in a limited range of forms) and remains frequent only in Navarre, Aragon, Murcia and eastern Andalusia, and parts of Spanish America, where it often the preferred affectation form.

-in plays a similar role to that of -ito and -ico (it is affectionate and diminutive): pajarin ‘sweet little bird’, pequeñin ‘tiny, wee (fellow)’. However, it lends itself to numerous lexicalizations: comodín ‘joker (in cards)’, futbolín ‘table football’, etc. I may represent a development of the Latin suffix -inus, used to indicate the young of certain animals (e.g. PALUMBRINUS (PULLUS) ‘young dove’) and throughout its history has been commonest in the western half of the Peninsula (Pgt.-inho, Gal.-iño, Ast.-ín, León and Extremadura -ino). It has always had some frequency in Castilian, but cannot be freely applied to any base.

-illo although affectation (and diminutive) for most of its history, this suffix now often has a slightly pejorative tone: asuntollo ‘an unimportant piece of business’, empleillo ‘rotten little job’, novellillo ‘piffling novel’, envidiosillo ‘pettily envious’. It also gives rise to frequent lexicalizations: camilla ‘stretcher’, casilla ‘pigeonhole’, molinillo ‘coffee grinder’, pitillo ‘cigarette, fag’. It is the first of the suffixes considered here which has a clear Latin antecedent, -ellus, which was certainly diminutive and probably affective. Its descendant, -ello, was the commonest Old Spanish affective diminutive, and was gradually replaced by -illo, spreading from the Burgos area (see 2.4.2.5). In the Golden Age, -illo is still dominant (and still affective) but since that time has been challenged by -ito as the ‘normal’ Castilian diminutive form and has acquired its current slightly pejorative tone (except in Andalusia).

Lat. -ELLUS was applied directly to nouns and adjectives in -US and -A, but in the case of nouns and adjectives with other endings an infixed - iterative was inserted. Since the /k/ of this infix was commonly intervocalic, it evolves to OSp. /d/’, MSP. /β/’ *PAUPERCELLU (for PAUPERCELO) > OSp. POBREZIELLO > MSP. POBRECILLO. The extension of the use of this infix is dealt with at the end of this section.

-ejo This suffix is generally pejorative in value and has diminutive sense when applied to nouns: animalejo ‘wretched creature’, callejejo ‘alley’, lugarejo ‘tin-pot village’, medianejo ‘poor to middling’. It gives rise to a certain number of lexicalizations: candilejas
‘footlights’. The Latin antecedent of -ejo, -ulus, was attached directly to nouns in -us and -a (e.g. FLAMMA → FLAMMULA) but to other nouns by means of the infixed -(i)c- (PAUPER → PAUPER- CULUS). In this respect, -ulus was like -ellus (see ·illo above) and the two suffixes were probably also similar in their affective/diminutive value. However, unlike -ellus, -ulus was atomic (see 2.3.1), and in the case of nouns in -us and -a there was a strong tendency in spoken Latin to replace atomic -ulus with tonic -ellus: ROTA → ROTULA ‘little wheel’ > ROTELLA (> rodilla ‘kneecap’ > ‘knee’). Where -ulus continued to be productive (i.e. when combined with the infixed -ic-: -iculus), it lent itself to frequent lexicalization, including cases where the suffixed form displaced the original base-form, which thus lost all affective/diminutive value: OVICULA (for OVIS) ‘sheep’, AURICULA (for AURIS) ‘ear’, APICULA (for APIUS) ‘bee’ (whence oveja, oreja, abeja). The change in value, from affective to derogatory, is already apparent in Old Spanish, where -ejo has a value similar to that which it has in the modern language.

-uelo is now most frequently pejorative in tone and may have diminutive value: autoruelo ‘insignificant writer’, asexual ‘mean little eyes’, gentezuelo ‘petty riffraff’, gorrizuelo ‘nastily small and fat’. It easily lends itself to lexicalization: habichuelo ‘French bean’, hoyuelo ‘dimple’. Its Latin antecedent, -ulus, was at first inherently atomic (see 2.3.1) and was applied to forms which displayed hiatus between the two final syllables: FILIUS → FILLIOLUS, FLUVIUS → FILLVIOLUS. Under such phonological conditions, spoken Latin transferred the stress from the antepenultimate to the penultimate syllable (see 2.3.1, end), thus converting the suffix from atomic to tonic and ensuring its continued identity (FILLIOLUS > hijuelo ‘offspring’). The suffix -uelo was very frequent in Old Spanish and apparently retained the affectionate value of its Latin ancestor. During the Golden Age, it continues to be well represented, but then declines in frequency and acquires its current predominantly pejorative tone. It has already been remarked that -uelo is often attached to its base by means of the infixed -(e)z, borrowed from structures like OSP. simpleszelo: ladronzuelo ‘little thief’.

-ete has above all a jocular tone (making it especially open to ironic use) and generally diminutive value: comedietta ‘insubstantial little play’, curite ‘jolly old priest’, pillete ‘young rascal’, tacete ‘pretty stingy’. It appears particularly frequently in lexicalized forms: boquete ‘hole’, camioneta ‘van’, chincheta ‘drawing pin’, salmonete ‘red mullet’. This suffix has the same (obscure) origin as ·ito but has reached Spanish as a result of borrowing of French/Occitan/Catalan words in -et. It may appear with the infixed which originated in ·cellus (trenecete ‘a joke of a train’), but may not be applied to bases with absolute freedom.

-uco has pejorative tone and diminutive value, where it appears (albeit infrequently) in the standard language: casuco ‘mean little house’, frauluco ‘petty little friar’. It gives rise to occasional lexicalized forms: hayucos ‘beech mast’. By contrast, in the Cantabria region, this suffix is the commonest affectionate morpheme. Its origin is uncertain, but it may be a variant of ·ico, on the pattern of other suffixes which share the same consonantal pillar combined with different vowels (e.g. ·azo, ·izo, ·uazo).

-ucho is (like ·uco) pejorative in tone and often diminutive in value: aldeichuca ‘miserable little village’, animaluchu ‘wretched creature’, fesichu ‘rather ugly’, medicucho ‘tenth-rate doctor’, tabernuchu ‘grotty little bar’. It provides only occasional lexicalizations: agulluchu ‘eaglet’. Its origin is unclear, as is its history in Spanish.

-ôn As applied to words denoting or referring to people, this suffix is generally pejorative, and augmentative in the sense that it implies ‘increase’ or ‘excess’ of some quality: feôn ‘very ugly’, maricôn ‘pansy’, mujeron­a ‘hefty great woman’, sargento­na ‘bossy woman’, valentôn ‘boastful’, zampon ‘very greedy’. As applied to non-personal concepts, -ôn still implies ‘excess’ and is pejorative in tone, except where ‘excess’ can be interpreted as desirable: caseron ‘great big house’, gotôn ‘big drop’, novelon ‘boringly long novel’. There are many cases of lexicalization of words containing -ôn: abejon ‘drone’, pimentôn ‘paprika’, velôn ‘oil lamp’; however, a number of lexicalized cases show that -ôn may have diminutive value: cordones ‘shoelaces’, ratôn ‘mouse’, tapôn ‘plug, stopper’, terrôn ‘clod of earth; sugar lump’.

The notion of ‘excess’ was present in the Latin antecedent, -one, of this suffix (e.g. násó, násonis ‘big-nosed’) and probably also the pejorative tone which springs from the notion of excess. However, we have already seen (4.14.2.1.) that -ôn extended its role to form lexical derivatives of various types. As an affective suffix, -ôn may now be added with considerable (but not complete) freedom to a wide range of bases.

-azo is similar to -ôn in its derogatory tone and its ‘augmentative’ value: acentazo ‘heavy, unpleasant accent’, broncazo ‘big row’, olaza ‘threatening wave’. Again like -ôn, its pejorative note may be replaced by an approving tone where ‘excess’ is seen as desirable: bodaza ‘slap-up wedding’, torazo ‘fine big bull’. There are
These include -angano (curdengano 'wretched priest'), -ango (querindango 'pathetic lover'), -astre (pillastre 'miserable young scoundrel'), -astro (camastro 'miserable bed'), -engue (blandengue 'contemptibly feeble (person)'), -ingo (señoritengo 'contemptible young gentleman'), -orio (papelorios 'rubbish papers'), -orrío (villorrio 'tenth-rate village'), -orro (chistorro 'crude joke'), -utre (franchute 'Frog (= Frenchman)'), -uza (gentuzca 'rabble, scum').

It should be noted that affective suffixes (and, to a much lesser extent, the suffixes responsible for lexical derivation (4.14.2.1)) may appear in sequences of two or more items joined to a single base. Such sequences usually consist of suffixes with the same or similar affective quality: chiquitillo 'little boy', chiquitín 'tiny', riachuelo 'insignificant stream', valentonazo 'utterly boastful man'. Where there is an apparent conflict of affective values, it is usually the case that the first suffix combines with the base to form a lexicalization and only the second retains affective value: salonicito 'attractive little room', caperucita 'charming little hood'.

The appearance (and origin) of the infix -(e)ñez- has been noted (see -illo, -ejo, above). Although the morpheme from which it descends was originally placed between -ELLUS and a base which ended otherwise than in -(e)ñez, and continues in such forms (e.g. florecilla), this infix has extended its use in the course of time in two ways; firstly, there are cases where the infix has come to be used before other affective suffixes (ladronzuelo, florecita); and secondly, there are many cases in which the infix has come to be used with bases ending in lor or lór: manecita, huertecillo, pueblito, viejito, etc. The latter forms are often ones in which the base displays a diphthong in either the final or the penultimate syllable, but are far from universal in the Spanish-speaking world; in general, American and Canadian Spanish prefers forms without infix (manita, huertito, pueblito, viejito).

For further discussion of the history of the ‘diminutives’ from the Middle Ages, see González Ollé (1962) and Náñez (1973). For additional detail on the current value of the suffixes considered here, see Alemany (1920), Alonso (1935), Alvar and Pottier (1983: 363–80) and Gooch (1970), from the last of whom a number of glosses have been taken.

4.14.3 Composition

Creation of new vocabulary by compounding two or more lexemes is today a relatively frequent process in Spanish. Such composition reveals differing degrees of fusion between the contributing elements, ranging from simple juxtaposition (where the second element modifies the first, e.g. tren correo 'mailtrain', ciudad dormitorio 'dormitory town, bedroom community') through union without modification (e.g. sordomudo 'deaf and dumb', abrelatas 'tin opener', tocadiscos 'record-player'), union via modification to lor of the final vowel of the first
element (e.g. machihembra 'tongue and groove', rajiblanco 'red and white'), to parasyntactic composition, in which two lexemes are compounded at the same time as a suffix is added (e.g. estadounidense 'American', sietemesino 'seven-month'). The classification adopted here is based upon the grammatical function of the words which enter into the compound and examples will be drawn from the various types of compounding just mentioned.

**Noun + noun.** It is normal for the second noun to modify the first: perro guardián 'guard dog', hombre rana 'frogman', buque-hospital 'hospital ship', aguamiel 'mead', telaraña 'spider's web', zarzamora 'blackberry (bush)'. In aguamiel 'sleet', machihembra 'tongue and groove', puerto espín 'porcupine', the two elements are co-ordinated rather than showing subordination of one (the second) to the other. Parasyntactic composition can be seen in salpimentar 'to season'.

**Noun + adjective.** This combination produces, on the one hand, co-ordinated nominal expressions (e.g. aguardiente 'eau-de-vie', camposanto 'cemetery', guardia civil 'civil guard', Nochebuena 'Christmas Eve') and, on the other, a wide range of adjectival expressions in which the second element is syntactically subordinated to the first (e.g. barbirrojo 'red-bearded', cariancho 'broad-faced', cejijunto 'bushy-browed', corniabierta 'wide-horned', cuellilargo 'long-necked', maniabierta 'open-handed', ojinegro 'black-eyed', patizambo 'knock-kneed', peligudo 'thorny (problem)', puntiagudo 'pointed', rabicorto 'short-tailed'; cabizbajo 'crestfallen' no doubt represents a modification of earlier *cabezibajo*, which accords with this pattern).

**Adjective + noun.** In this case, we find co-ordinated expressions which function as nouns: bajamar 'low tide', cortocircuito 'short circuit', cortometraje 'short film', extremaución 'extreme function', midiodia 'midday'.

**Adjective + adjective.** The result of composition (inevitably of co-ordinating type) between two adjectives may give rise to a new adjective (agridulce 'bitter-sweet', rajiblanco 'red and white (striped)', sordomudo 'deaf and dumb', verdinegro 'dark green'), or to a noun (altihajos 'vicissitudes', claroscuro 'chiaroscuro').

**Verb + noun.** This combination, in which the verb is imperative or present indicative and the noun functions as its direct object, has always constituted a frequent type of composition in Spanish: abre-latas 'tin opener', aguafiestas 'spoilsport', cortaplumas 'penknife', cortafuego 'fire-break', espantapájaros 'scarecrow', guardacostas 'coastguard', guardarropa 'wardrobe', hincepí 'firm footing', pasatiempo 'pastime', picamaderos 'woodpecker', portaviones 'aircraft carrier', portavoz 'spokesperson', rompecabezas 'puzzle', rompehielos 'icebreaker', sacacorchos 'corkscrew', saltamontes 'grasshopper', tragaperras 'fruit machine'. Less frequently, the noun of the compound does not function as direct object of the verb, but has one of a variety of other roles: cortafrio 'cold chisel', girasol 'sunflower', trotaconventos 'go-between'.

**Pronoun + verb.** This type is rare, but appears in the very common expression quehacer 'chore'.

**Verb + verb.** Such co-ordinated combinations, with or without the copula *el*, provide new nouns: duermelula 'the state of being half-asleep', pasapasa 'sleight of hand', de quitapón 'detachable', vaiven 'swaying; bustle'.

**Syntagmatic compounds.** These compounds, all with nominal function, originate in a variety of sequences of words, as they occur in connected speech: ciempies 'centipede', correvedile 'gossip', hazmerreir 'laughingstock', metomentodo 'busbody', paderneastro 'Lord's Prayer', tentemozo 'prop', tentempié 'snack'.

In addition to the types of composition discussed, we find a further type, which is of recent origin and now relatively frequent, and which represents a position intermediate between composition and prefixation. A 'prefixoid' (an element typically ending in *el* or *el*, but also in *el* or *el*) is used through writing from a Greek or Latin noun or adjective) is combined with a second element, which may be an inherited or borrowed word. It has been suggested (Pratt 1980: 186–9) that the profusion of such formations is one manifestation of the influence of English on the modern Spanish language. The 'prefixoids' include aero- (aerofotografía 'aerial photography'), amb- (ambigenérico 'ambigeneric'), archi- (archiconocido 'known by all'), auto- (autorrerato 'self-portrait'), bio- (biomecánico 'biomechanics'), electro- (electrodomésticos 'electrical household appliances'), euro- (eurocomunista 'Euro-communist'), ferro- (ferroniquel 'ferromagnet'), filo- (filosófico 'admirinig of the Soviet Union'), hidro- (hidroelectricidad 'hydroelectricity'), macro- (macroempresa 'big business'), micro- (microfilm 'microfilm'), mon- (monocarril 'monorail'), moto- (motorsegadora 'motor scythe'), multi- (multigrado 'multigrade'), núcleo- (nucleo-electricidad 'nuclear energy'), poli- (polideportivo 'sports hall'), proto- (prototipo 'prototype'), radio- (radiofaro 'radio beacon'), semi- (semidesnudo 'half-naked'), tele- (teledirigido 'remote controlled'), telediario 'TV news bulletin', termo- (termoconector 'thermocupula'), tri- (tricolor 'three-coloured'), zoo- (zoozocopia 'zoogeography').