30 Dialects of Spanish and Portuguese

JOHN M. LIPSKI

30.1 Basic Facts

30.1.1 Historical Development

Spanish and Portuguese are closely related Ibero-Romance languages whose origins can be traced to the expansion of the Latin-speaking Roman Empire to the Iberian Peninsula; the divergence of Spanish and Portuguese began around the ninth century. Starting around 1500, both languages entered a period of global colonial expansion, giving rise to new varieties in the Americas and elsewhere. Sources for the development of Spanish and Portuguese include Lloyd (1987), Penny (2000, 2002), and Pharies (2007). Specific to Portuguese are features such as the retention of the seven-vowel system of Vulgar Latin, elision of intervocalic /l/ and /n/ and the creation of nasal vowels and diphthongs, the creation of a “personal” infinitive (inflected for person and number), and retention of future subjunctive and pluperfect indicative tenses. Spanish, essentially evolved from early Castilian and other western Ibero-Romance dialects, is characterized by loss of Latin word-initial /f-/ and /s/, the diphthongization of Latin tonic /ɛ/ and /ɔ/, palatalization of initial C+L clusters to /ʎ/, a complex series of changes to the sibilant consonants including devoicing and the shift of /ʃ/ to /x/, and many innovations in the pronominal system.

30.1.2 The Spanish Language Worldwide

Reference grammars of Spanish include Bosque (1999a), Butt and Benjamin (2011), and Real Academia Española (2009–2011). The number of native or near-native Spanish speakers in the world is estimated to be around 500 million. In Europe, Spanish is the official language of Spain, a quasi-official language of Andorra and the main vernacular language of Gibraltar; it is also spoken in adjacent parts of Morocco and in Western Sahara, a former Spanish colony. In the Americas, Spanish is the official language throughout South America except Brazil, Suriname, French Guiana, and Guyana; in the Caribbean nations of Cuba, Puerto Rico, and the Dominican Republic; and in Mexico and all of Central America except for Belize. Unofficially, it is widely used in Belize, Haiti, Aruba, the US Virgin Islands, and in the United States, where the nearly 45 million speakers make the US a strong contender for second place among the world’s Spanish-speaking countries; Canada is home to nearly half a million Spanish speakers. Spanish is also residually present in the Philippines and the Mariana Islands. Spanish is the third most widely used language on the Internet (after English and Chinese).
30.1.3 The Portuguese Language Worldwide

Reference grammars of Portuguese include Cunha and Cintra (1984), Perini (2002), and Thomas (1974). Worldwide, the number of Portuguese speakers is estimated to be between 215 and 250 million, with the higher number including second-language speakers. Beyond Portugal, Portuguese is the official language of Brazil, Cape Verde, Guinea-Bissau, São Tomé and Príncipe, Angola, Mozambique, and East Timor. It still has some vitality in Goa (India) and is spoken natively in much of northern Uruguay and in the northeastern Argentine province of Misiones. Portuguese also has co-official status in Macau and Equatorial Guinea.

30.1.4 Mutual Intelligibility

At the level of educated speech, all varieties of Spanish are highly mutually intelligible. At the colloquial or vernacular level, differences in pronunciation and vocabulary and to a lesser extent morphosyntax result in considerable divergence, and although such differences are easily resolved in face-to-face encounters, they may represent obstacles to comprehension in passive listening situations. Differences between the Portuguese varieties of Portugal (extending to Lusophone Africa and Asia) and those of Brazil are often considerable, and Brazilians in particular frequently experience difficulty in understanding spoken European Portuguese, although easily comprehending the written language.

In their written forms, Spanish and Portuguese share a high degree of mutual intelligibility, once simple transpositions are mastered. In areas along the Spanish-Portuguese border and along the borders between Brazil and Spanish-speaking nations, mutual intelligibility of spoken Spanish and Portuguese is facilitated by familiarity, but regional and social varieties of the two languages often diverge to the point of limited mutual comprehension away from border regions. Although Spanish and Portuguese were once end points of a dialect continuum containing intermediate varieties such as Leonese and Extremeño, only (recently standardized) Galician enjoys contemporary vitality within the continuum.

30.2 Spanish: Main Sources

30.2.1 General

Moreno Fernández (2005) surveys recent corpora of oral Spanish. The most extensive multi-genre resource is the Corpus del Español searchable database (www.corpusdelespanol.org). Another research tool is the Corpus de Referencia del Español Actual (CREA) of the Real Academia Española (corpus.rae.es).

The first major attempt to consolidate transcriptions of spoken Spanish (and Portuguese) was the Estudio Coordinado de la Norma Lingüística Culta de las Principales Ciudades de Iberoamérica y de la Península Ibérica, known as the Norma Culta project (PILEI 1971–1973). Several of the transcriptions have been consolidated on a CD-ROM collection (Samper Padilla et al. 1998). The original recordings, mostly made with reel-to-reel tape recorders, have never been centrally archived or digitized; although copies of some of the recordings can be readily located, in most cases the published transcriptions cannot be compared with the original recordings.

30.2.2 European Spanish

Sources of information about dialect variation in Spain include Moreno Fernández (2009), Zamora Vicente (1967), and the articles in Alvar (1996a). Numerous traditional monographs have also been published on the traditional speech of single villages or rural sectors, usually


### 30.2.3 Latin American Spanish


Linguistic atlases are available for Colombia (Instituto Caro y Cuervo 1981), Costa Rica (Quesada Pacheco 2010), Mexico (Lope Blanch 1990), Nicaragua (Chavarría Ubeda 2010), and New Mexico and southern Colorado (Bills and Vigil 2008). Norma Culta transcriptions for Spanish American cities include Caravedo (1989) for Lima; Gutiérrez Marrone (1992) for La Paz; Heras Poncela (1999) for Guadalajara; Lope Blanch (1971, 1979) for Mexico City; Martorell de Laconi, Hondrogianis, and Soto (2000) for Salta, Argentina; Morales and Vaquera de Ramírez (1990) for San Juan, Puerto Rico; Rabanales and Contreras (1979) for Santiago; Rodríguez Cadena (2009) for Barranquilla, Colombia; Rosenblat and Bentivoglio (1979) for Caracas; and Universidad Nacional de Buenos Aires (1987) for Buenos Aires.

### 30.2.4 Spanish in Africa and Asia

For the Spanish of Equatorial Guinea, see Lipski (1985, 2000) and Quilis and Casado-Fresnillo (1995); for that of the Philippines, see Lipski (1987) and Quilis and Casado-Fresnillo (2008); for Western Sahara, see Tarkki (1995); for Morocco, see Sayahi (2004, 2005, 2006).

### 30.3 Portuguese: Main Sources

#### 30.3.1 General

Major corpora of written Portuguese include the *Corpus do português* (www.corpusdoportugues.org); the Reference Corpus of Contemporary Portuguese (CRPC) (www.clul.ul.pt); the Tycho Brahe project, a searchable database of historical Portuguese texts (www.tycho.iel.unicamp.br); and the Colonia Corpus of Historical Portuguese (corporavm.uni-koeln.de/colonia/).

#### 30.3.2 European Portuguese

In Portugal, the *Atlas Linguístico-Etnográfico de Portugal e da Galiza* (ALEPG) began collecting materials in the 1970s, and representative data are found in Ferreira *et al.* (2008). The Azores are represented by the *Atlas Linguístico-Etnográfico dos Açores* (www.culturacores.azesores.gov.pt/alea/).
30.3.3 Brazilian Portuguese


30.3.4 Portuguese in Africa and Asia


30.4 Spanish Dialect Zones and Characteristics

30.4.1 Spain

The principal division of Iberian Spanish dialects is north-south, with the northern dialects of Castile (including the capital, Madrid), León, Cantabria, the Basque Country, Aragon, and Spanish-speaking areas of Catalunya differentiated as a group from the southern varieties of Extremadura, Andalusia, and Murcia. The Canary Islands constitute a separate dialect cluster, most closely related to western Andalusian Spanish. The Spanish of Gibraltar is essentially that of the neighboring Spanish province of Cádiz, modified by contact with English. Modern standard European Spanish, based mostly on the Castilian dialect and therefore often called Castellano, is promulgated by the Real Academia Española.

30.4.1.1 Phonetics and Phonology

Most of Peninsular Spain, except for the southwestern provinces and part of the Valencia/Alicante region, distinguishes the sibilant phonemes /θ/ and /s/, for example, caza [‘ka.θa] “hunting” versus casa [‘ka.sa] “house.” In regions where this contrast is neutralized, [s] predominates in urban areas and [θ] in surrounding provincial regions. In northern Spain, /s/ receives an apico-alveolar realization [s]. In most of Spain, the opposition /ˈʃ/ - /ʃ/ has been neutralized in favor of /ʃ/, but /ˈʃ/ is retained in parts of northern Spain and sporadically elsewhere in the Peninsula as well as in much of the Canary Islands. Word-final /n/ is velarized to [ŋ] in the northwest and in Extremadura, Andalusia, and the Canary Islands. In the southern regions, syllable- and word-final /s/ is aspirated to [h] or elided. In eastern Andalusia, loss of word-final /s/ is reflected in laxing of the preceding vowel. The singular-plural distinction is therefore maintained through vowel quality, as in perro [‘pe.ro] “dog” versus perros [‘pe.rs] “dogs.” In southern Spain and the Canary Islands, word-final /l/ and /ɾ/ are routinely elided, whereas in preconsonantal contexts the two liquids are frequently neutralized, most often in favor of [ɾ].
Handbook of Dialectology

30.4.1.2 Morphosyntax

Most of Peninsular Spain employs le and les as masculine direct object clitics; the etymological lo and los predominate in the southwest and the Canary Islands. There is an increasing tendency in Peninsular Spain to employ the present perfect tense rather than the simple preterite in contextual frames that do not include the moment of speaking, for example, mi amigo ha llegado ayer, lit. “my friend has come yesterday.”

30.4.2 Latin America

The most robust dialect classification of Latin American Spanish includes the following major regions: Mexico and Guatemala; Honduras, El Salvador and Nicaragua; Costa Rica; the Caribbean basin (Cuba, Puerto Rico, Dominican Republic, Venezuela, northern Colombia, and Panama); the interior of Colombia; the Pacific coast of Colombia, Ecuador, and Peru; the highlands of Ecuador, Peru, Bolivia, and northwestern Argentina; Chile; Paraguay, eastern Bolivia and northeastern Argentina; central and southern Argentina and Uruguay. Most varieties of Spanish in the United States are directly related to those of immigrant populations’ home countries, principally Mexico, Puerto Rico, Cuba, the Dominican Republic, El Salvador and Guatemala, and Colombia (Lipski 2008a, Otheguy and Zentella 2011), but a small pocket of descendents of Canary Islanders is still found in southeastern Louisiana (Coles 1999, Lipski 1990).

30.4.2.1 Phonetics and Phonology

Unlike most of Spain, all varieties of American Spanish have a merger of /θ/ and /s/, realized as [s]. The principal phonological variable in Latin American Spanish is the opposition /s/ - /ʃ/ (e.g., se calló “he/she stopped talking” versus se cayó “he/she fell down.” This opposition is maintained in Paraguay, Bolivia, northeastern Argentina, a few areas in central Colombia, and most of highland Peru and Ecuador (in parts of Ecuador /ʃ/ is realized not as a lateral but as [ʃ]); elsewhere these sounds are merged as /ʃ/.

The major phonetic variables of Latin American Spanish involve the realization of /x/, /ʃ/, and the trill /ɾ/; and syllable- and word-final /s/, /n/, /l/, and /ɾ/. The posterior fricative /x/ is a weakly aspirated [h] in the Caribbean zone and the Pacific coast of Colombia, Ecuador, and Peru, whereas in Chile /x/ is a palatal fricative [ç] before front vowels, sometimes followed by a glide [j], for example, gente “people.” Intervocalic /ʃ/ is weak and often elided in contact with front vowels in coastal Peru, Central America, northern Mexico, and New Mexico and Colorado. In most of Argentina and Uruguay, /ʃ/ has traditionally been realized as a voiced fricative, [ʒ], but devoicing to [ʃ] has extended from Buenos Aires and Montevideo to much of the remaining territory. The nominally trilled /ɾ/ receives a fricative realization, [ʃ], throughout the Andean highlands and much of northern Argentina, and variably in Chile, Paraguay, Guatemala, and Costa Rica (in Costa Rica alternating with retroflex [ɹ]). Final /s/ is variably aspirated or elided throughout Latin America except in the Andean highlands and most of Mexico, Guatemala, and Costa Rica. Word-final /n/ is velarized to [ŋ] in the Caribbean, Central America, the Pacific coast of Colombia, Ecuador, and Peru, and variably in the Andean highlands. Syllable-final /l/ and /ɾ/ are variably neutralized throughout the Caribbean dialect cluster and in parts of Chile.

30.4.2.2 Morphosyntax

Latin American differs from European Spanish in the absence of the informal second-person plural pronoun, vosotros, and the corresponding object clitic, os; ustedes is the sole second-person plural pronoun. The second-person singular familiar pronoun is tis (instead of tis) in all of Argentina, Paraguay, Bolivia, most of Uruguay and Chile, certain regions of Colombia,
Ecuador, and Venezuela, and vestigially in Peru. *Vos* predominates in all of Central America, as well as in western Panama, parts of the Mexican state of Chiapas and small pockets of Cuba. The verb forms that accompany *vos* vary widely, ranging from reflexes of Peninsular Spanish *vosotros* conjugations to the inflexions corresponding to *tú* (Páez Urdaneta 1981).

Most Latin American varieties of Spanish employ the masculine direct object clitics *lo* and *los*; use of *le* and *les* for direct objects is characteristic of Paraguay and Ecuador and variably of Mexico. In the Southern Cone, direct object clitics can accompany animate direct object nouns (e.g., *Lo conozco a Juan* “I know John”) and in the Andean highlands clitic doubling occurs freely with all direct objects (e.g., *Lo veo el carro* “I see the car”). In much of South America the past subjunctive in subordinate clauses is replaced by present subjunctive, for example, *El profesor me aconsejó que estudie* [not *estudiaría*] mucho (“The teacher advised me to study a lot”). In the Caribbean dialect cluster and sporadically elsewhere, infinitives with overt subjects are used in preference to subjunctive forms: *La fiesta empezó antes de yo llegar* (“The party began before I arrived”) [not *antes que yo llegara*]. Also found throughout the Caribbean is the non-inversion of subject pronouns and verbs in questions, for example, *Cómo tú te llamas?* (“What is your name?”), or, *¿Dónde nosotros podemos comer?* (“Where can we eat?”). Found in much of Colombia, Venezuela, Ecuador, and the Dominican Republic (and also in Brazilian Portuguese) is the affirmative use of intrusive *ser* (“to be”): *Lo conocimos fue en la fiesta* (“(where) we met him was at the party”); *Tenemos es que trabajar mucho* (“We have to work a lot”).

### 30.5 Portuguese Dialect Zones and Characteristics


#### 30.5.1 Portugal, Africa, and Asia

The principal dialect division in Portugal is north versus south, with an approximate transition to the north of Coimbra. Lisbon is located in the southern zone and is the model for standard European Portuguese. African and Asian dialects generally follow a southern model, but since most are spoken in contact with other languages and often as a second language, there is considerable within-country variability. Audio samples of the principal varieties of Portuguese are available at cvc.instituto-camoes.pt.

#### 30.5.1.1 Phonetics and Phonology

The northern varieties retain /tʃ/ (e.g., *chave* “key”) in opposition to /ʃ/ (e.g., *caixa* “box”), a distinction neutralized as [ʃ] elsewhere. The diphthongs *ei* [ej] and *ou* [ow] are retained in the north but reduced to simple vowels in central and southern Portugal. In greater Lisbon, *ei* is realized as [ɛj]. Atomic /a/ is realized as [e], as is /a/ before nasals; atomic /o/ is raised to [u], and atomic /i/ and /e/ are generally neutralized to a high unrounded vowel [ɯ] and sometimes devoiced or elided. Northern Portuguese has lost the /b/ - /v/ distinction. The treatment of sibilants also varies by region. In an area that includes much of Tras-os-Montes, Alto Minho, and Beira-Alta, a four-way distinction is maintained, for example, /s/ (cego “blind”), /z/ (fazer “do, make”), /ʃ/ (senhor “sir,” passo “step”), and /ʒ/ (coisa “thing”). In parts of Minho, Douro, Beira Alta, and Beira Baixa, the two sibilant points of articulation have been neutralized and only apico-alveolar sibilants /ʃ/ and /ʒ/ are found, whereas in the remainder of the country only alveolar-dental sibilants /s/ and /z/ are used (Azevedo 2005: 186; Cintra 1995: 28). In most of Portugal, syllable- and word-final /s/ and /z/ are palatalized to [ʃ] and [ʒ], respectively. The dialects of the Azores and Madeira Islands generally follow southern Portuguese patterns. In Madeira, /l/ is palatalized after [i] and [j] as
Handbook of Dialectology

in vila “village,” stressed /a/ is often realized as [ɔ] (e.g., casa “house”) and the vowel /u/ is fronted to [y] (e.g., tudo “all”).

30.5.1.2 Morphosyntax
European and African Portuguese have generally substituted the infinitive for the etymological gerund in progressive constructions: estou a trabalhar “I am working” instead of estou trabalhando.

30.5.2 Brazil
Dietrich and Noll (2004) and Ilari and Basso (2006) describe dialect variation in Brazil. Brazilian linguists generally distinguish the following general dialect zones: the Northeast; Bahia; Mineiro (centered on Minas Gerais state); Fluminense/Carioca (centered on Rio de Janeiro); Paulista (centered on São Paulo); southern or Gaúcho; and the Amazonian region. No one city is considered the model for standard Brazilian Portuguese, but the educated speech of São Paulo and Rio de Janeiro comes the closest, although the palatalization of syllable-final /s/ in Rio is not emulated by other speakers. Brazilian Portuguese in general differs systematically from European varieties in several ways.

30.5.2.1 Phonetics and Phonology
Atonic /e/ and /o/ are raised in Brazil to [i] and [u], respectively. Prevocalic /b/, /d/, and /g/ are always occlusive, not fricative or approximant as in Europe, and /t/ and /d/ become pre-palatal affricates before /i/ (tio > [ʧiy] “uncle,” dia > [ʤia] “day”) except in the southernmost states. A velar fricative [x] replaces the trill [r] both in intervocalic position (carro “car”) and syllable-finall (por favor “please”), although in word-final position /r/ is frequently elided. Throughout Brazil, syllable-final /l/ is vocalized to [w] thus Brasil [bra.’ziw] as opposed to the velarized [l] found in European varieties. Most Brazilians insert a post-vocalic glide [j] in stressed syllables ending in /s/, thus nós [nɔjs] (“we”), or faz [fais] (“do,” 3s), and an epenthetic vowel after borrowed words ending in consonants other than /r/, /l/ or /s/, thus Nova York(i); the same occurs word-internally after syllable-final stop consonants, for example, adjilvogado (“lawyer”). Intervocalic /n/ is pronounced as a nasal glide [ŋ], for example, senhor (“sir”). Among regional dialects, the Carioca (Rio de Janeiro) variety is noted for palatalizing syllable-final /s/ to [ʃ] ([ʒ] before voiced segments). Palatalization also occurs in some northeastern dialects (e.g., Ceará), but only before dental and alveolar consonants. In rural São Paulo state, the Caipira dialect is noted for having a retroflex syllable-final [ɾ], for example, in porta (“door”). In casual speech, loss of the final /r/ in verbal infinitives is widespread, as is the delateralization of /ʎ/ to [ʃ] (mulher > muié[r], “woman”).

30.5.2.2 Morphosyntax
Spoken Brazilian Portuguese eschews object clitics in favor of full pronouns, for example, vejo ele (“I see him/it”; cf. European vejo-o). Progressive verb forms use estar+gerund, for example, estou falando (“I am talking”), rather than the European estar a + infinitive (estou a falar). In spoken Brazilian Portuguese, a gente (“the people”) is used pronominally in preference to nós “we,” and for the majority of the country (except for the southernmost states) você “you” is the familiar second-person pronoun instead of European tu; formal address is achieved with o senhor/a senhora (“sir, madam”). European Portuguese vós is not used; only vocês expresses second-person plural. Double negation (repeating não before and after the verb) is very common, especially when a negative response is being emphasized, and in casual speech the first instance of não often disappears, for example, (não) tenho não (“I don’t have”). Sei não alternates with sei lá (“I don’t know”).
In casual, vernacular speech, plural /-s/ is placed only on the first element of plural noun phrases, for example, _aquelas coisa(s) nova(s)_ (“those new things”). Also found in much vernacular speech, though condemned by prescriptivists, is the use of the third-person singular verb form for all other person-number combinations except first-person singular, for example, _vocês foi lá_ (“you (pl.) went there,” cf. standard _foram_), or _nós trabalha na cidade_ (“we work in the city,” cf. standard _trabalhamos_). In many vernacular varieties, the final /-s/ of first-person plural verbs is not pronounced, and in verbs of the first conjugation (in –ar), the ending becomes –emo, for example, _nós trabalhemo/trabalhamo na cidade_.

### 30.6 Current Research Trends

A survey of recent trends in Spanish and Portuguese dialectology is given in Lipski (2008b; also 1989). Earlier taxonomic and rural-oriented studies have given way to approaches that focus on dynamic urban environments with particular emphasis on socio-phonetic variation. The _Proyecto para el Estudio Sociolingüístico del Español de España y de América_ (PRESEEA; at preseea.linguas.net) is a coordinated effort to produce and consolidate sociolinguistic corpora representative of regional and social variation throughout the Spanish-speaking world.

Dialect variation has also informed linguistic models. Regionally distributed syntactic phenomena such as null subjects, non-inverted questions, and double negation have resulted in expanded theories, for example, by Bosque (1999b), Camacho (2006), Duarte (1995), Kato (2000), Ordóñez and Treviño (1999), and Toribio (2000). Regionalized pragmatic features such as politeness strategies have entered the picture, for example, the studies in Placencia and García (2006). Phonological theory has also been applied to dialect variation (e.g., Hualde 1989, Lipski 1999, Morris 2000).

### 30.7 Future Research

Vital to both the Iberian Peninsula and Latin America is the updating of now outdated dialect atlas materials to reflect the increasingly urbanized linguistic ecology of the twenty-first century. In addition, vast areas of Latin America are lacking accurate descriptions altogether. At least two other promising directions for future research can be mentioned.

The first is the systematic study of regional and social patterns of intonation, particularly in spontaneous speech. A consistent framework for the study of Spanish and Portuguese intonation is now widely accepted (e.g., Beckman _et al._ 2002, Frota 2013, Sosa 1999, Truckenbrodt _et al._ 2008), and many sentence-types have been classified for both European and American varieties of Spanish and Portuguese. Comparative studies are beginning to move beyond idealized and laboratory-elicited patterns to search for the defining traits of entire dialect regions.

Another facet of Spanish and Portuguese dialect variation is stable bilingual contact as a source of dialect features. There are studies of Ibero-Romance varieties that have arisen from sustained dialect or language contact, for example, Mirandese (Quarteu and Frías Conde 2002) and Barranquenho (Alvar 1996; Clements 2009: Chap. 8) along the Portugal-Spain border and “Fronterizo” in northern Uruguay (Elizaincín _et al._ 1987, Elizaincín 1992). In many other cases, analyses of regional and social dialects are handled separately from accounts of “interference” or code-switching in bilingual contact environments, even in cases where hundreds of thousands of speakers are involved. Bilingualism is a fact of life for millions of speakers of Spanish and Portuguese, making the integration of the dialectal traits of bilingual speech communities into general accounts of Spanish and Portuguese dialect variation a crucial step.
REFERENCES


García Mouton, Pilar, and Francisco Moreno Fernández. *Atlas lingüístico (y etnográfico) de Castilla-La Mancha*. www2.uah.es/alecman/


