

BASQUE

R. L. Trask

Language Name: Basque. **Autonym:** *euskara*. (Dialect variants *euskera*, *eskuara*). In Spanish: *vasco* or *vascuence*. In French: *basque*.

Location: The Basque-speaking area (called by Basques *Euskal Herria*) straddles part of the border of northern Spain and southwestern France, along the Bay of Biscay. It extends for about 100 miles from west to east and for about 30 miles from north to south. In Spain, it encompasses the provinces of Navarre, Bizkaia, Gipuzkoa, and Álava. In France it includes the former provinces of Labourd, Basse-Navarre, and Soule.

Family: None. Basque is genetically isolated. Specialists are satisfied that there is no shred of persuasive evidence that Basque is discoverably related to any other living language, and the not infrequent statements to the contrary in the non-specialist literature may be safely disregarded.

Related Languages: None, apart from its ancestral form Aquitanian (see below).

Dialects: The dialectal diversification is fairly substantial, enough to impede communication to some extent. Linguistically, however, the differences are quite superficial, consisting chiefly of differing vocabulary and some rather low-level phonological rules. Non-specialists writing about the language have often greatly exaggerated the degree of dialectal diversification.

The current classification is that of the Basque linguist Luis Michelena, who recognized nine dialect areas; however, the ninth, Roncalese, in the Roncal Valley in Pyrenees, has become extinct within the last few years. One other dialect, called Southern, is attested to in the sixteenth century but is long extinct.

The contemporary dialects are, roughly from west to east (locations cite modern Spanish provinces and former French provinces now incorporated into the *département* of Pyrenees-Atlantique): (1) Bizkaian (or Vizcayan) (B), in Bizkaia and western Gipuzkoa; (2) Gipuzkoan (or Guipuzcoan) (G), in most of Gipuzkoa; (3) High Navarrese (HN), in northern Navarre and eastern Gipuzkoa; (4) Lapurdian (or Labourdin) (L) in Labourd; (5) Low Navarrese (LN), in Basse Navarre; (6) Salazarese (S), in Salazar Valley in Pyrenees; (7) Aezkoan (A), in Aezkoa Valley in Pyrenees; and (8) Zuberoan (or Souletin) (Z), in Soule.

Following are a few of the more prominent dialectal differences: The French Basque varieties (L, LN, and Z) retain the consonant /h/ as well as distinctively aspirated plosives; the aspiration was long ago lost in all Spanish varieties. Z and R retain distinctively nasalized vowels; these have been lost in all other varieties. Z (only) has acquired a contrastive front rounded vowel and contrastive voiced fricatives. The ordinary main verb 'have' is *ukan* in the French varieties but *eduki* in the Spanish ones. There is considerable lexical differentiation between Basque dialects. For example, 'speak' is *berba egin* in B, *itz egin* in G, and *mintzatu* in the French varieties.

There is no significant social variation, save that educated speakers tend to adjust their speech toward the norms of the new standard language and to prefer indigenous words and neologisms to loanwords where possible.

Number of Speakers: 660,000 according to the 1991 census. Fewer than 80,000 of these are on the French side. This figure includes a few thousand people who speak Basque as a second language.

Origin and History

Basque is unquestionably the last surviving pre-Indo-European language in western Europe. In Roman times, while most of Gaul was occupied by Celts, the southwest (roughly the medieval Duchy of Aquitaine) was occupied by a quite distinctive people called Aquitanians. Their Aquitanian language is preserved only in the form of some 400 personal names and 70 divine names in Latin texts, but many of these names are so transparently Basque that specialists are satisfied that Aquitanian (Aq.) was an ancestral form of Basque (Bq.). Examples: Aq. *Cison*, Bq. *gizon* 'man'; Aq. *Andere* Bq. *andere* 'lady'; Aq. *Nescato*, Bq. *neskato* 'young girl'; Aq. *Sembe*, Bq.

seme 'son' (< **senbe*); Aq. *Ombe* and *Vmme*, Bq. *ume* 'child' (< **unbe*); Aq. *SA.HAR*, Bq. *zahar* 'old'; Aq. *Oso-*, Bq. *otso* 'wolf'. The phonological system of Aquitanian is strikingly similar to that of the pre-Basque of the Roman era reconstructed independently.

Aquitanian is only sparsely attested south of the Pyrenees in Roman times, although it was probably spoken in at least part of the modern Spanish Basque Country. Toponymic evidence shows that it must also have been spoken in the Pyrenees as far east as the valley of Arán, in territory which today is CATALAN-speaking.

There is clear evidence that Aquitanian/Basque co-existed with at least one Indo-European language in Roman times. As

Table 1: Consonants

	Labial	Dental	Alveolar	Retroflex	Palato-Alveolar	Palatal	Velar	Glottal
Stops	p b	t d		ʈ ɖ			k g	
Aspirated	ph bh	th dh		ʈʰ ɖʱ			kh gh	
Fricatives			s z		ʃ			h
Affricates						c j		
Aspirated						ch jh		
Nasal	m		n				ŋ	
Lateral			l					
Flap			r	ɽ				
Glide	w					y		

Politically, the Bengali dynasties of Pal (750–1162), Sen/Barman (1096–1228) and Dev (1170–1260) (years sometime overlapping due to separate independent kingdoms) ruled before the foreign invasion began with the Turkish (1204–1338) and continued with the independent Muslim Sultans (1338–1538), Moghuls (1575–1757) and the British (1757–1947). During the Muslim rule of Bengal the lexicon of Bangla drew heavily on SANSKRIT and (Turko)-Perso-Arabic sources, as the administrative language of the region was PERSIAN.

The earliest specimen of Bangla is found in a set of 47 songs called *Caryas* written by the teachers of a Mahayana Buddhist sect around the first millennium (950–1200). The manuscript was discovered in Nepal in 1907 (and later published in 1916), where some of the scholars possessing older manuscripts escaped after the Turkish invasion of Bengal in 1200—a period marked by destruction of manuscripts and persecution of their authors or possessors.

Fortunately, there is a copious Middle Bangla literature of which the earliest and most important is *Sri-Krishna Kirtan* of Chandi-dasa belonging to the latter half of the 14th century. Late Middle Bangla witnessed the development of the *Vaisnava* literature through the influence of the saint Chaitanya (1483–1533). This period also saw the development of an artificial literary language *Braja-Buli* (the speech of the Vraja, the birthplace of Krishna), which was a curious mix of Maithili, Bangla and Western HINDI.

New or Modern Bangla evolved a prose style due to western influence. Its greatest exponent was Rabindranath Tagore (the English spelling for the surname Thakur) who won the Nobel prize in literature in 1913.

The British departure in 1947 resulted in the partition of the subcontinent where West Bengal was left as part of India and East Bengal as the sole province of East Pakistan. In 1971 the war to free Bangladesh from Pakistan cost the lives of an estimated 3 million speakers of Bangla.

Orthography and Basic Phonology

The Bangla script is originally derived from the Brahmi script current from 400 B.C. to A.D. 300, a derivative of the Semitic

branch of ARAMAIC. But it was reshaped by the dominant Devanagari of the Northern and Western India from the seventh century onwards. Below is a sample of five Bangla alphabets on the extreme right as derived from the original Brahmi on the extreme left.

𑀓 𑀔 𑀕 𑀖 𑀗 𑀘	<ɔ>
𑀙 𑀚 𑀛 𑀜 𑀝	<k>
𑀞 𑀟 𑀠 𑀡	<c>
𑀢 𑀣 𑀤 𑀥	<ch>
𑀦 𑀧 𑀨 𑀩	<n>

From the Middle Ages, the Indian notion of *Akṣaras*, or the system of words being divided into syllables not closed by a consonant, was always treated somewhat mystically and was a result of the system of writing. The number of primary letters (or *maatrakṣara*) in Sanskrit, and following that in the vernacular languages, is 50, with 16 vowels and 34 consonants. Vowels include the retroflex /ɳ/, the vocalic /ri/, the *Anusvara* or voiceless /n/ and the *Visarga* /hi/ which are not phonemic (apart from the ones used as part of the spelling system) in Bangla. The same holds for the consonant /kṣ/.

Spelling rules for Bangla are extremely unscientific due mainly to the vast gap between the script and the pronunciation, a result of the heavy Sanskritization of Bangla since the twelfth century and carried on vigorously throughout the Middle Ages. For example, although Bangla does not distinguish between /s f ṣ/ (which are phonemically /ʃ/), primary school students are still taught the three different alphabets which in fact bear phonetically descriptive names, namely, *dantya* ('dental') 𑀓 for [s], *talbya* ('palatal') 𑀔 for [ʃ], and *muṛdhanya* ('retroflex') 𑀕 for [ʂ]. The script also marks contrasts now lost between short and long high vowels. That the

BANGLA

Tanmoy Bhattacharya

Language Name: Bangla. The name comes from the term *Bangaalah*, the name of the land comprising the states of West Bengal, part of Assam (both in India), and East Bengal (now Bangladesh) first used about 1200. **Alternate:** Bengali is the Colonial English designation for the language; however, the use of the term Bangla is becoming increasingly common in linguistic English. **Autonym:** *Bangla*.

Location: Bangla is the language of the Indian state of West Bengal and is also spoken in the states of Assam, Bihar, and Orissa. Bangla is the national language of the country of Bangladesh.

Family: Part of the Indic group of the Indo-Aryan (IA) branch of the Indo-European family of languages.

Related Languages: The bordering languages are ASSAMESE in the northeast, NEPALI in the north, MAITHILI and Magahi in the west, and ORIYA in the southwest. Several aboriginal languages and dialects also form part of its border. These are the non-related Austro-Asiatic languages SANTALI, HO and Mundari, and two Dravidian dialects Malto and Oraon (or Kurukh) to the west; a number of dialects belonging to six different groups of the Tibeto-Burman branch of the Tibeto-Chinese family to the North and East like Lepcha, Limbu, Banjong-ka (Sikkim), Lho-ke (Bhutan), Bodo, Garo, Maitei, Lushai, etc. Another Austro-Asiatic language, this time of the Mon-Khmer group, Khasi, is spoken to the east.

Dialects: Bangla is divided first into two main branches, Western and Eastern (or Vanga or Eastern Bangla). The Western branch is further divided into (i) Radha (South) (ii) Varendra (North Central) (iii) Kamrupa (North Bengal). Radha is subdivided into South Western Bangla and Western Bangla, the standard colloquial form of Bangla spoken around Calcutta.

Social Variety. Bangla is divided into codes, classifiable in terms of a High versus Low dichotomy called a diglossia. The H(igh) and L(ow) codes are differentiated in terms of the way the inflectional and pronominal systems in the *Sadhu* or H code of standard written Bangla contrast with those in the *colit* or L code. For example, the *Sadhu* and *colit* versions of the sentence 'They (distal)¹ (are) going²' are *tahara jaiteche* and *tara jacche* respectively [here and elsewhere, superscripts represent the sequence in the original]. The use of the H form for written prose became the norm for Bangla in the early nineteenth century as a result of that period's process of Sanskritization. It is rarely in use now. The L written norm uses the South Western morphology.

Number of Speakers: In the 1991 census, Bangla had 58,541,519 speakers in West Bengal (88,752 sq km) plus 11,054,219 other Indian speakers, and an estimated 107 million in Bangladesh (143,998 sq km), a total of 177 million subcontinental speakers in 1991, plus diasporic speakers for whom systematic figures are unavailable. The ethnolinguistic survey of LSA puts the total number of Bangla speakers at 189 million.

Origin and History

Around 600 B.C. outposts of the Aryan colonialization were being set up in the east inhabited by pre-Aryan tribes like Radhas and Vangas—ancestors of the people of Bengal. These were considered to be barbarous lands not suitable for high-caste Brahmins of the north and the midland to settle. Bengali toponymy suggests that these earlier tribes were speakers of Dravidian or some Austro-Asiatic language(s). However, by 700 B.C., a distinct Eastern branch (Pracya) of the Old IA was in existence, and by the fourth century Bengal was part of the well-established Aryan kingdom Magadh. The language spoken in these areas had certain distinct characteristics all of which are traceable in its current descendants. For example, the Old IA vowel [ʌ] became the default vowel [ɔ]; epenthesis with /i/, /j/ for IA /s f s/; interchangeability of /l/ and /r/; IA /kʂ/ became /kʰh/; using the affix -rɔ for the genitive -e, -ē for the instrumental, -e for the locative, -l- for the passive participle of the verbal noun and -b- for the future, and the roots for the auxiliary being *ho*, *ah*, *rah*, *ach*. This mother-dialect of the current

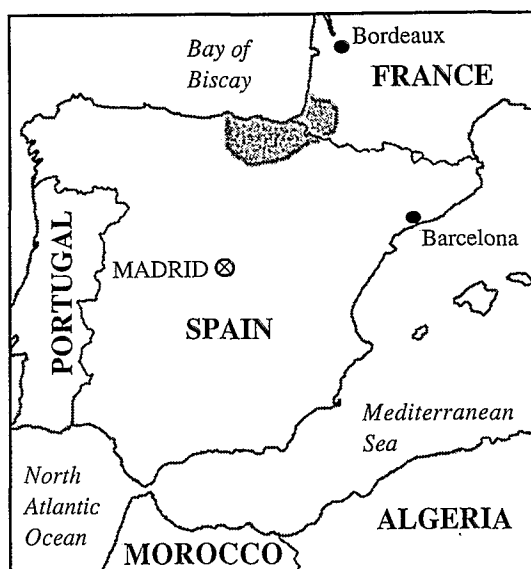
Eastern IA languages was called the *Maagadhii Apabhrañjaa* (MA), the second word meaning 'speech fallen from the norm'. This characterization is consistent with the depiction by the Sanskrit dramatists prior to 100 B.C. of lower caste characters speaking a /ʃ/ dialect. However, by the sixth century, MA was well-established and this is probably the language that the 7th century Chinese traveller Hiuen Tsang described when he said that the whole East India spoke a similar tongue. The following is a rough chronology of the different periods of Bangla:

1. **Pracya** ('Eastern'): 700 B.C.
2. **Middle IA:**
 - (i) Early Middle IA: 300 B.C.
 - (ii) Transitional Middle IA
 - (iii) 2nd Middle IA: A.D. 300
 - (iv) Late Middle IA: A.D. 800 (MA)
3. **Old Bangla:** 1100
4. **Early Middle Bangla:** 1400
5. **Late Middle Bangla:** 1600
6. **New Bangla:** 1800

Table 1: Consonants

	Labial	Dental	Alveolar	Palatal	Velar	Glottal
Stops	p, (ph) b	t, (th) d		tt dd, j	k, (kh) g	
Fricatives	f		s/z/x	j		(h)
Affricates			ts/tz/tx	j		
Nasals	m		n	ɲ		
Liquids			l r, rr	ll		
Glides				j		

far as we can judge, the western part of the modern Basque Country was entirely Indo-European-speaking. Some time after the collapse of the western Roman Empire, Basque was extended into what is now the western part of the Basque-speaking region, and also into much of the Rioja and Burgos, both of which were Basque-speaking in the early medieval period. Since that time, the language has been steadily losing ground to SPANISH (and Catalan) in the south and east. In the sixteenth century, the southern frontier of the language lay in the Ebro Valley; today, the language has disappeared from all but one small corner of Álava, from western Bizkaia, and from the greater part of Navarre. In the north, most of Aquitania was romanized rather early, but Basque survived in what is now the French Basque Country. The Basque-Gascon frontier in the north has been stable for many centuries, although French Basque is now under severe pressure from FRENCH.



Basque is spoken along the coast of the Bay of Biscay in northern Spain and southwestern France (shaded area).

Aside from the Aquitanian materials, the first documentation of Basque is the Emilian Glosses, usually dated to around 950. Thereafter, we find an increasing number of glosses, glossaries, individual words and phrases, complete sentences, brick inscriptions, personal names, place names, songs, verses, personal letters, and other materials. The first book published in Basque was Bernard Dechepare's *Linguae Vasconum Primitiae*, a collection of poems published in 1545. Since that time there has been a steady flow of publications in the language.

Orthography and Basic Phonology

Basque has always been written in the Roman alphabet. Before the 1960s, there was no standard orthography, and writers often used Romance spelling conventions supplemented by various additional devices to represent sounds absent from Romance. The modern standard orthography was promulgated by the Royal Basque Language Academy in 1964 and is now almost universally used.

The modern standard alphabet consists largely of the following letters: *a b c d e f g h i j k l m n n o p r s t u x z*. The letters *c q v w y* are not considered part of the alphabet, but they are used in writing foreign words and names. The digraphs *dd ll ts tt tx tz* represent single segments, but these are regarded as sequences of letters, not as single characters.

A notable feature of Basque orthography is that it rejects Romance conventions which are unmotivated in Basque. So the phoneme /k/ is always written *k* and never as *c* or *qu*; /g/ is always written as *g* and never as *gu*; /s/ is always written *z* and never *c* or *ç*. Hence we have the town names *Gernika* (Sp. *Guernica*) and *Zegama* (Sp. *Cegama*) and provinces *Bizkaia* (Sp. *Vizcaya*) and *Gipuzkoa* (Sp. *Guipuzcoa*).

The letter *h* is written where the French Basques have it word-initially or intervocalically, but not after a consonant. Thus the standard orthography writes *hau* 'this' and *ehun* '100' rather than the *au* and *eun* traditionally used by Spanish Basques, but it writes *ipar* 'north' rather than the *iphar* traditionally used by French Basques.