SANSKRIT LANGUAGE AND LITERATURE. The most important branch of the Indo-European family of language in Asia is Aryan or Indo-Iranian, with two main divisions: Iranian and Indo-Aryan. Languages belonging to the latter are spoken to-day by 250 million people in India, where they are the dominant languages except in the south, in Ceylon and the Maldive Islands, and throughout western Asia and Europe by colonies of Gypsies. As languages of administration they spread at one time far into central Asia where now is Chinese Turkestan; and as vehicles of Buddhism they have influenced through translation the whole of central and eastern Asia; Hindu civilization carried their vocabulary into the East Indies and the Malay peninsula. The oldest documents of Indo-Aryan are composed in Sanskrit. These are the Vedic texts. Their exact date is unknown, but it is probable that the oldest of them belong to the latter half of the second millennium B.C. They were probably composed before the Aryans had learned the use of writing; but being religious texts which it was essential to preserve unchanged they were handed down by an oral tradition which by various controls was made exceedingly exact. Even at the time of entry into India there must have been some dialectical differences in the language as spoken by the various invading tribes; with their further extension into India itself these differences became accentuated. The language of the Vedic texts shows a certain mixture, but is in the main founded upon the dialect of the north-west of India. With the advance of Aryan culture into the Punjab and the Gangetic plain the eastern dialects gained in importance; and this eastern influence, discernible in the earliest texts, constantly gains ground.

The most archaic of these texts is the Rig-veda, a collection of liturgical hymns; this is followed by the Atharva-veda, consisting chiefly of magical formulas, of prayers, curses and incantations. Considerably later come the first compositions in prose, commentaries on the Veda and philosophical treatises termed Brāhmaṇas and Upanishads. Although originally preserved as a religious language, Sanskrit was finally used for secular purposes. The earliest inscription in Sanskrit dates from 150 A.D., but it became the regular language of official inscriptions only in the 4th century A.D. But long before that the grammarians (of whom the most celebrated was Pāṇini in the 4th century B.C.) had fixed it as a learned language to which alone a strict interpretation confines the name Sanskrit, "the perfected," but which may conveniently be termed Classical Sanskrit as opposed to the Vedic or older form. As a literary language it is still cultivated; and a vast literature – philosophical, narrative, lyrical, dramatic, technical – has been written in it. Standing, in form at least, between the later Vedic and the Classical, but more nearly approaching the latter, is the language of the two great epics, the Mahābhārata and the Rāmāyaṇa, in which the influence of the spoken language can clearly be seen. For in the

meantime the spoken dialects continued to develop. Some of these, such as Pali and the Prakrits (*qq. v.*), were in turn crystallized and used as religious languages by new sects as well as for secular purposes before Sanskrit was so employed. From these vernaculars are eventually derived the modern Indo-Aryan languages (*q.v.*). Sanskrit, the literary language, although preserving the sound-system of Vedic practically unchanged, did not escape the influence of its descendants. The grammar was considerably changed, chiefly in a simplifying and normalizing direction; meanings of words were altered and developed, and vast quantities of new words were gradually absorbed after being given a Sanskrit form. On the other hand Sanskrit has continued to influence the spoken languages.

Sounds. The sound-system of Sanskrit consisted of:

(a) Fourteen vowels, viz.: Twelve simple vowels: *a, ā, i, ī, u, ū, ṛ, ṝ, ḷ, (ḹ), e, o.* Two diphthongs: *ai, au.*

(b) Thirty-six consonants, viz.: Five series of stops and nasals: guttural: *k, kh, g, gh, ṅ*; palatal: *c, ch, j, jh, ñ*; cerebral: *ṭ, ṭh, ḍ, ḍh, ṇ*; dental: *t, th, d, dh, n*; labial: *p, ph, b, bh, m.* Four semi-vowels: *y, r, l, v.* Three sibilants: palatal *ś,* cerebral *ṣ,* dental *s.* Two aspirates: voiced *h,* unvoiced *ḥ* (*visarga*). A nasal with loose closure of the lips *ṃ* (*anusvāra*), and another nasal *m̐* (*anunāsika*) which probably was a simple nasalization of the vowel rather than a consonant proper.

Vowels. The vowel-system was inherited practically unchanged from the common Indo-Iranian period, the only specific Sanskrit development being the change of the Indo-Iranian short diphthongs **ai, *au* (=Avestic *aē, ao*) to *e, o* (still classed as diphthongs by Indian grammarians) and of the long diphthongs **āi, *āu* to *ai, au.* Thus to Avestic *daēvo haomo uxδāiš gāuš* Sanskrit corresponds with *deváḥ somáḥ ukthāíḥ gáuḥ.* The comparative simplicity of the vowel-system and the great predominance of the vowels *a* and *ā* were due to the fact that in the Indo-Iranian period four distinct Indo-European sounds – *e, o, a, n̥* (*ṅ, n̥, m̥*) – had coalesced in the one sound *a* (similarly *e, o, ā* had all become *ā*). The only trace left of the earlier differentiation was the fact that an original guttural before *ă* representing older *ĕ* had became a palatal: thus we have *sácate,* "follows," from **sekʷetai* (Greek *hépetai,* Latin *sequitur*), but *sákṣat,* "he shall follow" beside Gk. *hépsetai.* Similarly the 12 I.E. diphthongs – *ei, oi, ai, eu,* **[955]** *ou, au, ēi, ōi, āi, ēu, ōu, āu* – were merged in four Skt. sounds *e, o, ai, au.* In this way the vowel alternation (especially between *e* and *o*)*,* a characteristic feature of Indo-European, was largely lost. Nevertheless, where the word contained a sonant the alternation, although

2

diminished in scope, was still discernible: *e.g.,* I.E. *ei : oi : i* became Skt. *e : i* (*e.g.,* Gk. *leípei : loipós : élipe* = Skt. *recati : rekaḥ : aricat*), I.E. *e, o : ē, ō* became Skt. *a : ā.* The correspondences thus resulting – *a : ā; i : e : ai; u : o : au; ṛ : ar : ār* – being associated with particular grammatical formations were still further developed in Sanskrit into a system, which, early recognized by the Indian grammarians, was used by them in their description of the formation of the language.

Consonants. The consonant-system has remained much truer to the original Indo-European. It is characterized by the rarity of spirants and by the opposition of unaspirated and aspirated stops (both surd and voiced) in each series. In preserving the voiced aspirate stops unchanged, Sanskrit and its descendants (for most of the modern Indo-Aryan languages still possess these sounds) are unique among the Indo-European languages, in which these sounds either became surd aspirates and later spirants (as in Greek and Latin) or lost their aspiration (as in Iranian, Balto-Slavonic, Armenian, Albanian, Germanic, Celtic). Thus to Greek *phérō,* Latin *fero,* Eng. *bear,* Sanskrit corresponds with *bhárāmi,* to Gk. *éthēke,* Lat. *fēcit,* Eng. *deed* with *ádhāt.*

Some consonants were restricted in their use: *ñ* appeared only before or after palatals, *ṅ* only finally or before gutturals (and where a guttural had subsequently disappeared), *ṇ* only between vowels and semivowels or before cerebral stops, *ḥ* only finally or before sibilants and surd gutturals or labials, *ṃ* only finally or before consonants. Neither aspirate nor *h* or *ś* ended a word. At the end of a sentence only the following consonants were used: *k, ṭ, t, p, ṅ, n, m, ḥ.*

Of the palatals *ch* appears as a single consonant only initially: elsewhere it is always doubled unless preceded by another consonant, for it corresponds to the Indo-European group sk_1: *e.g., chindánti* "they cut" = Latin *scindunt, gácchāmi* "I go" = Greek *báskō, vā ñchā* "wish" = Germ. *Wunsch. j* represents two I.E. sounds (I) palatal g_1 (= Gk. *g,* Avestic *z*): *jánaḥ* "birth" = Av. *zanō,* Gk. *génos,* Lat. *genus;* (2) velar g^w before an original *e* or *ĭ* (= Av. *j*): *jīváḥ* "alive" = Av. *jīvō,* Lat. *vīvos, cf.* Gk. *bíos,* Eng. *quick, jh* does not belong to the Indo-European part of the vocabulary, but appears only in onomatopoeic and borrowed words, or in words taken from the vernacular in which the frequent Sanskrit groups *dhy* and *hy* became (*j*) *jh.*

Of the sibilants *ś* corresponds to I.E. palatal k_1: in this Sanskrit agrees with the other eastern I.E. languages (Balto-Slavonic, Albanian, Armenian, Iranian) which have an *s* or *sh* sound as opposed to the *k* (*h* in Germanic) of the western languages: *e.g., śátam* "100" = Av. *satem,* Lithuanian *šiṁtas,* but Gk. *he-katón,* Lat. *centum,* Eng. *hundred.* Before the surd dentals this *ś* became *ṣ: viśáti* "enters", *viṣṭáḥ* "entered". But *ṣ* also corresponds to

I.E. *s* when preceded by *i, u, r* or *k,* agreeing in this with Iranian and partly with Armenian and Balto-Slavonic. The chief innovation in the consonant-system is the introduction of the third series, the cerebrals or linguals (better termed retroflex). The dentals were formed by the tip of the tongue striking the roots of the teeth, the cerebrals by the tip of the tongue, bent backwards, striking further back on the palate. In the oldest Sanskrit they are derived from the dentals when immediately preceded by *ṣ* or **ẓ* (which subsequently disappeared): thus **diśtas* (= Lat. *dictus)* became **diṣtas* and then *diṣṭáḥ;* **miẓdham* (= Gk. *misthós)* became **miẓḍham* and then *mīḍhám.* Later also under the influence of a preceding *r* or *ṛ* dentals became cerebrals: such words are loans from the popular dialect, but they begin to appear in the literary language even as early as the Rigveda: thus *víkaṭaḥ* "monstrous" is derived from *víkṛtaḥ* "strange". Some words, however, contain cerebrals which cannot be explained as derived from dentals in either of these ways. It is probable that these sounds were characteristic of both the families of languages, Muṇḍā and Dravidian, which the Aryan speakers of Sanskrit found in possession of India; and the appearance of them in Sanskrit and its descendants (and in a few of the adjoining Iranian languages such as Baluchi and Pashto) can scarcely be unconnected with this fact.

The Indian grammarians emphasized the difference in the pronunciation of consonants according as they came at the beginning, in the middle or at the end of a word or before other consonants. These differences, though slight, have ended in transforming the whole Sanskrit consonant-system in its descendants, the modern Indo-Aryan languages. Final consonants were pronounced without being fully exploded: in the spoken languages these final consonants had disappeared before the middle of the 3rd century B.C. A similar pronunciation is assigned to consonants before stops: in the spoken languages these were early assimilated (*e.g., suptáḥ* became *sutto). y* and *v* were pronounced more strongly initially than intervocalically: in the modern languages initial *y*- has remained or become *j,* but intervocalic -*y*- has been lost. [This contrast is observable even in the dialect of the Rigveda. For in the phonetically weak position of the termination and in certain accessory words intervocalic -*dh*-   has lost its occlusion and become -*h*-: *ihí* "go" = Greek *íthi, śemahe* "we lie" = Gk. *keímetha.* Further in all words intervocalic -*ḍ*- has become -*ḷ*-, although other more easterly dialects still preserve it (through whose influence it was afterwards re-established in Classical Sanskrit): thus Classical *nīḍaḥ* "nest" (from **niẓḍa-)* is *nīḷaḥ* in Vedic.]

Accent. Our knowledge of the accent of Sanskrit words is derived from its marking in the more important Vedic texts and from the statements of grammarians. It was predominantly a musical or tonic, not a stress, accent. Three different types are generally

4

distinguished, the *udātta* "raised", *anudātta* "unraised" and *svarita* the rising-falling accent following the *udātta*. Generally the position of the *udātta* agrees, as far as can be ascertained by comparison, with that of the chief word accent of Indo-European. The rhythm of Sanskrit was purely quantitative; in verse metre depended only on the number and the length of the syllables composing a line, and was entirely independent of accent. In this respect, the Yashts of Avestan literature are similar.

GRAMMAR

  Nouns. Sanskrit makes a perfect distinction between nouns (including adjectives, pronouns and indeclinables) and verbs. The declension of nouns comprises three numbers – singular, dual and plural; and eight cases, viz., nominative, accusative, instrumental, dative, ablative, genitive, locative, vocative. In the majority of declensions ablative and genitive are not distinguished in the singular, nor dative and ablative in the plural, while the dual universally has only three separate forms, – (1) nom.-acc.-voc., (2) inst.-dat.-abl., (3) gen.-loc. These confusions were inherited from Indo-European, but the tendency to confuse still further the forms and functions of the cases continued within Sanskrit itself, the genitive especially enlarging its sphere at the expense of other cases. This process, continued in the spoken languages down to the present day, has resulted in a noun declension, which for the most part consists of two cases only – (1) a direct case founded on the Sanskrit nominative and accusative; (2) an oblique case founded on the Sanskrit genitive or, perhaps, in some forms, the dative. Among the numbers the dual by itself was reserved to express natural pairs (such as *akṣī* "the eyes," *karṇáu* "the ears"), or any two objects or persons already referred to or present in the mind of the speaker or hearer; otherwise the use of the word for "two" itself was required.

  Although the idea of case was in the main expressed by the termination (as, *e.g.,* the nominative by *-s* in *devás* = Latin *deus,* the accusative by *-m* in *devám = deum),* it was sometimes accompanied by change in the pre-terminational element, and was connected with a shift of accent:

	Skt.	Greek	Skt.	Greek	Skt.	Greek
nom.	*pā́t*	*pṓs* (Doric)	*dyáuḥ*	*Zeús*	*pitā́*	*patḗr*
acc.	*pā́dam*	*póda*	*dyā́m*	*Zen-a*	*pitáram*	*patéra*
instr.	*padáḥ*	*podós*	*diváḥ*	*Diós*	*pitúḥ*	cf. *patrós*
dat.	*padé*	cf. *podí*	*divé*	cf. *Dií*	*pitré*	cf. *patrí*
loc.	*padí*	*podí*	*dyávi*	Lat. *Jove*	*pitári*	
voc.			*dyàuḥ*	*Zeũ*	*pitar*	*páter*

There was, however, one important class of stems, namely masc. **[956]** and neut. in *-a* and fem. in *-ā* (corresponding to the *-o* and *-ā* stems of Greek and Latin), in which the accent was fixed either on the stem or on the termination throughout the paradigm without any change in the form of the stem, of the type *áśvaḥ* = Gk. *híppos*, Lat. *equos*. There had, moreover, been a tendency in the pre-Sanskrit period to fix the form even of the variable stems, especially those consisting only of a root: thus the long vowel was generalized in *vāc-* "voice" (nom. *vā́k*):

	Skt.	Latin	Avestic	Greek
acc.	*vā́cam*	*vocem*	*vācim*	
gen.	*vācáḥ*	*vocis*	*vacō*	*opós*

But in the process of normalization the vowel declension (and especially with the stems in *-a -ā*) was predominant, and even in Sanskrit itself there was a considerable transference of consonant-stems and root-stems to this declension: *e.g.*, *pā́da-*"foot," *dvā́ra-* "door", *dánta-* "tooth" replaced *pād- pad-, dant- dat-, dvār- dur-*. This proceeded apace in the spoken tongue, for the declension of the modern languages rests exclusively on vowel-stems and of these chiefly on the *-a -ā* stems.

As in Indo-European, the pronominal declension differed from the nominal, that of the personal pronouns radically, that of the demonstratives and others to a varying degree. In the personal pronoun each number had a different root (*t[u]vám* "thou," *yūyám* "you"), while that of the first person had different roots also for the nominative and the oblique cases (*ahám* "I," *mā́m* "me," *vayám* "we," *asmā́n* "us"). The tendency of the two declensions, the nominal and the pronominal, to influence each other is observable: *e.g.*, the nominal ending of the neuter sing. nom. and acc. *-m* as in *kím* "what?" (= Lat. *quid*) replaced the *-d* or *-t* still found in other pronouns (*tát* "that" = Lat. *is-tud*). In the spoken language this mutual influence continued to react until practically all distinction between the two types of declension was lost. Perhaps the most satisfactory explanation of the various elements that stand between Indo-European demonstrative pronouns and their case endings is to be found in Hittite where the original pronominal stems were appended to the sentence connective (Skt. "tasmin" = Hit. "ta [sentence connective] + smi [pronoun]"; Sk. "tas" = Hit. "ta+as"), etc.

Numerals. The numeral-system, built upon a decimal basis, is that of Indo-European. There are separate names for the numbers up to 10; from 11 to 19 compounds of the

units with the word for 10 partly correspond with those of Greek and Latin: *e.g., d(u)ắdaśa* "I2" = Gk. *dōdeka*, Lat. *duōdecim*. Above that the tens (20, 30, etc.) were in origin probably compounds expressing a number of tens, and the intermediate numbers were formed by compounds of the units with these: *e.g., pañcá* "5" (Gk. *pente*, Lat. *quinque*), *pañcāśát* "50" (*cf.* Gk. *pentēkonta*, Lat. *quīnquāginta*), *páñcapañcāśat* "55". There are separate words for 100 and 1000. The special word for 100,000 (*lakṣaḥ*, whence modern *lākh*) is post-vedic. The numerals from 1 to 19 are adjectives; the rest are substantives.

Gender. Traces of the distinction of gender between animate (= masculine and feminine) and inanimate (= neuter) are seen in the use of the neut. *udakám* "water as a medium of floating, etc." and the fem. *ắpaḥ* "water as personified as sentient beings". But for the most part in Sanskrit gender is grammatical and largely independent of nature, except that male and female living beings are usually (though not always) respectively masculine or feminine. There were three genders – masculine, feminine and neuter. The neuter was distinguished by its termination (or lack of termination) in the nominative and accusative, and by the fact that the form of the accusative was the same as that of the nominative. But the masculine and feminine substantives were primarily distinguished only by the form of an adjective, if there were one, in agreement with them (as often, *e.g.,* in French or German). But there was a tendency to reserve certain types of stem for one or other gender. Thus the *-a* stems in Sanskrit are reserved only for masculine and neuter nouns (although in Greek they may still be feminine, *e.g., hē híppos* "the mare"); and the *-ā* stems are mostly feminine. In the older language the *-ī* stems might be masculine or feminine; in the later language they are almost all feminine. Similarly even in Classical Sanskrit *-i* stems are either masculine or feminine; but in the spoken language there was a growing tendency to confine them to feminine: thus words of this declension which are masculine in Sanskrit become feminines in the modern languages: *agníḥ* m. "fire" (=Lat. *ignis*) becomes Hindi *āg* f.

Verbs. In the Vedic language the verbal system is of considerable complexity. A verb might have various stems, viz. present (sometimes more than one), aorist (three in number, root, *e.g., ásthām* = Gk. *éstēn*, strong, *e.g., aricam* = Gk. *élipon*, sigmatic, *e.g., araikṣam* = Gk. *éleipsa*); perfect (characterized by reduplication and peculiar terminations); future (rare in the old language). The significance of the stems lay in the mode of action they expressed. Very generally speaking, the present indicated continuous action, the aorist momentary action, the perfect a state resulting from past action. The various present stems indicated various types of present-stem action, such as intensive, repetitive, inchoative, causative, desiderative, etc. Each of the first three stems had five moods – indicative expressing fact, injunctive and subjunctive expressing will

and futurity, optative expressing wish, imperative expressing command. In the indicative of the present, perfect and future stems there were two tenses, present and past; the aorist stem was in the indicative confined to the past tense. Each tense had three persons and three numbers – singular, dual and plural. Finally each tense could be conjugated in two voices with different terminations – active and middle. Among the parts of the infinite verb there was connected with each stem a participle which could be either active or middle, present, past, or perfect, and independent of tense stems a past participle, one or more infinitives (for the most part case-forms of verbal nouns), a gerundive (= future passive participle) and an indeclinable participle or gerund. The total number of possible forms belonging to any one verb is thus very great.

The idea of time was expressed in the indicative primarily by the termination: the primary endings expressing present and future time, the secondary endings past (and also, to some extent, future) time. Past time was usually still further defined by the prefixing of a particle, *a* or *ā*, before the verb (corresponding to the Greek and Armenian augment), e.g., *dádhāmi* "I place," *ádadhām* "I placed" = Gk. *títhēmi, étithēn*. The active voice indicated that the action performed had reference to some person or thing other than the doer, the middle that the action had reference to the doer in some way or other: e.g., active *taṇḍulā́n nenekti* "he washes the rice" (French *il lave),* middle *pāṇī́ nenikte* "he washes his hands" (*il se lave les mains*). As in Greek, the middle could also be used to express passive sense, but this use was later confined to the perfect and especially the aorist stems, while a special present stem (characterized by the suffix *-yá-* with middle terminations) was used for the present passive: e.g., *labhate* "he takes," *labhyate* "is taken". The expression of person was inherent in the verbal form, and pronouns were used only when emphasis was desired. As with the nouns, there were two types of stems – the athematic with variable accent and variable stem form to which the terminations were directly attached (type *ás-ti* "he is" = Gk. *esti,* Lat. *est*), the thematic with invariable accent and stem between which and the termination the vowel *a* was inserted (type *váh-a-ti* = Lat. *veh-i-t*). There was a growing tendency for verbs to be brought into this class from the other, since its conjugation was the more normal; thus *yunák-ti* (3rd plur. *yuñj-ánti*) "joins," contrasted with the already thematic Lat. *jung-i-t,* became later *yuñj-a-ti* (3rd plur. *yuñj-a-nti*). This type almost alone survives in the modern languages.

This verbal system was greatly simplified in Classical Sanskrit. The injunctive mood almost, and the subjunctive entirely, disappeared or were incorporated in the imperative. The aorist and perfect stems remained only in the indicative, and the aorist participle disappeared. The infinitive was reduced to one form only (the accusative of a noun in *-tu-,* corresponding in form to the Latin supine, e.g., *jñātum* "to know" = *[g]nōtum),* and

the absolute to two, one for simple, one for compound verbs. This left the present stem predominant; and though the aorist and the perfect survived in the indicative, their meaning was scarcely, **[957]** if at all, to be distinguished from that of the past tense of the present stem (*i.e.,* the imperfect). Finally even in Sanskrit itself the imperfect tended to be replaced by the past participle used as a finite verb. Among the spoken languages in some areas the aorist seems to have been developed further as a past tense, but as a whole the verb of the modern languages possesses a present system based on the Sanskrit present stem and a past system based on the Sanskrit past participle.

Syntax. In its main lines the syntax of Sanskrit is Indo-European, and the primary uses of its cases, moods and tenses can be paralleled from other Indo-European languages. Its most peculiar development was in the region of composition. Compounds, usually of not more than two members, are common in the Vedic texts. Even in the Epic their extension is considerable, and compounds consisting of three or four separate members are not rare. In the later, more artificial, language compounds may be met with which extend over a page or more.

Vocabulary. Sanskrit shows a constant change and development of vocabulary. Part of this is in the natural evolution of already existing material. The main constituent of the vocabulary of Vedic Sanskrit is Indo-European, but even here there are some words which are not Indo-European. One source of such words is undoubtedly the languages which the Aryans found in India; and recently J. Przyluski has demonstrated that some (such as *bāṇáḥ* "arrow," *mayūraḥ* "peacock," *karpāsam* "cotton") were received from *Muṇḍā* or Austro-Asiatic languages. Another source for new words was found in the spoken Aryan languages: words taken from these were given a Sanskrit form. In this way even original Sanskrit words returned under a new form: we have already seen how *víkṛta-,* having in the spoken language become *vikaṭa-,* was readopted in that form in the literary tongue, and retained side by side with *víkṛta-* (R. L. T.).