# 12
# English as an Indo-European Language

*Philip Baldi*

Like most of the more than 5,000 languages in the world, English belongs to a language family, that is, a group of languages that are related to each other genetically and share a common ancestry. The "genes" they share are inherited linguistic features which have been transmitted through time over the history of the languages in question. The notion of a language family is founded on the observation that two or more languages may contain features of lexicon (vocabulary), phonology (sound), morphology (word structure), and syntax (grammar) which are too numerous, too fundamental, and too systematic to be due to chance, to general features of language design (typology), or to borrowing through contact. The language family to which English belongs is known as the Indo-European (IE) language family, and the common ancestor from which the Indo-European languages derive is called Proto-Indo-European (PIE). The subgroup within Indo-European to which English belongs is Germanic, specifically West Germanic.

As we begin our exploration of English as an IE language, we will first spend some time discussing the methods by which languages are classified genetically, how these methods help us to separate linguistic structures that are inherited from those which are not, and how they are used to access the past, including the preliterary past, of languages such as English.

How do we know that languages share "genetic material," and are therefore to be grouped within the same language family? We begin with a few simple illustrations with languages which will be familiar to most readers.

Everyone knows that the "Romance" languages (such as French, Italian, and Spanish) are all in some way descended from Latin. What this means is that the Romance languages are all "sister" languages, and that they stem from a common ancestor, thereby forming a genetic group (more specifically a subgroup). We know this on independent factual grounds, based on the documented history of the Roman Empire and its spread throughout early Europe. But even in the absence of historical records tracing the spread of the Romans and their language in its various forms, we

*Table 12.1*   *One* through *ten* in some Romance languages

|        | Italian | French | Spanish | Portuguese | Rumanian | Latin    |
|--------|---------|--------|---------|------------|----------|----------|
| one    | uno     | un     | uno     | um         | unu      | unus     |
| two    | due     | deux   | dos     | dois       | doi      | duo      |
| three  | tre     | trois  | tres    | tres       | trei     | tres     |
| four   | quattro | quatre | cuatro  | quatro     | patru    | quattuor |
| five   | cinque  | cinq   | cinco   | cinco      | tint     | quinque  |
| six    | sei     | six    | seis    | seis       | sase     | sex      |
| seven  | sette   | sept   | siete   | sete       | sapte    | septem   |
| eight  | otto    | huit   | ocho    | oito       | optu     | octo     |
| nine   | nove    | neuf   | nueve   | nove       | nao      | novem    |
| ten    | dieci   | dix    | diez    | dez        | dzate    | decem    |

*Table 12.2*   *One* through *ten* in some Germanic languages

| English | Dutch | German | Swedish | Yiddish |
|---------|-------|--------|---------|---------|
| one     | een   | eins   | en      | eyns    |
| two     | twee  | zwei   | två     | tsvey   |
| three   | drie  | drei   | tre     | dray    |
| four    | vier  | vier   | fyra    | fir     |
| five    | vijf  | fünf   | fem     | finf    |
| six     | zes   | sechs  | sex     | zeks    |
| seven   | zeven | sieben | sju     | zibn    |
| eight   | acht  | acht   | åtta    | akht    |
| nine    | negen | neun   | nio     | nayn    |
| ten     | tien  | zehn   | tio     | tsen    |

would arrive at the same conclusion of linguistic relatedness through the comparison of the modern languages. Consider, for example, the lower numerals in selected "major" Romance languages (table 12.1), written in standard orthography (which may obscure features of pronunciation).

Of course the existence of similarities among these five Romance languages is easy to explain. They share a common ancestor language (Latin), and have inherited the lower numerals directly from this source; i.e., the words are "cognates" and the languages are "sisters." But there are equally compelling data from languages whose ancestor can only be inferred because, unlike Latin, it was never written down. Consider the modern members of the Germanic subgroup (table 12.2).

Despite the obvious relatedness and common ancestry in the Romance and Germanic examples just cited, such connections are not always obvious. And even when it is convincingly established that the languages in question are in some sort of historical relationship, it is by no means an easy step to determine what the ancestor

might have looked like, when and where it was spoken, or what other languages might be related, perhaps more distantly (i.e., as "cousins" rather than "sisters"). The Latin-Romance connection is deceptively simple because of what we know about the lines and stages of transmission between the historical end points (Latin and Italian, for example). The Germanic case is somewhat more difficult because of the absence of an attested ancestral language (there are older Germanic languages, such as Gothic or Old Icelandic, but these are not proto-systems). Nonetheless, the evidence for relatedness among these languages is just as powerful as with the Romance languages. We just don't have a written ancestor.

Are such resemblances enough to prove a genetic relationship among languages? Are we forced to conclude from these displays of vocabulary in a limited field (here, lower numerals) that the languages in each group are derivable from some common ancestor? Surely there are other explanations available to account for the likenesses – borrowing through language contact, for example. Languages exchange vocabulary without regard for family membership; need and prestige are the two primary factors which govern the borrowing process. The languages which make up the respective Romance and Germanic subgroups have been in close cultural and geographical contact for millennia, so might it not be conceivable that they all just borrowed the numbers 1–10 from one or the other of them, or perhaps some other language?

For the lexicon to be used even as a preliminary guide to possible genetic relationships, we need more examples of potential cognates than a few (admittedly impressive) sets of numerals. In particular we need vocabulary items which, like the numerals, are part of the "core" vocabulary, i.e., words which are unlikely to have been borrowed, and which exist in sufficient quantity to exclude the possibility of chance (see table 12.3).

Like the numerals, these words come from deep in the core of the lexicon. They are not technical terms, like *computer* or *fax*, nor do they represent culturally transportable items such as *pizza* or *sushi*. And there are countless numbers of sets like them, eliminating the factor of chance. The only reasonable way to account for these similarities is to treat the words as cognates, and to assume that they are derived from a common source. We call that source language "Proto-Germanic."

*Table 12.3*  Some "core" Germanic vocabulary

| English | Dutch | German | Swedish | Yiddish |
|---------|-------|--------|---------|---------|
| love | liefde | Liebe | ljuv "sweet" | libe |
| to live | leven | leben | leva | lebn |
| to fly | vliegen | fliegen | flyga | flien |
| hand | hand | Hand | hand | hant |
| house | huis | Haus | hus | hoyz |
| my, mine | mijn | mein | min | mayn |
| mother | moeder | Mutter | mo(de)r | muter |
| name | naam | Name | namn | nomen |

Classifying languages based on vocabulary similarities represents only the first step in the historical process. To complete the task, we have to take a closer look at the properties of the words we have assembled to determine the degree of systematicity which holds across the languages. If the languages are indeed related (as we know these to be), the correspondences in vocabulary should be matched by systematic correspondences in phonology and morphology as well (syntax is somewhat more problematic). The principle of regularity is the cornerstone of the comparative method, by which linguists reconstruct the parent language and its intermediate stages based on the comparative analysis of the descendant languages. So, if say English and Swedish are related, and if there is a correspondence such that Eng. /m/ corresponds to Swed. /m/ in a given phonetic environment, then it should be the case for every /m/ (see the examples for "my," "mother," "name"); likewise for /l/ (see "love," "live," "fly") or for /v/ and /b/ (see "love" and "live" in English and German). As we work out the details of such correspondence sets we make inferences about the ancestral sound, which in the first two cases would be postulated as *m and *l (with the * designating a hypothetical reconstructed segment). For every set of words in which Swed. /m/ corresponds to Eng. /m/ in a given phonetic environment, we claim that both derive from a common proto-sound *m in Proto-Germanic. The same principle holds as we work to progressively more distantly related languages, such as Latin and (Old) English, or Greek and Sanskrit, using the oldest available data as we work backwards in time, all the way to PIE. Needless to say the correspondences become less and less obvious with deeper time spans and the need for auxiliary explanatory mechanisms such as analogy and secondary sound change increases, but the method is sophisticated enough that it can reveal correspondences over millennia of distance in first attestation, say between Old English (ca. 600 CE) and Ancient Greek (ca. 800 BCE) or Hittite (ca. 1750 BCE).

## The Indo-European Language Family

The term "Indo-European" refers to a family of languages which by about 1000 BCE were spoken over a large part of Europe and parts of southwestern and southern Asia (see figure 12.1).

The dating and location of a unified PIE is controversial in many respects, but the most widely held opinion among specialists puts the protolanguage in the area of the Pontic-Caspian steppes north of the Black and Caspian Seas at about 3500 BCE, after which it began to diversify into the descendant subgroups through phases and stages which are matters of debate (more than a few locales and time horizons have been proposed). Though the concept of "Indo-European" is linguistic, the term is originally geographic, referring to the location of the easternmost (India) and westernmost (Europe) languages at the time the family was securely identified in the eighteenth and nineteenth centuries. In comparison with some of the other 250–300 language families of the world, the IE family is relatively small. It contains about 140 languages

Publisher's Note:
Permission to reproduce this image
online was not granted by the
copyright holder. Readers are kindly
requested to refer to the printed version
of this chapter.

*Figure 12.1* Distribution of Indo-European Languages, circa 500–1000 BCE (from P. Baldi (2002) *The Foundations of Latin*. New York: Walter de Gruyter; map 1, p. 37). Reprinted with permission

(many extinct), more than 90 of which belong to Indo-Iranian; these 140 or so languages are classified into 11 subgroups, one of which is Germanic, where English is located. By contrast, the Austronesian language family of the Pacific has some 800 languages in a large number of subgroups, and the Bantu family (Africa) has as many as 400 languages. Of course it is important to distinguish the number of languages in a family from the number of speakers, or the geopolitical importance of the languages in question (as evidenced by their status as second languages, or as a *lingua franca*). By these latter criteria the Indo-European family, specifically through the colonial and global languages such as French, Spanish, and especially English, has a unique standing among the language families of the world.

The family tree represents graphically some of the more important and recognizable members of the IE family (figure 12.2). We offer here a few words about each subgroup, its dating, and its overall importance for our understanding of PIE and its history.

## Anatolian

Completely extinct, the Anatolian languages were unknown until archeological excavations in Boğazköy, Turkey in the early twentieth century uncovered the royal archives of the ancient Hittite city of Hattušaš. The original trove of about 10,000 clay tablets (now about 25,000), dating from the seventeenth to the thirteenth centuries BCE, was deciphered from its cuneiform script and shown to be representing an Indo-European language now called Hittite. The discovery, classification, and eventual detailed analysis of the Anatolian languages, but especially Hittite, has impacted IE studies significantly. Before Hittite, PIE was reconstructed with a "look" that resembled the older IE languages, in particular Baltic, Slavic, Greek, Latin, and Sanskrit. But Hittite, though demonstrably older, does not share a number of structural features with the "classical" IE languages, and in many cases displays characteristics which can be shown to significantly predate those in other IE languages. Two of the more famous of these archaisms were the existence of several sounds (called "laryngeals") that had been lost in the other subgroups, often leaving a trace; and the absence of the "classical" three-way gender system (masculine-feminine-neuter) in substantives in favor of a two-way animate-inanimate system. Accounts of differences such as these between Hittite and the other IE languages have challenged the traditional look of reconstructed PIE and its chronology, prompting some scholars to view the Anatolian languages as sisters, rather than daughters, of PIE, with both descending from a more remote protolanguage called "Indo-Hittite."

## Indo-Iranian

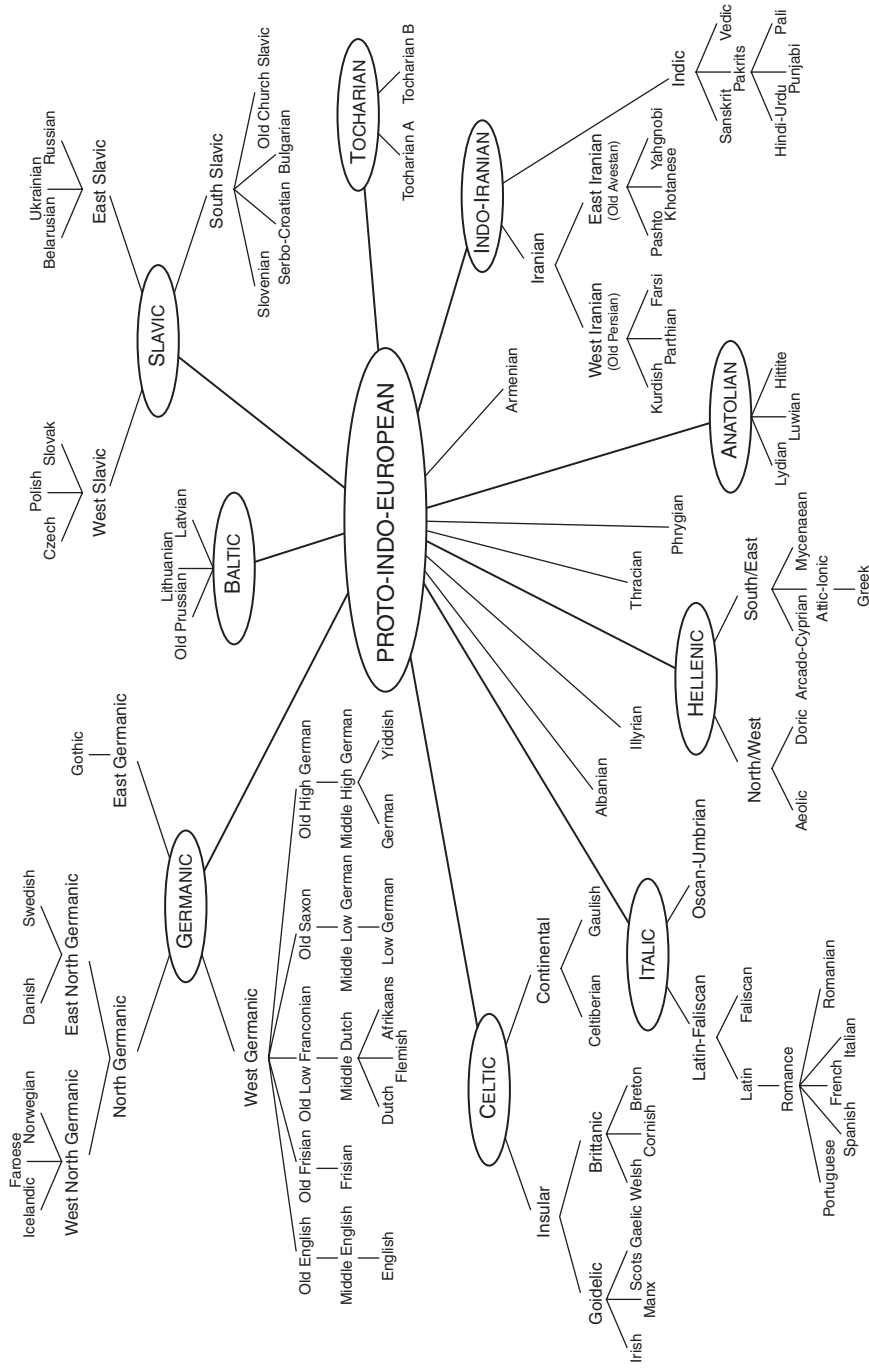This subgroup contains two closely related subdivisions, namely Indic (Indo-Aryan) and Iranian.

*Figure 12.2*  Major Indo-European branches and languages

### Indic (Indo-Aryan)

The languages of the Indic group are classified into three historical periods, namely Old Indic (1500–600 BCE), Middle Indic (600 BCE–1000 CE), and Modern Indic (since 1000 CE). The most ancient language is Vedic, an archaic form of Sanskrit whose oldest documents are dated by some to about 1200–1000 BCE, though others consider them to be older. A closely related form of Vedic is Classical Sanskrit, which was codified in the work of the grammarian Pāṇini ca. 500 BCE, and in which several important literary texts are written. The oldest Middle Indic texts are in Pāli (sixth to fifth century BCE), followed by the Aśoka inscriptions (ca. 250 BCE) and some Jainist religious writings from about the same period. Modern Indic is one of the largest and most heterogeneous of the IE subgroups, with perhaps as many as ninety different languages. Among the best known of them are Hindi-Urdu, Marathi, Punjabi, and Gujurati.

### Iranian

Ancient Iranian has two important representatives. The chief one of these is Old Avestan (also known as Gathic Avestan), dating from about 600 BCE, possibly earlier. The second important member of Ancient Iranian is Old Persian, a Western Iranian language, which may date to as early as 500 BCE. Western Middle Iranian is represented by Middle Persian (Pahlavi) and Parthian, while the Eastern Middle Iranian languages are Sogdian, Khotanese, Khorasmian, and Bactrian. Modern descendants of Iranian are Modern Persian (Farsi), Pashto, and Kurdish.

### Greek

Also known as Hellenic, the Greek branch contains some of the oldest testaments of Indo-European. Attested inscriptionally from as early as the eighth century BCE, Greek has textual monuments in the Homeric epics the *Iliad* and the *Odyssey*, which may be as old as 800 BCE. Even older than these are the Linear B tablets from Crete, Pylos, and other ancient locales which represent a form of Greek called Mycenaean and may be from as far back as the fourteenth century BCE. The two principal subdivisions are between South/East Greek (comprising Attic-Ionic, Arcado-Cyprian, and Mycenaean), and North/West Greek (comprising Aeolic and Doric). The main dialect of Greek is Attic, the literary language of Athens in which standard Classical Greek literature was composed. Standard Modern Greek developed from Attic-Ionic.

### Italic

The Italic subgroup of Indo-European consists of many genetically connected languages from ancient Italy which share certain distinctive characteristics. There are two main Italic subdivisions. The more important of the two, Latin-Faliscan, is

represented chiefly by Latin, one of the most important IE languages and arguably the most important language in the development of Western Civilization. Latin is identifiable in some short inscriptions from the seventh century BCE, though the first continuous literature stems from the third century BCE. Faliscan is known only from inscriptions, the oldest of which dates to the early seventh century BCE. Latin survives in the modern Romance languages, which developed from spoken varieties of the language in various parts and at different times and social circumstances in the history of the Roman Empire and beyond. The best known of the Romance languages are Italian, French, Spanish, Portuguese, Catalan, and Rumanian; less well known are Dalmatian, Rhaeto-Romansch, Ladino, Sicilian, Sardinian, Occitan, and many other local and social varieties. The second Italic subdivision is called Osco-Umbrian (also Sabellic or Sabellian). There are no modern descendants of this branch, which comprises Oscan (attested as early as the fifth century BCE), Umbrian (perhaps as early as 300 BCE), South Picene (fifth to sixth century BCE), and a number of fragmentary languages. Some classification schemes put Italic in a special subunity with Celtic known as "Italo-Celtic."

### Germanic

The Germanic subgroup, which includes English among its members, is widespread geographically and is internally heterogeneous. The oldest attestations of Germanic are the Scandinavian Runic inscriptions, the oldest of which dates from the first century CE. The Germanic languages are conventionally separated into three geographic subdivisions. The first, East Germanic, contains only a single well-attested language, Gothic. Gothic is the language with the oldest continuous documents in Germanic, the biblical translation by Bishop Wulfila from around the second half of the fourth century CE. The second subdivision of Germanic is North Germanic, whose principal representative is Old Icelandic (also called Old Norse). Apart from the Runic inscriptions, the oldest material in North Germanic comprises Norwegian and Icelandic sagas and legal texts from the ninth century. Modern North Germanic languages are Icelandic, Faroese, and Norwegian in one group, and Danish and Swedish in another. The final group, West Germanic, is the most expansive and internally diverse of the Germanic languages; its descendants include German, Yiddish, Dutch, Flemish, Afrikaans, and English, with its many varieties worldwide. (See further ENGLISH AS A GERMANIC LANGUAGE.)

### Celtic

The languages of the Celtic subgroup are traditionally divided into two main geographical sections, the Continental and the Insular. The Continental group, made up of Celtiberian (Hispano-Celtic), Lepontic, and Gaulish, is extinct. The oldest material from the Continental group is from the sixth century BCE. The Insular Celtic languages show up materially somewhat later. Split into two groups, Goidelic and

Brittanic (Brythonic), the Insular languages are first attested in some Ogham Irish sepulchral inscriptions from around 300 CE. The Goidelic group is made up of Irish, Scots Gaelic, and the extinct Manx. Brittanic comprises Welsh, the most robust of the modern Celtic languages, Breton, spoken in Brittany (France), and the extinct Cornish. Some classification schemes put Celtic in a special subunity with Italic known as "Italo-Celtic."

## Tocharian

Discovered in archeological excavations around the turn of the twentieth century in Chinese Turkestan, the two varieties of Tocharian (usually called simply "A" and "B") have added modestly to the Indo-European base. The documents of the languages, mostly religious and some commercial, are relatively late, stemming from the period of about 500–700 CE.

## Baltic

Sometimes grouped with the Slavic languages to form a composite intermediate branch called "Balto-Slavic," the Baltic subgroup survives in two modern languages, Lithuanian and Latvian (Lettish), which together make up the East Baltic subdivision. Many other Baltic languages have become extinct, including a language called Old Prussian, which was spoken until the early eighteenth century and represents the West Baltic subdivision. The oldest Baltic material, the Old Prussian Basel Epigram, dates to as early as 1369 CE, while the oldest Lithuanian texts stem from the early sixteenth century, and the oldest Latvian material is probably datable to 1585.

## Slavic

Often grouped with Baltic as "Balto-Slavic," the Slavic languages fall into three geographical subdivisions. The first, South Slavic, comprises Bulgarian, Macedonian, Serbian, Croatian, Slovenian, and the extinct Old Church Slavic, in which the bulk of the oldest (tenth century) Slavic materials are written. The second Slavic subdivision is West Slavic, which comprises Czech, Slovak, Polish, Kashubian, and some others. And finally there is East Slavic, made up of Russian, Ukrainian, and Belarussian.

## Armenian

Armenian is first attested in religious documents and translations from the fifth century CE. It shows a great deal of influence from neighboring languages, including Greek, Arabic, Syriac, and Persian, so much so in fact that it was first misclassified as a dialect of Iranian.

### Albanian

Unknown linguistically until the fifteenth century CE, Albanian shows a great deal of influence from neighboring languages such as Greek, Slavic, and Turkish, as well as from Latin. This made its secure identification as a branch of Indo-European somewhat problematic when the IE languages were being classified in the eighteenth and nineteenth centuries. The first document in Albanian is a baptismal formula from the fifteenth century. There are two principal dialects, Gheg and Tosk.

### Fragmentary languages

In addition to the 11 major subgroups, there are also many apparently unaffiliated languages which survive only in fragments such as glosses and sporadic inscriptions. These languages provide enough information to be classified as IE, but not much beyond that. Included among the fragmentary IE languages are Ligurian (northern Italy, possibly related to Celtic), Messapic (southern Italy, possibly connected with Illyrian), Sicel and Sicanian (Sicily), Venetic (northeastern Italy), Thracian (in the area of modern Bulgaria and southern Romania), Phrygian (in the area of modern central Turkey), Illyrian, from the Dalmatian coast area of the Adriatic), and several others.

### Aspects of the structure of PIE

The extensive comparison of the daughter languages and their analysis according to the comparative method and other established methodologies has led to a protolanguage that has been reconstructed in considerable detail. In this section we will identify some of the more prominent features of reconstructed "classical" PIE, especially those relevant for the history of English, largely omitting revisions, including laryngeals, based on Anatolian evidence.

### Phonology

Table 12.4 shows the correspondences between selected consonant and vowel segments in several ancient IE languages and the oldest Germanic languages. Reconstructed PIE initiates the correspondences.

Table 12.5 provides a few illustrative lexical reconstructions. (See further PHONOLOGY: SEGMENTAL HISTORIES.)

### Morphology

#### Nominal and pronominal morphology

"Classical PIE," that is, the PIE reconstructed before the integration of Anatolian evidence into the protolanguage, is considered to be an inflectional (fusional) language

*Table 12.4*  Phonological correspondence among some PIE languages

| PIE | Hitt. | Skt. | Lat. | Gk. | Goth. | OIc. | OHG | OE |
|------|-------|--------|-------|----------|-------|------|-------|-------|
| p | p | p | p | p | f | f | f | f |
| t | t | t | t | t | þ | þ | d | þ |
| k | k | ś | k | k | h(j) | h | h | h |
| $k^w$ | ku | k/c | qu | p/t/k | hw/w | hv | hw/w | hw |
| b | p | b | b | b | p | p | p/pf | p |
| d | t | d | d | d | t | t | z | t |
| g | k | j | g | g | k | k | k | k |
| $g^w$ | ku | g/j | gu/u | b/d/g | qu | kv | q | cw/k |
| $b^h$ | p | bh | f(b) | ph | b | b | b | b |
| $d^h$ | t | dh | f(d) | th | d | d | t/d | d |
| $g^h$ | k | h | h | kh | g | g | g | g |
| $g^{wh}$ | ku | gh/h | f | ph/th/kh | w | w | w | w |
| s | s | s | s | s | s | s | s | s |
| m | m | m | m | m | m | m | m | m |
| n | n | n | n | n | n | n | n | n |
| l | l | l/r | l | l | l | l | l | l |
| r | r | r/l | r | r | r | r | r | r |
| w/u | w | v | v | Ø | w | v | w | w |
| y/i | y | y | j | h/z | j | Ø | j | g(y) |
| a | a | a | a | a | a | a | a | æ |
| e | e, a | a | e | e | i | e | e | e |
| i | i | i | i | i | i | i | i | i |
| o | a/ā | a | o | o | a | a | a | æ |
| u | u | u | u | u | u | u | u | u |
| ā | ā | ā | ā | ā/ē | ō | ō | ō | ō |
| ē | ē | ā | ē | ē | ē | ā | ā | ǣ |
| ō | ā | ā | ō | ō | ō | ō | ō | ō |

(Hitt. = Hittite; Skt. = Sanskrit; Lat. = Latin; Gk. = Greek; Goth. = Gothic; OIc. = Old Icelandic; OHG = Old High German; OE = Old English)

in which case markers on nominals (nouns, adjectives, pronouns) indicate their grammatical relationship to other words in a sentence, and mark gender and number agreement among words in phrases. The protolanguage is traditionally reconstructed with eight (occasionally nine) cases which indicate grammatical and semantic distinctions such as subjecthood, objecthood, direction towards, dislocation from, temporality, exchange, possession, agency, and instrumentation. The cases are known as the

*Table 12.5*  Some PIE reconstructions, based on "core" vocabulary from IE languages

|  | Hitt. | Skt. | Lat. | Gk. | Goth. | OIc. | OHG | OE | PIE |
|---|---|---|---|---|---|---|---|---|---|
| three | teri- | tráyaḥ | trēs | treîs | þreis | þrīr | drī | þrī | *trei- |
| seven | šipta- | saptá | septem | heptá | sibun | siau | sibun | seofon | *septm̥ |
| cow | wa-wa-(i)-ˣ | gáuḥ | bōs | boûs | ⊕ | kýr | chuo | cū | *gʷou- |
| I | ūk | ahám | ego | egṓ(n) | ik | ek | ih | ic | *eg- |
| foot | pata-ˣ | pā́t | pedisº | podósº | fōtus | fōtr | fuoz | fōt | *ped- |
| heart | kard- | ⊕ | cordisº | kardíā | haírtō | hjarta | herza | heorte | *kerd- |
| sheep | ḫawi-ˣ | áviḥ | ovis | o(w)is | ⊕ | ǣr | ouwi | ēowu | *owi- |

ˣ The form is Hieroglyphic Luwian.
º The genitive case reveals the stem.
⊕ The cognate form is not found in this language.

*Table 12.6*  A sample noun declension (Lat. *servus* "servant")

| Case | Singular | Plural |
|---|---|---|
| Nominative | *servus* | *servī* |
| Genitive | *servī* | *servōrum* |
| Dative | *servō* | *servīs* |
| Accusative | *servum* | *servōs* |
| Ablative | *servō* | *servīs* |
| Vocative | *serve* | *servī* |

nominative, genitive, dative, accusative, ablative, locative, instrumental, and vocative. Number refers to the quantification of entities in a phrase; the protolanguage had three numbers (singular, dual, and plural), as well as three genders (gender is an unfortunate term which simply means a kind of noun class), namely masculine, feminine, and neuter. Adjectives followed the same pattern of inflection as nouns, and agreed in gender and number with their head noun. Pronouns are marked by their own more-or-less unique endings. Latin provides a useful analog to the PIE system, though without the locative and instrumental (table 12.6).

In the Latin sentence *Marcus serv**um** vocat* "Marcus calls the servant," Marcus' role as subject is marked by the ending *-**us*** and the servant's role as object is indicated by *-**um***. The order of the words is grammatically irrelevant (Latin, like PIE, usually puts the verb at the end). When words occur as members of a constituent (word group), their membership is indicated by shared endings marking case, number, and gender, as in *velōcī equō* "to the swift horse" [dative-singular masculine], *malōrum animālium*

"of the bad animals" [genitive-plural-neuter], or *ab aliīs fēminīs* "from the other women" [ablative-plural-feminine]. (See further HISTORY OF ENGLISH MORPHOLOGY.)

### Verbal morphology

PIE verbs are synthetically complex amalgamations of meaningful elements which indicate grammatically and semantically significant categories. The PIE verb encoded two voices, active and mediopassive (voice reflects the role of the subject); a number of tenses (tense locates the verbal action temporally: at least the present, imperfect, aorist, perfect, and possibly a future are usually postulated); and mood, which indicates the factual content of the utterance from the speaker's point of view: at least the indicative, imperative, and optative moods are reconstructed, and occasionally the subjunctive. Voice, tense, and mood markers are attached to stems indicating aspectual categories (e.g., whether the action is continuous or punctual), and the entire complex is indexed to the subject by means of person/number markers. Verbs can be transitive (i.e., they can govern an object as in "Mary sees Bill") or intransitive ("Sarah walks to school"), though there is no specific formal marking on the verb to distinguish the transitive and intransitive types.

Once again Latin can be instructive, though it is not a perfect replica of PIE: a verb form like *am-ā̆ba-t* in *rēx amābat* "the king used to love" contains a stem form (*am-*), which indicates the lexical meaning "love"; a mood marker (*-ā̆-*), marking indicative (factual) mood; a tense/aspect marker (*-ba-*) which marks continuous past action; and finally a person/number/voice marker (*-t*), which indicates third person singular in the active voice. If we change the example to the passive *rēx amābātur* (*am-ā̆-bā̆-t-ur*) "the king used to be loved," the marker of passivity is the final *-ur*; in the plural *rēgēs amābantur* (*am-ā̆-ba-nt-ur*) "the kings used to be loved" note that the person/number marker is now *-nt-*.

### Syntax

Fusional languages like PIE and many of its descendants (including Old English, though not to the same extent as Latin, Greek, or Sanskrit) have fundamentally different syntactic patterns from languages like Modern English or French. The reason has much to do with word order, and the fact that a good deal of the syntax of fusional languages is conveyed in morphological expressions, such as case endings. In Modern English, for example, the order of elements in a sentence is grammatically fixed: except in stylistically marked utterances such as "Bagels, I like," the subject precedes the verb, and the object follows the verb in simple sentences. It is not grammatical to say "Him John sees" or "Sees him John" to mean "John sees him." But in fusional languages like PIE, word order is a stylistic, not a grammatical device. Latin is illustrative again: *Marcus mē vocat* "Marcus calls me" represents the preferred (unmarked) order of elements, but *mē vocat Marcus* or *Marcus vocat mē* have the same semantic value as

*Marcus mē vocat*. That is because the grammatical indication of subject (*Marcus*) and object (*mē*) is being carried by the endings, not the position of the words relative to each other; furthermore, the verb *vocat* is indexed by the final *-t* to the third person nominal subject *Marcus*, and couldn't possibly go with *mē*. PIE (like Old English) was dominantly verb-final (*John him sees*). Verb-final languages have certain properties such as: they use postpositions (*the world over*); adjectives typically precede the noun they modify (*the proud winners*), also true for genitives (*Susie's exam*); comparative constructions have the order standard-marker-adjective (*Louis than taller* [= *taller than Louis*]); and relative clauses precede the noun they qualify (*who teach English professors* [= *professors who teach English*]). (See further HISTORY OF ENGLISH SYNTAX.)

The ways in which many of these features of PIE descended into Germanic and on to English are discussed in ENGLISH AS A GERMANIC LANGUAGE, in this volume.

## ACKNOWLEDGMENT

## REFERENCES AND FURTHER READING

Fortson, B. W., IV (2004). *Indo-European Language and Culture: An Introduction*. Oxford: Blackwell.

Mallory, J. P. & Adams, D. Q. (1997). *Encyclopedia of Indo-European Culture*. London: Fitzroy Dearborn.

Onions, C. T. (1969). *The Oxford Dictionary of English Etymology*. Oxford: Clarendon Press.

Pokorny, J. (1951–9). *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.

Ramat, A. G. & Ramat, P. (eds.) (1998). *The Indo-European Languages*. London: Routledge.

Szemerényi, O. (1996). *Introduction to Indo-European Linguistics*. 4th edn. Oxford: Oxford University Press.

Watkins, C. (2000). *The American Heritage Dictionary of Indo-European Roots*. 2nd edn. Boston: Houghton Mifflin.