

Origins of the English Language

from "The Origins of the English Language: A Social and Linguistic History", by Joseph M. Williams.

The Indo-European Family of Languages

Our English vocabulary is not something to be studied in isolation but is related in one way or another to many of the other languages of the world. The proper beginning for us, therefore, is to view the place of English in perspective, amid the many tongues of mankind.

Those students who have studied German undoubtedly have noticed a remarkable similarity between that language and their own. The German word *Milch* is very close in sound to the English *milk*; likewise, the German *Wasser* and English *water*, *Brot* and *bread*, *Fleisch* and *flesh* closely resemble each other, not to mention a great many additional examples. Perhaps we can see this similarity best if we place side by side in systematic form the words for *mother*, *father*, and *brother*, as they appear in various tongues.

English	German	Dutch	Danish
mother	Mutter	moeder	moder
father	Vater	vader	fader
brother	Bruder	broeder	broder

While the spelling of these words makes their similarities obvious, we would be even more struck by the likeness if we heard them pronounced. For instance, a German "v" (as in *Vater*) sounds the same as an English "f"; also, a brief consideration of one's own speech will show that the sounds represented by "t," "th," and "d" are closely related.

Now at first glance we might conclude that the similarities among these languages are due to borrowing, that because these languages are spoken by people living relatively close to one another, such words were adopted from one of the languages by the others. This is not the case, however, as we shall see later. For one thing, there is usually no need for languages to borrow such simple, fundamental words.

If we extend the table to cover a wider area, we shall find the same similarity, though not to so great a degree.

Spanish	French	Latin	Greek
madre	mère	mater	meter
padre	père	pater	pater
*	frère	frater	*(phrater)

* There are some gaps and irregularities in this pattern; the Spanish for brother is *hermano*, related to English *germane*, while the Greek *phrater* means "member of a clan" (originally, "of a brotherhood").

We have said that the similarities between these languages was not to be explained in terms of borrowing. The cause of their resemblance lies rather in the fact that they are descendants of a single parent language. Thus, most of the major languages of Europe, and some of the languages of Asia, belong to one family known as the Indo-European family of languages. (There are exceptions like Finnish and Hungarian, which belong to the Ural-Altai family of languages.)

The original Indo-European parent language became extinct long before written records existed. It was spoken by a prehistoric people whose homeland was somewhere in Eastern Europe. The location of their origin, however, remains an unsolved problem. Sometime between 3000 and 2000 BC this primitive people, still living in the Stone Age, began a series of migrations. During the centuries that followed, successive waves of Indo-Europeans moved westward into Europe and southward into Persia and India, and the language of these invaders for all practical purposes obliterated the languages of the earlier inhabitants, in much the same way that, in the United States, English has superseded the languages of the American Indians.

Possibly the only survivor of the pre-Indo-European tongues of Europe is Basque, spoken in a mountainous corner of Spain, where the original inhabitants were able to maintain their way of life against the invaders.

But language is constantly changing, and as the various groups of Indo-Europeans became isolated from one another in the vast tracts of Europe and Asia, the language of each group began to evolve in its own peculiar fashion. Eventually a number of distinct dialects arose, and in the course of centuries there was no long one common tongue, but a series of completely different languages,

each of which in turn produced still more numerous descendants, and so on, until the present multitude of languages in Europe and Southwest Asia emerged. In diagrammatic form, the relationship of the Indo-European languages resembles a family tree, some branches having died out, others having given rise to many descendants. The modern descendants of the common Indo-European parent may be divided into eight principal groups or branches. Four of these belong to the Western or *Centum* subfamily, and the other four belong to the Eastern or *Satem* subfamily. (This classification is made on the basis of the word for "hundred" in the various languages, whether it more nearly approximates the Latin *centum* or the Zend *satem*.)

Teutonic

We should perhaps consider the Teutonic branch first, for to it belongs English. The primitive Teutonic, which antedates the earliest written records, eventually came to be divided geographically into three groups of languages, East, North, and West. The East Teutonic languages did not survive into modern times, but we know their principal representative, Gothic, from early translations of the New Testament into that language. The language of the Vandals, the barbarian raiders who sacked Rome in 455 AD, also belonged to this group. The North Teutonic languages are spoken today in the Scandinavian countries, Denmark, Norway, Sweden, and Iceland. West Teutonic is represented principally by modern German, Dutch, Flemish, and English.

Italic

When Rome was only a small village of rude huts on the banks of the Tiber, there were several Italic languages having equal status with Latin; but, as Rome achieved a dominant position in the ancient world, these disappeared, and Latin alone remained. The modern descendants of Latin, usually called Romance languages, show by their geographical distribution something of the extent of the Roman Empire. In France and Spain the Roman conquest resulted in the complete displacement of the earlier languages by Latin. So today French, Spanish, and Portuguese, as well as Italian, are classed as Romance languages. Far to the east, Romanian is likewise a descendant of the language of Roman colonists and soldiers.

Hellenic

The Hellenic branch of the Indo-European family is today represented by modern Greek, which is the descendant of the classical Greek of Plato and Aristotle and the common Greek dialect of the eastern Mediterranean area in which the New Testament was written.

Celtic

More than two thousand years ago the Celtic languages were spoken throughout a wide area of western Europe, generally comprising Gaul (France), part of Spain, and Great Britain. But as we have seen, after the Roman conquest, Latin replaced the Celtic languages on the Continent. And, as we shall see later, Teutonic Anglo-Saxon largely replaced the Celtic speech in Britain. In sections of the British Isles which were difficult for the invaders to reach, however, Celtic languages still exist, notably in Ireland and Wales.

Balto-Slavic

The Slavic languages are spoken in Russia, on the Balkan Peninsula in Bulgaria and Yugoslavia, and in Poland and Czechoslovakia.

The Baltic groups is today of far less importance. It comprises mainly Lithuanian and Latvian (or Lettish), the languages of two of the Baltic States forcibly absorbed by the Soviet Union in 1940, subsequently becoming independent nations with the breakup of the USSR.

Indo-Iranian

The oldest literary works in any Indo-European language are written in an Indian language, Sanskrit. As early as 1500 BC a number of very beautiful sacred books were composed in this language. Because of its antiquity and therefore closer resemblance to its Indo-European parent, Sanskrit is of great interest to linguists; only when European scholars became familiar with it did they realize fully the common origin of the many languages which we have been discussing. Sanskrit, however, is no longer spoken, but like Latin, it has become a learned language.

From early Indian dialects called Prakrits existing at the same time as Sanskrit ultimately came many of the languages of present day India and Pakistan, such as Hindi, Bengali, and Nepali.

Romany, the language of the Gypsies, likewise belongs to this group, although the Gypsies have wandered widely throughout Europe and even America; but their homeland was long ago in the northwestern part of India.

Closely related to the Indic group is the Iranian. Here is to be found the language of Iran (Persia). Modern Persian (Farsi) contains a great many Arabic elements, but it is nevertheless considered Indo-European. Between India and Iran is the small nation of Afghanistan, where likewise an Iranian language is spoken.

Armenian

Modern Armenian, the sole member of this group, is a language native to a small area east of the Black Sea, partly in Turkey and partly in Russia.

Albanian

The only surviving representative of the Albanian branch is Modern Albanian, spoken in a small nation just north of Greece.

Because all of these languages have come from a common ancestor, they are called cognate languages, and the similarities between them, which are not confined merely to vocabulary but include elements of grammar as well, might be compared to the physical similarities which between brothers and sisters, or between cousins.

We must remember, however, that the Indo-European is only one of a number of language families throughout the world. Universal agreement has not been reached as to the exact extent of many of these families. The following is a list of some of the more important language families, together with representative languages.

Semitic: Hebrew, Arabic, Syrian, Ethiopic, Phoenician (extinct)

Hamitic: the ancient Egyptian of the Pharaohs, Berber (or Libyan)

Indo-Chinese: Chinese, Burmese, Siamese, Tibetan

Ural-Altaic: Mongolian, Finnish, Estonian, Turkish, Magyar (Hungarian)

Dravidian: the languages of southern India and Ceylon.

Malay-Polynesian: the languages of the Malay Peninsula, Indonesia, the Philippines, and many of the Pacific islands, including Hawaii.

African Negro

Japanese and Korean

American Indian

Between the various families no definite natural relationship has been established. Thus, while English and Greek can be traced back to a common ancestor, and are cognate languages, this is not true of, say, English and Hebrew, or Chinese and Turkish.

The Background of English Vocabulary

Descent from a common ancestor is not the only type of relationship between languages. A relationship which is more obvious and which has greatly influenced the vocabulary of our languages is that which has come about through borrowing; and to see the effect of this, we should know something of the history of the English language.

During the Stone and early Bronze Ages, the British Isles were inhabited by a non-Indo-European race, about which next to nothing is known. Around 1000 BC, however, the Celts began to arrive in Britain and to conquer the prehistoric inhabitants, eventually completely wiping out their language. Thus, for some centuries, the languages of the British Isles were of the Celtic group, and a few of their descendants, Welsh and Irish, for instance, are spoken today.

In the first century AD, the Romans began the conquest of Britain, and for the next several hundred years were in control of most of it. Romanization was proceeding on the island in much the same fashion as on the Continent, when the Roman occupation was cut short by the withdrawal of the legions, which were sent to buttress the tottering Roman Empire against the onslaughts of the eastern barbarians. The language of the British Celts, therefore, was never completely replaced by Latin.

But something did happen to the Celtic languages of Britain, for it has already been pointed out that English is a Teutonic language. In the 5th century, after the Roman troops had been withdrawn, the Celts, having been exposed for some time to the softening effects of civilization, found themselves unable to resist the incursions of the fierce northern tribes, which had never been subdued by the Romans. The Celtic leaders sought the aid of

certain Germanic peoples living on the Continent to help them in their struggle. The Germans, however, found the island so much to their liking that they decided to remain, and in fact they conquered for themselves most of what is now England. These Teutonic tribes were known as the Angles, Saxons, and Jutes; their language is called Anglo-Saxon, and it became the basis of modern English. As a matter of fact, the English language took its name from that of one of the tribes.

From the very first, however, Anglo-Saxon or Old English was subjected to outside influences. Even before the Angles and Saxons came to Britain, they had come in contact with Roman civilization. For a time large parts of Germany were under Roman Domination, and from the Roman soldiers and the inevitable traders who traveled in their wake, the languages of the Teutonic tribes received a large number of Latin words. These words generally indicate the new products and concepts which were acquired from contact with a higher civilization. Thus, when they arrived in England, the Anglo-Saxons already had borrowed such words as *stræt* (from Latin *via strata*, "paved road"), which became *street* in modern English. Likewise, for example, came the words *ciese* (cheese), *win* (wine), *cuppe* (cup), and *pund* (pound).

Old English Period (450 to 1150)

One the Anglo-Saxons were in Britain, contact with the earlier inhabitants brought some Celtic words, mostly place names like *Kent* and *London* into Old English. Then, too, some words from Latin and Greek, such as *altar*, *candle*, and *priest*, were introduced during this period by Christian missionaries sent from Rome.

Beginning in the 8th century, England was subjected to repeated invasions by Danish marauders, many of whom settled in the areas of Britain which they had conquered. Despite the valiant efforts of English leaders, like King Alfred the Great, the invaders could not be dislodged, but continued to live side by side with the Anglo-Saxons. The Danes were eventually assimilated, and their language has likewise left its mark upon English.

Middle English Period (1150 to 1500)

The greatest event in the shaping of the English language, however, was the Norman Conquest. In 1066, William of Normandy defeated King Harold at the Battle of Hastings and became ruler of England. The Anglo-Saxon nobles were replaced by a French-speaking aristocracy who regarded themselves as primarily Frenchmen, and who looked upon England as an occupied territory. Norman-French became the official language of the country, the language of law courts, the schools, and the army.

In such a situation, the speech of the conquerors was bound to exert an influence upon English. Yet English remained as the language of the masses and kept its basically Teutonic structure. For a time, Norman-French, spoken by the nobles, and English, spoken by the lower classes, existed side by side in the conquered land, without affecting each other as much as is often supposed.

Around 1200, the Norman rulers of England lost many of their holdings on the Continent, including Normandy, and they henceforth began to regard themselves as Englishmen rather than as Frenchmen. By 1500, English has reasserted itself, but it was an English far different from the language of the Angles and Saxons. In the preceding three centuries many French-speaking aristocrats, as they learned English, had automatically introduced many French words into their speech when they could not remember the new language, or when the English synonym had been forgotten through disuse. Many who spoke English as a native language had learned French, not at this time the Norman-French of their early conquerors, but Parisian French, which had become the cultural and commercial language of Europe.

Since French was more polished than their own language, which had declined somewhat in the centuries when it was used mainly by the uneducated classes, they naturally borrowed French words to make up deficiencies in their native tongue. Often both the French word and its English equivalent were kept, sometimes with varying shades of meaning, and this has tended to make English vocabulary rich and varied. Thus we have both "begin," which is native English, and "commence," which is of French origin, as well as

the following examples: "sin" and "crime", "wretched" and "miserable", "shun" and "avoid".

During the centuries when the distinction between Frenchman and Englishman was disappearing in England, the English vocabulary was enriched by the addition of thousands of French words. The language of this period is called Middle English, and it reached its fullest development in the writings of Geoffrey Chaucer. Notable also during the 14th century in England was the translation of the Bible into English by John Wycliffe.

Modern English Period (1500 to present)

By 1500, a new force had begun to exert itself on the English language, this time the result of an intellectual movement rather than of a military conquest. The movement was the Renaissance or Revival of Learning, which was marked by the rapid advance of the sciences, a renewal of interest in the Greek and Roman classics, the rise of nationalism, and by such events as the Protestant Reformation, the invention of the printing press, and the discovery of the New World.

The growth of national consciousness brought with it a desire on the part of Englishmen and Frenchmen and Italians and others to write books in their native tongues rather than in Latin, which had been the universal language of learned men throughout the Middle Ages. Many English writers, however, felt that their own language was still not sufficiently developed to meet the demands of the new learning; therefore, to remedy what they considered deficiencies in vocabulary, they borrowed wholesale from Latin, which most of them knew almost as well as English. Further, the revival of interest in the ancient classics brought with it a flood of new ideas, especially from Greece; and when a new concept is imported from another culture, there is a strong tendency to import the word that denotes the concept as well.

The additions to English from Greek and Latin entered English during or after the Renaissance. Not only were words borrowed in great numbers directly from the pages of Greek and Roman authors, but countless new words were also coined, by combining elements from the classical languages. In fact, the practice of using Greek and Latin as the basis for new scientific terms is one that still continues.

Thus we have seen, in outline, the major forces which have shaped English vocabulary. But any contact with foreign nations brings with it new additions; and English, perhaps because of its conglomerate background, has shown a greater tendency to borrow than most languages. The sources of its words are, therefore, from the whole world, often reflecting the great extent of English based commerce and colonization.

The modern European languages, especially those of France, with its leadership in cooking and fashion, and Italy, with its preeminence in the arts, have continued to supply us with a store of words. From the Near East, Persian (Farsi), for instance, has given to English such common terms as "check", "divan", "pajamas", and "tiger". Arabic has supplied "cotton", "admiral", "syrup", etc. In the Middle Ages, Arabian science was considerably in advance of the European; a number of our early scientific terms consequently came to use from the East. Many of these, "algebra", "alcohol", and "alkali", for example, can be recognized from the fact that they begin with "al-", the Arabic definite article.

From the languages of India, long ruled by the British, have come "punch", "bungalow", "loot", "thug", and "dungaree." The Far East has produced fewer loan words, yet the Chinese has given "tea", "typhoon", and "catsup", while Japan has supplied "tycoon" and "kimono". From Malayan have come "bamboo" and "bantam". The islands of the Pacific have given us "tattoo", "taboo", and "hula".

The borrowings by English in the New World remind us of the different cultures with which the colonists and pioneers came in contact. English has absorbed words from American Indian languages, such as "caucus", "raccoon", "hickory", and "skunk". Some words came from the French in America; "prairie" and "butte" testify to the extensive early French exploration. "Levee", "picayune", and "bayou" came from the French settlements in Louisiana. The Dutch colonists in New York added "boss", "cookie", "stoop", and "scow". Later, from the Spanish-speaking culture of the Southwest English acquired words like "ranch", "canyon", "stampede", and "mustang". In the 19th century, immigrants from various countries added some of

their native stock words, such as “pretzel”, “hamburger”, and “delicatessen” from the German.

Thus we have seen something of the composite nature of English, and the many sources of its words. No other important language possesses such a complexity and variety of vocabulary; and this richness, while it has increased the difficulty of learning English, has made it an extremely flexible instrument.
