LANGUAGES OF THE WORLD
The most widely studied language family in the world is the **Indo-European**.

Many of the most important languages of the world are Indo-European. These languages are official or co-official in many countries and are important in academic, technical and world organizations.

Examples: **English**, **Spanish**, **French**, **German**, **Russian**.

Indeed, more than half the world’s population speak one or more of these languages either as a mother tongue or as a business language.

Languages that are essential in multinational contexts or with large numbers of speakers.

Examples: **Portuguese**, **Hindi**, **German**, **Bengali**.

Some of the great classical languages of religion, culture and philosophy were Indo-European.

Examples: **Latin**, **Greek**, **Persian**, **Sanskrit**, **Pali**.
Languages that are scattered around the world as their speakers are part of diasporas.

Examples: Greek, Yiddish, Polish, Armenian, Romany, Italian, Punjabi, Gujarati.

The Indo-European languages tend to be inflected (i.e., verbs and nouns have different endings depending on their part in a sentence). Some languages (e.g., English) have lost many of the inflections during their evolution.

The Indo-European languages stretch from the Americas through Europe to North India.

The Indo-European Family is thought to have originated in the forests north of the Black Sea (in what is now Ukraine) during the Neolithic period (about 7000 BC). These people began to migrate between 3500 BC and 2500 BC, spreading west to Europe, south to the Mediterranean, north to Scandinavia, and east to India.

The Indo-European Family is divided into twelve branches, ten of which contain existing languages. I will describe each of these branches separately.
LANGUAGES OF THE WORLD
‘Indo-European Family’
The Celtic Branch

- This is now the smallest branch. The languages originated in Central Europe and once dominated Western Europe (around 400BC). The people migrated across to the British Isles over 2000 years ago. Later, when the Germanic speaking Anglo Saxons arrived, the Celtic speakers were pushed into Wales (Welsh), Ireland (Irish Gaelic) and Scotland (Scottish Gaelic).

- One group of Celts moved back to France. Their language became Breton spoken in the Brittany region of France. Breton is closer to Welsh than to French.

- Other Celtic languages have become extinct. These include Cornish (Cornwall in England - now being revived), Gaulish (France), Cumbrian (Cumbria), Manx (Isle of Man - another language being revived), Pictish (Scotland) and Galatian (spoken in Anatolia by the Galatians mentioned in the Christian New Testament).

- Welsh has the word order Verb-Subject-Object in a sentence. Irish has the third oldest literature in Europe (after Greek and Latin).
These languages originate from **Old Norse** and **Saxon**. Due to the influence of early Christian missionaries, the vast majority of the Celtic and Germanic languages use the Latin Alphabet.

They include **English**, the second most spoken language in the world, the most widespread, the language of technology, and the language with the largest vocabulary. A useful language to have as your mother tongue.

**Dutch** and **German** are the closest major languages related to English. An even closer relative is **Frisian**.

**Flemish** and **Afrikaans** are varieties of Dutch while **Yiddish** is a variety of German. **Yiddish** is written using the Hebrew script.

Three of the four (mainland) Scandinavian languages belong to this branch: (**Danish**, **Norwegian**, and **Swedish**). Swedish has tones, unusual in European languages. The fourth Scandinavian language, Finnish, belongs to a different family.

**Icelandic** is the least changed of the Germanic Languages - being close to Old Norse. Another old language is **Faroese**.
Gothic (Central Europe), Frankish (France), Lombardo (Danube region), Visigoth (Iberian Peninsula) and Vandal (North Africa) are extinct languages from this branch.

German has a system of four cases and three genders for its nouns. Case is the property where a noun takes a different ending depending on its role in a sentence. An example in English would be the forms: lady, lady’s, ladies and ladies’. The genders are masculine, feminine and neuter. German has three dialects spoken in northern Germany, southern Germany and Austria, and a very different form spoken in Switzerland.

English has lost gender and case. Only a few words form their plurals like German (ox, oxen and child, children). Most now add an s, having been influenced by Norman French.
The Latin Branch

- Also called the **Italic** or **Romance Languages**.
- These languages are all derived from **Latin**. Latin is one of the most important classical languages. Its alphabet (derived from the Greek alphabet) is used by many languages of the world. Latin was long used by the scientific establishment and the Catholic Church as their means of communication.
- **Italian** and **Portuguese** are the closest modern major languages to Latin. **Spanish** has been influenced by Arabic and Basque. **French** has moved farthest from Latin in pronunciation, only its spelling gives a clue to its origins. French has many Germanic and Celtic influences. **Romanian** has picked up Slavic influences because it is a Latin Language surrounded by a sea of Slavic speakers.
- Portuguese and Spanish have been separate for over 1000 years. The most widely spoken of these languages is Spanish. Apart from Spain, it is spoken in most of Latin America (apart from Portuguese speaking Brazil, and a few small countries like Belize and Guyana).
Romansh is a minority language in Switzerland. Ladino was the language spoken by Spain's Jewish population when they were expelled in 1492. Most of them now live in Turkey and Israel. Provincial and Catalan are closely related languages spoken in the south of France and the north-east of Spain, respectively.

Note that Basque (spoken in parts of Spain and France) is not an Indo-European language - in fact it is totally unrelated to any other language of the world.

Galician is a Portuguese dialect with Celtic influences spoken in the north west of Spain. Finally, Moldavian is a dialect of Romanian spoken in the Moldova. Under the Soviets the Moldavians had to use the Cyrillic alphabet. Now they have reverted back to the Latin alphabet.

Apart from Latin, other extinct languages include Dalmatian, Oscan, Faliscan, Sabine and Umbrian.

Latin had three genders and at least six cases for its nouns and a Subject-Object-Verb sentence structure. Most modern Romance languages have only two genders, no cases and a Subject-Verb-Object structure.
The Slavic Branch

- These languages are confined to Eastern Europe.
- In general, the Catholic peoples use the Latin alphabet while the Orthodox use the Cyrillic alphabet which is derived from the Greek. Indeed some of the languages are very similar differing only in the script used (Croatian and Serbian are virtually the same language).
- One of the oldest of these languages is Bulgarian. The most important is Russian. Others include Polish, Kashubian (spoken in parts of Poland), Sorbian (spoken in parts of eastern Germany), Czech, Slovak, Slovene, Macedonian, Bosnian, Ukrainian and Byelorussian.
- The Slavic languages are famed for their consonant clusters and large number of cases for nouns (up to seven). Many of the languages have three numbers for verbs: singular, dual and plural. Macedonian has three definite articles indicating distance; all are suffixes: VOL (ox), VOLOT (the ox), VOLOV (the ox here), VOLON (the ox there).
The Baltic Branch

- Three Baltic states but only two Baltic Languages (Estonian is related to Finnish).
- **Lithuanian** is one of the oldest of the Indo-European languages. Its study is important in determining the origins and evolution of the family. Lithuanian and **Latvian** both use the Latin script and have tones. Lithuanian has three numbers: singular, dual and plural.
- **Prussian** is an extinct language from this branch.
The Hellenic Branch

- The only extant language in this branch is **Modern Greek**.
- Greek is one of the oldest Indo-European languages. **Mycenaean** dates from 1300BC. The Ancient Greek of **Homer** was written from around 700BC. The major forms were **Doric** (Sparta), **Ionic** (Cos), **Aeolic** (Lesbos), and **Attic** (Athens). The latter is Classical Greek.
- The New Testament of the Christian Bible was written in a form of 1st Century AD Greek called **Koine**. This developed into the Greek of the **Byzantine Empire**. Modern Greek has developed from this.
Greek has three genders and four cases for nouns but no form of the verb infinitive. The language has its own script, derived from Phoenician with the addition of symbols for vowels. It is one of the oldest alphabets in the world and has led to the Latin and Cyrillic alphabets. The Greek Alphabet is still used in science and mathematics.

Until the 1970s Greek was a **Diglossic** language. This means that there were two forms: **Katharevousa** used in official documents and news broadcasts and **Demotic** used in common speech.

The Greek spoken in Cyprus includes many Turkish, Arabic and Venetian words and has a different pronunciation to the official Greek of Greece.
The Illyric Branch

- Another single language branch. Only **Albanian** (called **Shqip** by its speakers) belongs to this branch. It has been written in the Latin script since 1909; this replaced a number of writing systems including Greek and Arabic scripts. Albanian has many avoidance words. Instead of saying *wolf*, the phrase *may God close its mouth* is used. The definite article is shown by a suffix: **BUK** (*bread*) **BUKA** (*the bread*). Many noun plurals are irregular.

- There are two dialects that have been diverging for **1000 years**. They are mostly mutually intelligible. **Geg** is spoken in the north of Albania and Kosovo (Kosova). **Tosk** is spoken in southern Albania and north west Greece.

- The ancient **Illyric** and **Mesapian** languages, spoken in parts of Italy, are considered by some to be an extinct member of this branch.
The Anatolian Branch

- This branch includes the language of the Hittite civilisation which once ruled central Anatolia, fought the Ancient Egyptians and was mentioned in the Christian Bible's Old Testament. Other languages were Lydian (spoken by a people who ruled the south coast of Anatolia), Lycian (spoken by a Hellenic culture along the western coastal regions), Luwian (spoken in ancient Troy) and Palaic.

- All languages in this branch are extinct.
The Thracian Branch

- This branch is represented by a single modern language, Armenian. It has its own script.
- Armenian is spoken in Armenia and Nagorno-Karabakh (an enclave in Azerbaijan). The language is rich in consonants and has borrowed much of its vocabulary from Farsi (Iranian). Nouns have 7 cases and the past tense of verbs take an E prefix like Greek.
- Three extinct languages from this branch are Dacian (or Daco-Mysian - spoken in the ancient Balkan region of Dacia), Thracian and Phrygian (spoken in ancient Troy).
The Iranian Branch

- These languages are descended from **Ancient Persian**, the literary language of the Persian Empire and one of the great classical languages.

- The main language of this branch is **Farsi** (also called **Iranian**, **Dari** and **Persian**), the main language of Iran and much of Afghanistan. **Kurdish** is a close relation. Kurdish is spoken in Turkey, Syria, Iran and Iraq by the Kurds. It is the second largest of the Iranian languages after Farsi. In Turkey it was banned until recently.

- **Pashto** (also called **Pushtu** or **Pakhto**) is spoken in Afghanistan and parts of north west Pakistan. **Baluchi** is spoken in the desert regions between Iran, Afghanistan and Pakistan. These languages are written in the Nastaliq script, a derivative of Arabic writing. It is interesting that you cannot tell which family a language belongs to by the way it is written.

- **Ossetian** is found in the Caucasus mountains, north of Georgia. **Tadzhik** is a close relative of Farsi, written in Cyrillic and spoken in Tadzhikistan (of the former USSR) as well as northern Afghanistan.

- **Avestan** is the extinct language of the Zoroastrian religion. **Scythian** is an extinct language of a warrior people who once lived north of the Black Sea.
The Indic Branch

- This branch has the most languages. Most are found in North India. They are derived from Sanskrit (the classical language of Hinduism dating from 1000BC). This gave rise to Pali (the language of Buddhism), Ardhamagadhi (the language of Jainism) and the ancestors of the modern North Indian languages.

- Of the modern North Indian languages, Hindi and Urdu are very similar but differ in the script. The Hindi speakers are Hindus and use the Sanskrit writing system called Devanagari (writing of the Gods). Urdu is spoken by the Muslims so uses the Arabic Nastaliq script. These two languages are found in north and central India and Pakistan. Nepali is closely related to Hindi.
In India most of the states have their own language. These languages either use Devanagari script or a derivation (if the people are Hindus) or the Arabic Nastaliq script (if the people are Muslims).

- Bengali (West Bengal as well as Bangladesh), Bhili (Central India), Oriya (in Orissa), Marathi (in Maharashtra), Assamese (in Assam), Punjabi and Lahnda (from the Punjab), Maithili and Maghadi (from Bihar), Kashmiri (Kashmir - written mainly in Nastaliq), Sindhi (the Pakistan province of Sindh - also written in Nastaliq), Gujarati (Gujarat in western India), Konkani (in Goa, an ex Portuguese colony, uses the Latin script), Sinhalese (Sri Lanka - uses its own script derived from Pali), Maldivian (Maldives - with its own script based on Arabic).
And..
The most surprising language in this branch is Romany, the language of the Roma (also known as Gypsies - this is a derogatory term which should not be used). The Roma migrated to Europe from India.

Sanskrit had three genders as has Marathi; most modern Indic languages have two genders; Bengali has none.

The fascinating point about India is that the south Indian languages (like Tamil) are not Indo-European. In other words, Hindi is related to English, Greek and French but is totally unrelated to Tamil. North Indians visiting Madras (in the south) are as baffled by Tamil as a foreigner would be.
The Tokharian Branch

- **Turfanian** and **Kucheán** are recently identified extinct languages once spoken in north west China. Very little is known about this branch as only a few manuscripts dating from 600 AD are in existence. The languages disappeared around the 8th century AD. The closest relatives of these languages are from the Celtic, Anatolian and Latin branches.
<table>
<thead>
<tr>
<th>Celtic Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welsh : Irish Gaelic : Scottish Gaelic : Breton</td>
</tr>
<tr>
<td>Cornish : Gaulish : Cumbrian : Manx : Galatian</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Germanic Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td>English : Dutch : Flemish : Frisian : Afrikaans</td>
</tr>
<tr>
<td>German : Yiddish : Danish : Swedish : Norwegian</td>
</tr>
<tr>
<td>Faroes : Icelandic</td>
</tr>
<tr>
<td>Anglo Saxon : Old Norse : Frankish : Gothic</td>
</tr>
<tr>
<td>Lombardo : Visigoth : Vandal</td>
</tr>
<tr>
<td>Branch</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
</tbody>
</table>
| **Romance (Latin) Branch** | Italian : Sardinian : French : Provencal : Catalanian  
Spanish : Ladino : Galician : Portuguese : Romansh  
Romanian : Moldavian  
| **Slavic Branch**       | Russian : Belorussian : Ukrainian : Polish : Sorbian  
Czech : Slovak : Slovene : Croatian : Serbian  
Kashubian : Bulgarian : Macedonian : Bosnian  
Old Church Slavic |
| **Baltic Branch**       | Lithuanian : Latvian  
Prussian |
| **Hellenic Branch**     | Modern Greek  
Mycenaean : Koine : Byzantine Greek  
Classical Greek (Attic : Doric, Ionic, Aeolic) |
| **Illyric Branch**      | Albanian  
Illyric : Mesapian |
| **Anatolian Branch**    | Hittite : Lydian : Lycian: Luwian : Palaic |
| **Thracian Branch**     | Armenian  
Dacian : Thracian : Phrygian |
### Iranian Branch

- Farsi
- Pashto
- Baluchi
- Ossetian
- Tadzhik
- Persian
- Avestan
- Scythian

### Indic Branch

- Hindi
- Urdu
- Nepali
- Bengali
- Assamese
- Oriya
- Kashmiri
- Punjabi
- Sindhi
- Marathi
- Gujarati
- Bhili
- Lahnda
- Maithili
- Magahi
- Konkani
- Sinhalese
- Maldivian
- Romany
- Sanskrit
- Pali
- Ardhamagadhi

### Tokharian Branch

- Turfanian
- Kuchean

Extinct languages are in lighter type.
A List of The Eleven Languages in The World With The Most Native Speakers

- Chinese 1 billion
- Hindu-Urdu 365 million
- English 350 million
- Spanish 250 million
- Bengali 205 million
- Portuguese 175 million
- Russian 165 million
- Arabic 150 million
- Malay 150 million
- German 100 million
Some Statistics

- 45 out of the 50 most frequent words in English are function words; 5 are content words.
- 85% of the 1000 most frequent words in English are of old English origin.
- In words in the range 1,000 – 10,000, 40 – 60% are of French or Latin origin.
TYPES OF LANGUAGES
Linguistic typology studies and classifies languages according to different features. It makes it possible to classify languages into larger groups according to the way they use sounds, types of morphology they do or do not possess and how they order the elements in sentences. Some features can be found all over the world in unrelated languages. The mere fact that two languages share a feature does not imply that they are related.
One of the most common ways of classifying languages is by the most typical order of the subject (S), verb (V) and object (O) in sentences such as “The cat eats the mouse”:

- SVO (“The cat eats the mouse”),
- SOV (“The cat the mouse eats”),
- VSO (“Eats the cat the mouse”),
- OSV (“The mouse the cat eats”),
- OVS (“The mouse eats the cat”),
- VOS (“Eats the mouse the cat”).
English is an SVO language, because all the other orders are incorrect or change the basic meaning (“The mouse eats the cat” means something very different). In some other languages, such as Russian, all these sentences would be correct, though one order might be much more common (in this case, SVO). In some languages, the order can depend on different parameters. French is usually SVO, but SOV when the object is a pronoun.

The most common orders are SVO (found in English, Indonesian, Chinese, Spanish and thousands of other languages) and SOV (found in Japanese, Persian, Hindi and Turkish among others). VSO is less common (found in Standard Arabic and Irish), and the three other orders that put the object before the subject are found in less than 5% of the world languages. Anecdotally, Yoda speaks in OSV order (Strong with the Force you are, but when 900 years you reach, look as good you will not).
Some languages use tone to distinguish words. One of the best-known examples is Mandarin Chinese, which has four tones. For example, sī (with a high tone) means “thought”, sì (with a falling tone) means “four” and sǐ (with a low falling and rising tone) means “death”. Tone languages are common in East Asia (other Chinese languages, Burmese, Thai and Vietnamese among others) but are also found in Africa (Yoruba) and the Americas (Navajo).

Languages that do not have tones usually have stress (a syllable pronounced more strongly than the others). In some languages, the stress always falls on the same syllable (the first in Hungarian, the last in French), but in other languages stress is important and is used to distinguish words. English and Russian are such languages: “increase” is a noun while “increase” is a verb, and in Russian, “muka” means “flour” while “muka” means “torture”.
Languages can also be classified by the way they mark grammatical functions. In isolating languages such as Chinese, words usually consist of one morpheme (a linguistic unit that carries meaning), while in synthetic languages, they can consist of several morphemes. English is mildly synthetic (it has inflections such as the plural suffix in books), while languages such as Latin have many inflections. Some languages such as Inuktitut (spoken by the Inuit) are sometimes called polysynthetic because they use very many inflectional affixes, and one word in such languages can correspond to an entire sentence in other languages.
Synthetic languages can be divided into fusional and agglutinative languages. In agglutinating languages, each morpheme usually has a single function and words can consist of many morphemes. Turkish is a good example of an agglutinative language. The word evlerimde means “in my houses” and consists of the following morphemes: ev (house), -ler (plural), -im (my) and -de (in). In fusional languages, affixes can combine functions. In the Spanish word habló (“he spoke”), -ó simultaneously indicates both the past tense and a third person singular subject.

There are no clear-cut boundaries between these categories however and languages rarely fall perfectly into one category.
Why Do Languages Change?
Why Do Languages Change?

All living languages change.

- Sound shift is found in the history of all languages.
- Languages that evolve from a common source are genetically related. Genetically related languages were dialects of a same language.
  
  **Example:** English, German and Swedish were dialects of a postulated earlier form of Germanic called Proto-Germanic.

- Earlier forms of Romance Languages such as Spanish, French, Italian were dialects of Latin.
- Earlier forms of Proto-Germanic, Latin and other languages were dialects of Proto-Indo-European, a postulated ancestor.
All components of the grammar may change.

- Phonological, morphological, syntactic, lexical and semantic changes occur.
- Words, morphemes, phonemes and rules of all types may be added, lost or altered.
- The meanings of words and morphemes may broaden, narrow or shift.
- The lexicon may expand by borrowing, which results in loan words in the vocabulary. This is very common in language contact situations. It also grows through word coinage, blends, compounding, acronyms and other processes of word formation.
The study of linguistic change is called historical and comparative linguistics. Linguists use the comparative method to identify regular sound correspondences among the cognates of related languages and systematically reconstruct an earlier protolanguage. Internal reconstruction uses the same method applied to different stages of the same language. Where available, written texts are also used to inform linguists about language change.
Languages may also be classified according to certain characteristic such as a rich versus an impoverished morphology or according to whether their basic word order is subject–verb–object like English or subject–object–verb like Japanese.
No one knows all the causes of linguistic change. Some sound changes result from assimilation, a fundamentally physiological process of ease of articulation. Others, like the great vowel shift, are more difficult to explain. Some grammatical changes are analogic changes, generalizations that lead to more regularity, such as cows instead of kine and waked instead of woke.
1. Which one of branches originate from Old Norse and Saxon?

A) The Germanic Branch  
B) The Latin Branch  
C) The Indic Branch  
D) The Celtic Branch  
E) The Anatolian Branch
1. Which one of branches originate from Old Norse and Saxon?

A) The Germanic Branch
B) The Latin Branch
C) The Indic Branch
D) The Celtic Branch
E) The Anatolian Branch
2. Which one of language is most used language in the world with the most native speakers?

A) Turkish
B) Arabian
C) English
D) Chinese
E) Portuguese
2. Which one of language is most used language in the world with the most native speakers?

A) Turkish
B) Arabian
C) English
D) Chinese
E) Portuguese
3. 85% of the 1000 most frequent words in English are of ..........., fill the empty piece of sentence correctly.

A) English origin
B) Greek origin
C) French origin
D) Hindu origin
E) Persian origin
3. 85% of the 1000 most frequent words in English are of ..........., fill the empty piece of sentence correctly.

A) English origin
B) Greek origin
C) French origin
D) Hindu origin
E) Persian origin
4. Which one of branches addresses to the smallest area of languages?

A) The Celtic Branch
B) The Latin Branch
C) The Indic Branch
D) The Germanic Branch
E) The Anatolian Branch
4. Which one of branches addresses to the smallest area of languages?

A) The Celtic Branch
B) The Latin Branch
C) The Indic Branch
D) The Germanic Branch
E) The Anatolian Branch
5. ....... and ....... are the closest major languages related to English.

Complete the sentence with the correct selection.

A) Japanese and Chinese
B) Arabic and Persian
C) Dutch and German
D) Turkish and Greek
E) Spanish and Portuguese
5. ....... and ....... are the closest major languages related to English.

Complete the sentence with the correct selection.

A) Japanese and Chinese
B) Arabic and Persian
C) Dutch and German
D) Turkish and Greek
E) Spanish and Portuguese