

Lecture 1

LANGUAGE.
GENERAL NOTIONS.

PLAN

1. The nature of language
2. Standard English
3. Language origin
4. The Indo-European language family
5. Functions of language

Nature of Language

Language is a system for communicating.

Written languages use *symbols* (characters) to build words.

The language's vocabulary is the entire set of words.

The language's syntax and grammar are the ways in which the words can be meaningfully combined.

The language's semantics is the actual meaning of words and combinations of words.

Nature of Language

In computer science, human languages are known
as *natural languages*.

Nature of Language

Animal versus Human Communication

Animals communicate

- by sound
- by odor
- by body movement

Nature of Language

Human communication is dependent on both signals and symbols.

Symbols are sounds or gestures that have meaning for a group of people - it is the cultural tradition that gives it meaning

Standard English

There are four processes a variety goes through to become standardized

- 1) ***Selection*** - The variety must be chosen out of a group of competing varieties as the one to be developed into the standard form.
- 2) ***Codification*** - the standard variety is codified (written down) so that it is some way fixed as the standard.
- 3) ***Elaboration of Function*** - As the standard language is diffused socially and geographically (often through writing or education systems) it becomes necessary for it to be used in a wider variety of functions,
- 4) ***Acceptance*** - the variety must be accepted as the standard variety by an influential group of society.

Pidgin and Creole

A *pidgin* is a simplified, make shift language that develops to fulfill the communication needs of people who have no language in common but who need to occasionally interact for commercial and other reasons. Pidgins combine a limited amount of the vocabulary and grammar of the different languages.

At times, a pidgin language becomes the mother tongue of a population. When that happens, it is called *a Creole language*. As pidgins change into creoles over several generations, their vocabularies enlarge.

language origin

1. The divine source
2. Natural sounds
3. Bow-wow theory
4. Yo-heave-ho theory
5. The oral-gesture source

the Indo-European family

Many of the most important languages of the world are Indo-European.

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

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

Languages are scattered around the world as their speakers are part of diasporas.

the Indo-European family

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

the 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**).

the Indo-European family

The Germanic Branch

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.

the Indo-European family

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).

the Indo-European family

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 Indo-European family

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 Indo-European family

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 Indo-European family

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.

the Indo-European family

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 Indo-European family

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.

the Indo-European family

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 Indo-European family

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.

the Indo-European family

The Tokharian Branch

Turfanian and **Kuchean** 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.

Language functions

MICRO FUNCTIONS:

- **Physiological function** (releasing physical and nervous energy)
- **Phatic function** (for sociability)
- **Recording function**
- **Identifying function**
- **Reasoning function** (instrument of thought)
- **Communicating function**
- **Pleasure functions**

Language functions

MACRO FUNCTIONS:

- Ideational function
- Interpersonal function
- Poetic function
- Textual function