

**LINC-202 Historical Linguistics and South Asian Language
Families**

Dr. P. CHANDRAMOHAN

Associate Professor

CAS in Linguistics

Annamalai University

chandruling@gmail.com

**CENTRE OF ADVANCED STUDY IN LINGUISTICS
ANNAMALAI UNIVERSITY
ANNAMALAI NAGAR
TAMIL NADU
INDIA**

Syllabus

Objectives:

This paper introduces various approaches and classifications made in language analysis to the students. A study of historical linguistics leads one to understand the different types of linguistic changes taking place in the language cross time. It is also helpful to equip the students in reconstructing historical changes taken place in cognate language with the linguistic evidences and to enable the students to classify the language.

Unit – I: Introduction

Synchronic and diachronic approaches to language; Interrelationship between diachronic and synchronic data; Use of written records for historical studies; Language classification: Genealogical; language isolates; criteria for typological classification—agglutinative, inflectional, analytic, synthetic and polysynthetic; basic word order typology—SVO, SOV, etc.

Reading List

Arlotto, A. 1972
Lehmann, W.p. 1962

Unit- II: Linguistic Change and Reconstruction

Sound change; Well-known sound laws: Grimm's law, Varner's law, Grassman's law; Neogrammarian theory of gradualness and regularity of sound change; genesis and spread of sound change; phonetic and phonemic change; split and merger; conditioned vs unconditioned change; types of change—assimilation and dissimilation, coalescence, metathesis, vowel harmony, haplology, epenthesis; Loss of sound change, social motivation for change; lexical diffusion; analogy and its relationship to sound change; reconstructing the proto-stages of languages, internal reconstruction and comparative method—their scopes and limitations; relative chronology of different changes. Transformational-generative approach to sound change—rule addition, rule deletion, rule generalisation, rule ordering.

Reading List

Arlotto, A. 1972
Bynon, Theodora. 1977
Hock, H. H. 1986

Unit-III: Language Families of South Asia

Notion of language family, sub grouping within a family; family tree and wave models; criteria for identifying family relationships among languages; definition of the word 'cognate'; Indo-Aryan, Dravidian-, Austro-Asiatic, Tibeto-Burman; their geographical distribution, enumeration; characteristics.

Reading List

Arlotto, A. 1972
Bynon, Theodora. 1977
Hock, H. H. 1986

Unit-IV: Language Contact and Dialect Geography

Linguistic borrowing—lexical and structural; motivations—Prestige and need-filling (including culture-based); Classification of loan words—Loan translation, loan blend, calque, assimilated and unassimilated loans; Bilingualism as the source for borrowing; dialect, idiolect; isogloss; methods of preparing dialect atlas, focal area, transition area and relic area.

Reading List

Arlotto, A. 1972
Bynon, Theodora. 1977
Hock, H. H. 1986

Unit-V: Areal Features of South Asia

South Asia as a linguistic area—phonological—length contrast in vowels and consonants, retroflexion, open syllable structure; morphemic structure rules; morphological and syntactical—agglutination, ergativity, agreement; productive use of conjunctive participles; passives; causatives; echo words; phenomenon of reduplication; copulative compounds; compound verbs, relative clause construction; dative /genitive subject construction.

Reading List

Bynon, Theodora. 1977
Masica, C.P. 1976
Hock, H. H. 1986
Steever, S. 1997.

Suggested Readings:

Bloomfield, L. 1933. Language. New York: Henry, Holt Rinehart and Winston.

Benveniste, E. 1971. *Vocabulary of Indo-European Institutions* (Translated from French). Coral Gables: The University of Miami Press.

Brian D. Joseph, Richard D. Janda (eds.) 2003. *The Handbook of Historical Linguistics*. Oxford: Blackwell.

Bynon, 1977. *Historical Linguistics*. Cambridge: Cambridge University Press.
Campbell, Lyle. 2004. *Historical Linguistics: An Introduction*. Massachusetts: MIT Press.

Hitchcock, C. 1998 *The Common Cause Principle in Historical Linguistics* *Philosophy of Science*, Vol. 65, No. 3 (Sep., 1998), pp. 425-447.

Hock, H. 1988. *Principles of Historical Linguistics*. Mouton de Gruyter.

Hockett, C.F. 1958. *A Course in Modern Linguistics*. New York:

Macmillian and Co. Joseph, D. B. and R. D. Janda 2004. *The Handbook of Historical Linguistics*. Oxford: Blackwell.

Lehman, W.P. 1962. *Historical Linguistics*. New York: Holt Rinehart and Winston.

Introduction

Historical linguistics is the study of not only the history of languages as the name implies, but also the study of how languages change, and how languages are related to one another. The main job of historical linguists is to learn how languages are related. Generally, languages can be shown to be related by having a large number of words in common that were not borrowed. Languages often borrow words from each other, but these are usually not too difficult to tell apart from other words. When a related group of languages has been studied in enough detail, it is possible to know almost exactly how most words, sounds, and grammar rules have changed in the languages.

There are over 6000 distinct human languages in the world. One basic question is how did they all evolved? Historical linguistics is the branch of Linguistics that focuses on the inter connections between different languages in the world and or their historical development. Historical linguist investigates how languages develop gradually and change through time, how multiple person's child /children languages can arise from one past "parent" language, and how cultural contact between speakers of different languages can influence language development and evolution. We all are aware that English has changed over time; none of us would be able to understand old English as it was spoken many centuries ago. We also know of words such as "astronaut" that our great grandparents would not have known about. A basic assumption in historical linguistics is that languages are constantly changing. We need to think about them as one of the most dynamic areas of culture.

So the branch of linguistics which investigates the processes of language change, which attempts to identify all types of historical and prehistoric connections between languages and which tries to establish genetic relationships between them. Though there were a few earlier efforts of note, historical linguistics largely gained importance with the rise of comparative linguistics in late 18th century and it was the first branch of linguistics to be placed on a firm scholarly footing; indeed, it was almost the only kind of linguistics pursued before the rise of general linguistics towards the end of the 19th century. Classical historical linguistics was chiefly concerned with the study of internal history. But the subject was revolutionized in the 1960s by the introduction of sociolinguistic techniques, and today external history is seed as equally important.

History and development of historical linguistics

Modern historical linguistics dates from the late 18th century and grew out of the earlier discipline of Philology. It is the study of ancient texts and documents which goes back to ancient times. At first historical linguistics was comparative linguistics and mainly concerned with establishing language families and the reconstruction of prehistoric proto-languages using the comparative method and internal reconstruction. The focus was on the well-known Indo-European languages many of them which had long written histories. But since then, significant comparative linguistic work has been done on the Uralic languages, Austronesian languages and various families of Native American languages among many others. Comparative linguistics is a part of a more broadly conceived discipline of historical linguistics. For the Indo-European languages comparative study is now a highly specialized field and most research is being carried out on the subsequent development of these languages, particularly the development of the modern standard varieties.

Scholars have undertaken studies attempting to establish language families linking Indo-European, Uralic and other families into Nostratic [Nostratic is a macro family, or hypothetical large-scale language family, which includes many of the indigenous language families of Eurasia, although its exact composition and structure vary among proponents]. These attempts have not been accepted widely because the information necessary to establish relatedness becomes less available as the time depth is increased. The time depth of linguistic methods is limited because of chance word resemblances and varies between language groups, but a limit of around 10,000 years is often assumed. The dating of the various proto-languages is also difficult. Several methods are available for this but only approximate results can be obtained.

Initially all modern linguistics was historical in orientation - even the study of modern dialects involved looking at their origins. But Saussure drew a distinction between Synchronic and diachronic linguistics which is fundamental to the present day organization of the discipline. Primacy is accorded to synchronic linguistics and diachronic linguistics is defined as the study of successive synchronic stages. Saussure's clear separation is now seen to be crystallized. In practice, purely synchronic linguistics is not possible for any period before the invention of the gramophone. Written records always cover behind speech in reflecting linguistic developments and in any case are difficult to date accurately before the development of the modern title page. The work of sociolinguistics on linguistic variation has shown synchronic states are not uniform. The speech habits of older and younger speakers differ in ways which leads to language change. Synchronic variation is linguistic change in progress.

The biological origin of languages is in principle a concern of historical linguistics. But most linguists regard it as too remote to be reliably established by standard techniques of historical linguistics such as the comparative method. Less standard technique such as mass lexical comparison is used by some linguists to overcome the limitations of the comparative method. In this sense most linguists regard them as unreliable. The findings of historical linguistics are often used as basis for hypotheses about the groupings and movements of peoples particularly in the prehistoric period. In practice it is often unclear how to integrate the linguistic evidence with the genetic evidence. For instance, there are a large number of theories concerning the homeland and early movements of the Proto-Indo-Europeans, each with their own interpretation of the genetic record.

Historical linguistics (also called diachronic linguistics) is the study of language change. It has five main concerns:

- To describe and account for observed changes in particular languages;
- To reconstruct the pre-history of languages and determine their relatedness, grouping them into language families (Comparative linguistics);
- To develop general theories about how and why language changes;
- To describe the history of speech communities;
- To study the history of words, i.e. etymology.

Comparative linguistics

Comparative linguistics is a branch of historical linguistics that is concerned with comparing languages in order to establish their historical relatedness. Languages may be

related by convergence through borrowing or by genetic descent. Genetic relatedness implies a common origin or Proto-language, and comparative linguistics aims to construct language families to reconstruct proto-languages and specify the changes that have resulted in the documented languages.

Etymology

Etymology is the study of the history of words when they entered a language from which source and how their form and meaning have changed over time. In languages with a long-detailed history the etymology makes use of linguistics to study of how words change from culture to culture over time. However, etymologists also apply the methods of comparative linguistics to reconstruct information about languages that are too old for any direct information to be known. By analyzing the related languages with a technique is known as the comparative method. Linguists can make inferences about their shared parent language and its vocabulary. In this way the word roots have been found which can be traced all the way back to the origin.

Synchronic and Diachronic Approach

An unrecorded and hypothetical language which is ancestral to one or more attested languages and whose properties are deduced by some process of reconstruction, most often comparative reconstruction in characterizing the ancestor of language family. We describe the language in a particular period which is called descriptive linguistics. This is the basic for linguistic field. Since descriptive linguistic analysis a language as it exists at a given point of time, which is called synchronic study. Opposed to this the study of language at different points on the time dimension is called diachronic study.

Writing system

All human orthography is an attempt to set down visually the flow of spoken words. We all learn to speak our own languages well before we learn to write. It is possible to talk about writing systems on the basis of such graphitic factors as the size, style and configuration of the symbols, or the direction in which they are written. But this does not help us to understand what the graphemes are and how they are used. In principle, any of the systems to be described could be written in almost any set of graphitic conventions. Sometimes for example, several directions are used during the history of a language as in early Greek which at different periods was written right to left, left to right and even using alternate directions. A more useful approach to writing systems is to classify them into cases that show a clear relationship between the symbols and sounds of the language. The vast majority of present day systems are phonological and the non phonological systems are mainly found in the early history of writing which is where we begin. Throughout history a number of different ways of representing language in graphic media have been invented. These are called writing systems.



All spoken languages have phonemes of at least two different categories viz., vowels and consonants. That can be combined to form of syllables. As well as segments such as consonants and vowels in some languages use sound in other ways to convey meaning. Many languages use stress, pitch, duration and tone to distinguish meaning. In sign languages, the equivalent to phonemes are defined by the basic elements of gestures, such as hand shape,

orientation, location and motion, which correspond to manners of articulation in spoken language.

Writing systems represent language using visual symbols which may or may not correspond to the sounds of spoken language. The Latin alphabet was originally based on the representation of single sounds. So that words were constructed from letters that generally denote a single consonant or vowel in the structure of the word. In syllabic scripts each sign represents a whole syllable. In logographic scripts, each sign represents an entire word and will generally accept no relation to the sound of that word in spoken language. Because all languages have a very large number of words, no pure logographic scripts are known to exist. Written language represents the way spoken sounds and words follow one after another by arranging symbols according to a pattern that follows a certain direction. The direction used in a writing system is entirely arbitrary and established by convention. Some writing systems use the horizontal axis (left to right or right to left), while others such as traditional Chinese writing use the vertical dimension (from top to bottom). A few writing systems use opposite directions for alternating lines and others such as the ancient Maya script can be written in either direction and rely on graphic cues to show the reader the direction of reading.

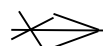

A writing system as a set of visible or tactile signs used to represent units of language in a systematic way. This simple explanation encompasses a large spectrum of writing systems with vastly different stylistic and structural characteristics spanning across the many regions of the globe. Writing provides a way of extending human memory by imprinting information into media less changeable than the human brain. However, many early philosophers like Plato, have branded writing as a loss to the human intellect. They argued that it makes the brain lazy and decreases the capacity of memory. It is true that many non-writing cultures often pass long poems and prose from generation to generation without any change and writing cultures can't seem to do that. But writing was a very useful invention for complex and high population cultures. Writing was used for record keeping to correctly counting agricultural products, for keeping the calendar to plant crops at the correct time and used for religious purpose like social functions. In past centuries, scientists had used writing as one of the "markers" of civilization. While it is true that writing systems appear to develop in agricultural and urban cultures by no means it is a requirement for civilization.

In Tamilnadu Writing was first developed sometime about 250 BC, when the Asokan Southern Brahmi script was adapted for Tamil. The earliest inscriptions in Tamil script proper are the *Pallava* copperplates of about AD 550. The Grantha script used in Tamil Nadu for Sanskrit since the 6th century was accommodated for Malayalam and Tulu. The Kannada–Telugu script is based on *Calukya* (6th century) inscriptions. Apart from these, Tamil has an old cursive script called *VaTTeluttu*, “round script,” and Malayalam possesses its own modern cursive form *Ko:leluttu*, “rod-script”. The ancient scripts ultimately capture is part or whole of a tongue spoken in ancient times. However, as you may have noticed all human languages evolve over time. As changes accumulate over time the ancient texts become unintelligible, if the knowledge of the language is lost. In some cases the texts can be read but cannot be understood. Ancient language had survived and evolved into later daughter languages that can be understood and then it becomes possible to know something the parent language as well as the processes involved in the development.

The power of a spoken over a written language is so noticeable as to be hardly valuable too mention. We all learn to speak our mother tongue well before we learn to write through formal education. Many people in the world who never learn to write their mother tongue whose languages have no written script; even though they all speak perfectly in their respective language. The earliest period in placing spoken words in visual form is the method of simply drawing pictures which look like the concrete objects that words refer to. In this system, the graphemes provide a recognizable picture of entities as they exist in the world. For instance a set of curly lines might represent the sea or a river and outlines of people and animals represent their living counterparts. There is no intention to draw the reality artistically or exactly, but the symbols must be sufficiently clear and simple to enable them to be immediately recognized and reproduced as occasion demands as part of a narrative. According to Arlotto in Sumerian the ancient language of Mesopotamia, the picture  represents a foot. The earliest stages of Chinese language the symbol  represented a horse. In both cases the symbol is a picture of a concrete object. It gives the information about what the word means, but tells nothing about how it might be pronounced. Ideographic writing is usually distinguished as a later development of pictographic. Ideograms or ideographs have an abstract or conventional meaning no longer displaying a pictorial link with external reality. Two factors account for this. The shape of an ideogram may so alter that it is no longer recognizable as a pictorial representation of an object; and its original meaning may extend to include notions that lack any clear pictorial form. In early Sumerian writing the picture of a starry sky came to mean, 'night', 'dark', or black'; a foot came to represent 'go', 'stand', and other such notions. It is rare to find a 'pure' ideographic writing system that is one in which the symbols refer directly to notions or things. Most systems that have been called ideographic in fact contain linguistic elements. The symbols stand for words in the language or parts of the symbols represent sounds. The Sumerian, Egyptian, Hittite and other scripts of the early period were all mixture of pictographic, ideographic and linguistic elements.

The **cuneiform method** of writing dates from the 4th millennium BC, and was used to express both non phonological and phonological writing systems in several languages. The name derives from the Latin meaning 'wedge – shaped' and refers to the technique used to make the symbols. A stylus was pressed into a tablet of soft clay to make a sequence of short straight strokes. In later periods harder materials were used. The strokes are thickest at the top and to the left reflecting the direction of writing. At first the symbols were written from top to bottom; later they were turned onto their sides and written from left to right. The earliest cuneiform was a development of pictographic symbols. Subsequently the script was used to write words and syllables and to mark phonetic elements. It was used for over 3000 years by such cultures as the Sumerians, Babylonians, Assyrians, and Hittites, finally dying out as the Christian era approached. The latest cuneiform tablets date from the 1st century BC, in which the script could not be read until the 19th century, when several of the languages it represented were finally worked out.

The extremely important factor in the move to phonetic representation was the need to write down proper names. In a system of syllabic writing (a syllabary), each grapheme corresponds to a spoken syllable usually a consonant – vowel pair. The phonetically written word we have the name of the ancient Sumerian city Girsu. It was an attempt to put this name in writing simply joint together the symbol with the meaning. For instance the symbol

 which represent 'gir' (knife) with the symbol  represent 'su' (meat) joint

together make a word 'girsu' which means city. It gives rise to the type of orthography known as syllabic. In contrast to an alphabet each symbol represents a single sound and each symbol in a syllable stands for a whole group of sounds. With alphabetic writing, there is a direct correspondence between graphemes and phonemes which makes it the most economic and adaptable of all the writing systems. Instead of several thousand logograms or several dozen syllables the system needs only a relatively small number of units which proves easy to adapt to a wide range of languages. Most alphabets contain 20-30 symbols, but the relatively complexity of the sound system leads to alphabets of varying size.

Classification of Languages

Typological classification

The classifications of languages are classified according to their structural characteristics and not according to their ancestry. Typological similarities are not considered to be indicative of genetic relatedness. This is based on a comparison of the formal similarities which share similar grammar exists between languages. It is an attempt to group languages into structural types on the basis of phonology, grammar or vocabulary rather than in terms of any real or assumed historical relationship.

August Wilhelm von Schlegel (1767 – 1845) and others in the early 19th century, recognized three main linguistic types on the basis of the way a language constructs its words. Schlegel suggested a three way classification of languages as monosyllabic, affixing and inflecting. Monosyllable languages which are simply composed of invariable disjoint meaning sounds. Flexional languages which for the purpose of expressing relation can regularly vary their roots as well as their affixes. His classification is based on the elements of meaning in a language with or without changes or modification in the root system.

August Schleicher 1821-1868 went back to **Schelegel's** three-way classification. He proceeded from the assumption that all language has meaning and relation and suggested that all languages could be classified by virtue of the manner in which sound is used to express these two aspects. The basic meaning elements are the root and the relational elements are suffixes, prefixes, infixes and regular variation. Applying this notion to the classes of languages the following typology results were obtained.

- (1) Monosyllable languages which are simply composed of invariable disjointed meaning sounds.
- (2) Agglutinative languages which can link to these invariable sounds and sounds of relation.
- (3) Flexional languages which for the purpose of expressing relation can regularly vary their roots as well as their affixes.

Isolative language

An isolating language is a type of language with a low morpheme per word ratio. In the extreme case of an isolating language each word contains a single morpheme. There is only one element of basic meaning per word then the language was isolative. All the words in that language are unchanging and there are no endings grammatical relationships are shown through the use of word order. Chinese, Vietnamese and Samoan are these languages come under isolative type.

Example:

Chinese:

Wǒ mǎi júzi chī

I buy orange eat

“I bought some oranges to eat”

t'a dau nar ch'u

he to where go

“Where is he going?”

Here each word is separate from the others and each contains only one element of meaning. A language can be said to be more isolating than another if its correspondence between word and number of morphemes approaches 1:1 more than the other one. To illustrate the relationship between words and morphemes, the English word "rice" is a single word consisting of one morpheme only (also "rice"). This word has a 1:1 morpheme-word ratio.

Agglutinative language

An agglutinative language is a type of synthetic language with morphology that primarily uses agglutination. Words are formed by joining phonetically unchangeable affix morphemes to the stem. In agglutinative languages each affix is a bound morpheme for one unit of meaning instead of morphological modifications with internal changes of the root of the word or changes in stress or tone. In an agglutinative language the stems do not change, affixes do not fuse with other affixes and affixes do not change form conditioned by other affixes. The term was introduced by **Wilhelm von Humboldt** to classify languages from a morphological point of view. It is derived from the Latin verb *agglutinare*, which means “to glue together”.

Words are built up out of a long sequence of units with expressing a particular grammatical meaning in a clear one to one way. A sequence of five affixes might express the meaning of a word, for instance – one for each category of person, number, tense, voice and mood. Turkish, Finnish, Japanese, Dravidian and Swahili form words are in this way. ‘He who gets water for me’ in Swahili is **anayenipatia maji**, which can be analyzed as:

a – na – ye – ni – pat – i – a (maji)

He pre. Who me gets for

(water)

Turkish:

sehir- ler-e gid - iyor – um

ciry – plural- to go- pre. I

“I am going to the cities”

Tamil:

avan paLLi-il paTi-ttu-kkonTiru-kkir-a:n

He school-loc. study-pre.cont. – III per.sg.

“He is studying in the school”

Analytic languages

Analytic languages show a low ratio of morphemes to words. In fact the correspondence is nearly one to one. A closely related concept is the analytic language which in the extreme case does not use any inflections to indicate grammatical relationships. But which may still form compound words or may change the meanings of individual words with derivational morphemes either of which processes gives more than one morpheme per word. Sentences in analytic languages are composed of independent root morphemes. Grammatical relations between words are expressed by separate words where they might otherwise be expressed by affixes which are present to a minimal degree in such languages. There is little to no morphological change in words. They tend to be uninflected. Grammatical categories are indicated by word order or by bringing in additional words (for example, a word for “some” or “many” instead of a plural inflection like English -s). Individual words carry a general meaning (root concept) and gradations are expressed by other words. Finally in analytic languages context and syntax are more important than morphology.

Synthetic languages

Synthetic languages form words by affixing a given number of dependent morphemes to a root morpheme. The morphemes may be distinguishable from the root or they may not. They may be fused with it or among themselves. In that multiple pieces of grammatical information may potentially be packed into one morpheme. Word order is less important for these languages than it is for analytic languages. Since individual words express the grammatical relations that would otherwise be indicated by syntax. In addition there tends to be a high degree of concordance. Therefore, morphology in synthetic languages is more important than syntax. Grammatical relationships are expressed by changing the internal structure of the words. Typically by the use of inflectional endings which express several grammatical meanings at once. Latin, Greek and Arabic are clear cases. For instance, the -o ending of Latin amo; ‘I love’ simultaneously expresses that the forms in the first person singular, present tense, active and indicative. If on the other hand, there were several meaningful elements but those were in some way fused together or were modified in different context, than the language was inflective or synthetic. Sanskrit, Greek, Latin etc are under this category. In this type of languages there are two types of fusions. They are

Regular fusion – change in the root

Regular fusion –no change in the root.

There were several meaningful elements, but these were in some way fused together or modified in different context. i.e., in this context the concepts are combined. There are two sub types of synthesis according to whether morphemes are clearly differentiable or not. These subtypes are agglutinative and fusion. Morphemes in fusion languages are not readily

distinguishable from the root or among themselves. Several grammatical bits of meaning may be fused in to one affix. Morphemes may also be expressed by internal phonological changes in the root such as consonant gradation and vowel gradation or by supra-segmental features such as stress or tone which are of course inseparable from the root. The Indo-European and Semitic languages are the most typically cited examples of fusion languages. On the other hand not all Indo-European languages are fusion. For instance Armenian and Persian are agglutinative, while English and Afrikaans lean more analytic.

Polysynthetic languages

In 1836, Wilhelm von Humboldt proposed polysynthetic under this category of language classification. These languages have a high morpheme to word ratio. A highly regular morphology and a tendency for verb forms to include morphemes that refers to several arguments besides the subject. Another feature of polysynthetic languages is commonly expressed as 'the ability to form words that are equivalent to whole sentences in other languages'. The distinction between synthetic languages and polysynthetic languages is therefore relative. The place of one language largely depends on its relation to other languages displaying similar characteristics on the same scale. Most of the world's polysynthetic languages are native to North America. These types of languages typically have long 'sentence words' such as the Yupik word *tuntussuqatarniksaitengqiggtuq* which means "He had not yet said again that he was going to hunt reindeer." The word consists of seven morphemes with the meanings. Except for the morpheme *tuntu* "reindeer", none of the other morphemes can appear in isolation.

tuntu - *ssur* - *Qatar* - *ni* - *ksaite* - *ngqiggte* - *uq*

reindeer - hunt - future - say - negation - again - III per.sing.

Another common feature of polysynthetic languages is a tendency to use head marking as a means of syntactic cohesion. This means that many polysynthetic languages mark grammatical relations between verbs and their constituents by indexing the constituents on the verb with agreement morphemes. Further the relation between noun phrases and their constituents by marking the head noun with agreement morphemes. There are some dependent marking languages that may be considered to be polysynthetic because they use case loading to achieve similar effects and very long words.

Edward Sapir (1884-1939) proposed typological classification based on the inter relation of three sets of distinctions. They are grammatical concepts, grammatical processes and firmness of affixation. Grammatical concepts are based on the relation of one word in the sentence to another. Four classes of such concepts are distinguished. (a) Basic or existing concepts, where no such relation is involved (play). (b) Derivational concepts, which give an added or altered meaning to the root without involving the rest of the sentence (players). (c) The change from player to players requires a change in the associated verb (kills to kill). (d) Pure relational concepts which are purely abstract and relate the concrete elements to each other (the orders kill the player vs. the players kill).

Sapir identifies isolating languages as those which always identify the word with the root; affixing languages as those which affix to the writing system.

Col - coRkkal ke:l -ke:TTa:n

Pal - paRkkal kal -kaRpa:n

Root modifying elements (affixes) and symbolic (inflective) languages as those which empty internal modifications of vowels or consonants. Affixing may be of two kinds viz., agglutinative and fusion. Fusion itself can be further divided in the regular fusion which involves no change in the root.

SVO / SOV

Winfred P. Lehmann first proposed to reduce the six possible permutations of word order to just two main ones, VO and OV, in what he calls the fundamental principle of placement. He argued that the subject is not a primary element of a sentence. VO languages are primarily right-branching or head initial i.e., heads are generally found at the beginning of their phrases. Opposed to this OV languages have a tendency to favor the use of prepositions instead of postpositions. There are six theoretically possible basic word orders for the transitive sentence: subject–verb–object (SVO), subject–object–verb (SOV), verb–subject–object (VSO), verb–object–subject (VOS), object–subject–verb (OSV) and object–verb–subject (OVS). SVO language is a language in which the verb typically comes before the object. The great majority of the world’s languages are either SVO or SOV with a much smaller but still significant portion using VSO word order. The remaining three arrangements are exceptionally rare with VOS being slightly more common than OSV and OVS being significantly rarer than the two preceding orders.

In linguistic typology, subject–verb–object (SVO) is a sentence structure where the subject comes first, the verb second and the object third. Languages may be classified according to the dominant sequence of these elements. It is the most common order by number of speakers and the second most common order by number of known languages after SOV. SVO and SOV together account for more than 75% of the world’s languages. The label is often used for ergative languages which do not have subjects, but have an agent–verb–object order. The Romance languages also follow SVO construction except for certain constructions in many of them in which a pronoun functions as the object.

Greenberg’s word - order types

Group – 1

VO languages

Verb precedes object – She loves him

Auxiliary precedes main verb – We do drive safely

Adjectives follows noun – This is a big house

Genitive follows noun - This is priya’s umbrella.

Relative clause follows head – The boy (who/whom) we met yesterday is very nice.

Prepositions - There is a calf in the field.

Case marking absent

Comparative adjective precedes standard

Group – 2

OV languages

Verb follows object

Auxiliary follows main verb

Adjectives precede noun

Genitive precedes noun

Relative clause precedes head

Postpositions

Case marking present

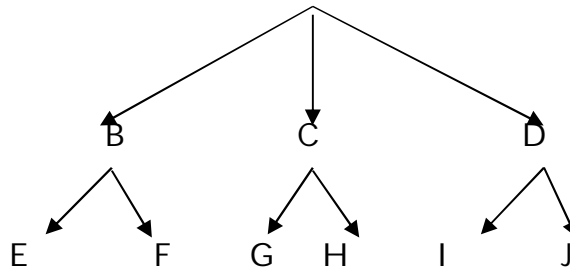
Comparative adjective follows standard

Genealogical classification

This classification is based not on the simple discovery of similar features in two or more languages and the correspondence is not sought in the philosophical or psychological constraints on human speech. The facts are based on the human history. The observed similarities and correspondences are accounted for by positing a definite contact of languages at some point in time past. The term ‘genetic’ strongly suggests a biological model or analog for the classification of languages and indeed linguists have employed biologically inspired models. A biological analogy employing the term ‘genetic’ should be applied to languages alone. Among cultural artifacts is virtually absent in the linguistics literature, where by long tradition going back over two hundred years most linguists have simply assumed the validity of a biological analog in linguistic classification. Within a biologically inspired framework there are at least two possible classes of interpretations of genetic relatedness. One could conceive of languages as unitary organisms and consider relatedness in a way analogous to that of individual animals or plants which can be related via lineages created through sexual or asexual reproduction. Alternatively, one could conceive of a language as a population either of speakers or of linguistic constructs or even of a population of speakers each with his/her idiolect. Population models of this sort might adopt a species analogy for understanding genetic relatedness.

Family tree model

The family tree model assumes that any set of related languages descends from a single ancestor according to the parthenogenesis model of biology. The family tree model was formed by **August Schleicher a 19th century's German linguist**. Within this model, families may be represented by means of diagrams like A represents the immediate ancestor of B, C, and D and the common ancestor of E, F, G, H, I, and J, all of which can be said to be related.



The classical family tree model assumes that the languages following this parthenogenetic analogy, that there can be no special genetic relationship between F and G other than their common descent from A. There is no linguistic feature or set of features which determine the genetic status of a language rather, it is the circumstance of its birth that determines this. In this way any word that is borrowed from another language does not affect its genetic status.

Sir William Jones

In **1786, Sir William Jones** an English jurist in India observed that Sanskrit the ancient classical and sacred language of India was systematically similar to Greek and Latin to his native languages and to earlier forms of English. This observation had been made before without widespread effect. After Jones's statement however scholars in Europe began systematic comparison of older forms of English and German with Latin, Greek, Sanskrit and other languages. Their work in the 19th centuries led to the classification of these languages in the Indo-European family and to the development of historical linguistics as a discipline.

The Romance languages which developed from Latin provide proof for such an explanation. Since French, Italian, Portuguese Spanish and other romance languages developed into independent languages. They furnish evidence in texts on how languages develop and how we can group languages by degrees of relationship. Some words listed below which shows similarities and prove it developed from attested Latin words with loosed some elements. Systematic similarities and difference in many words of the romance languages in Morphology and Syntax give us evidence assuming separate developments in the different areas in which Latin was once the common language.

The comparison of various languages led to the assumption that some languages are related that they are developed from a common source. This assumption came to be confirmed in language part through the linguistic situation in Western Europe. For common words in

French, Italian, Spanish, Portuguese and several other languages show consistent similarities and differences. For example compare the words “dear “and “field”

	French	Ital.	Span.	Port.
“Dear”	cher	caro	caro	caro
“Field”	champ	campo	campo	campo

The French (s) in these words spelled (Ch), corresponds consistently with the (K), spelled C; in the words of Italian, Spanish, and Portuguese. From these consistent correspondences we assume that at least some French (s) developed from earliest (K) through phonological change. This assumption is supported by other examples such as:

Fr.	Ital.	Span.	Port.	Cf. Latin	
chandelle	candela	candela	candeia	cande:la	“Candle”
chez	casa	casa	casa	casa	“House”

To assure ourselves that the change has truly lacteally indeed been in French we may look for French words in which the (K) has not changed. We find such words as in “School”

French	Italian	Spanish	Portuguese	Latin
e:cole	scuola	escuela	escola	shoal

Here the (K) was protected by the following (o). After comparing these and other words we propose rules for phonological equivalents in related languages. If we express the parallel elements, the rules may be written “=” (is read “corresponds to”).

$$\text{Fr.}(s) = \text{Ital.}(k) = \text{span.}(k) = \text{port.}(k)$$

We may also indicate relationships at different historical chronological periods and changes that have taken place. We then write rules such as:

$$\text{Lat.}(S) > \text{Ita.}(K) = \text{Span.}(k) \text{ Port.}(k) \text{ Fr}(S) < \text{Lat}(k)$$

(“>” is read became, developed to, “<” is read developed from”). We determine similar rules for morphological and syntactic patterns.

From correspondence like those presented here we also conclude that French, Italian, Spanish and Portuguese are members of one linguistic group of language family. The reason for their similarity is their common descent from one earlier language Latin through such comparison we can relate other languages. Assume that they developed from an earlier language and classify them as belonging to a specific language family. **Determination of language families is known as genealogical classification.**

Language Families

Language families are groups of languages said to be “genetically” related on the basis of having a common ancestor. Like living organisms languages develop and languages that come from the same ancestor called **“proto-language”** are part of the same language family. Such languages share several features and vocabulary items. Although these similarities are

hardly apparent when comparing two languages as it appears that distinct as say, English and Hindi both of the Indo-European family. As speakers move apart systematic changes occur over time until languages come to differ greatly from each other. Families are further subdivided into branches of languages that diverged from each other only after splitting from the family's common ancestor. These languages share more similarities with each other than with languages belonging to other branches within the family. Comparative linguistics compares languages in order to establish their historical relatedness. This can be done by comparing their phonology, grammar and vocabulary, even in cases where there are no written accounts of their ancestors.

Historical linguists normally use the comparative method to reconstruct a protolanguage. By examining several related languages for cognate's words bearing a similarity due to their common descent. Linguists can suggest the original forms from which the cognates arose. This method uses lexical terms such as pronouns, kinship terms, body parts and lower numbers which are terms most opposed to change. These methods have provided significant insight into the genetic links between different languages. It is important to remember that similarities can also take place from borrowing, linguistic universals or by chance. A language family is a group of languages interlinked by family background from a common ancestor and identified as the proto-language of that family. Similar to natural families the confirmation of relationship is visible through shared characteristics. In the same manner the languages of India grouped by major language families.

Families of Indian Languages

Indian languages have evolved from different stocks and are closely associated with the different ethnic groups of India. Broadly the Indian languages can be put into five groups. They are

- *Indo – Aryan*
- *Dravidian*
- *Tibeto – Burman*
- *Austro – Asiatic*
- *Andamanese*

Indo – Aryan

This is the most important family of Indian languages and comprises of all the principal languages of northern and western India such as Sanskrit, Hindi, Bengali, Marathi, Gujarathi, Punjabi, Sindhi, Rajasthani, Assamese, Oriya, Pahari, Bihari, Kashmiri and Urdu. It is part of the Indo- European family of languages which came to India with the Aryan. It is the biggest of language groups in India and accounts for about 74% of the total Indian population. The Indo-Aryan or Indic languages are the dominant language family of the Indian subcontinent spoken largely by Indo-Aryan people. They constitute a branch of the Indo-Iranian languages also a branch of the Indo-European language family. Indo-Aryan speakers form about one half of all Indo-European speakers. The largest in terms of native speakers are Hindustani (Hindi-Urdu about 240 million), Bengali (about 230 million), Punjabi (about 110 million), Marathi (about 70 million), Gujarati (about 45 million), Bhojpuri (about 40 million), Oriya (about 30

million), Sindhi (about 20 million), Sinhala (about 16 million), Nepali (about 14 million), and Assamese (about 13 million) with a total number of native speakers of more than 900 million.

Proto-Indo-Aryan or sometimes Proto-Indic is the reconstructed proto-language of the Indo-Aryan languages. Proto-Indo-Aryan is meant to be the ancestor of Old Indo-Aryan (1500–300 BCE) which is directly attested as Vedic and Mitanni-Aryan. i.e. early Old Indo-Aryan which include Vedic Sanskrit (1500-500 BCE) and late Old Indo-Aryan which include the Epic Sanskrit and Classical Sanskrit (500-300 BCE). In about the 4th century BC the Vedic Sanskrit language was codified and standardized by the grammarian Panini called “Classical Sanskrit”. Vedic has been used in the ancient preserved religious hymns, the foundational rule of Hinduism known as the Vedas. Mitanni-Aryan is of similar age to the language of the Rig-Veda, but the only evidence of it is a few proper names and specialized loanwords. The language of the Vedas – commonly referred to as “Vedic Sanskrit” by modern scholars is only slightly different from Proto Indo-Aryan the proto language of the Indo-Aryan languages.

Middle Indo-Aryan (Prakrits)

Outside the learned subject of Sanskrit, vernacular dialects (Prakrits) continued to develop. The oldest attested prakrits are the Buddhist and Jain canonical languages pali and Ardha Magadhi respectively. By medieval times, the Prakrits had diversified into various Middle Indo-Aryan dialects. “Apabhramsa” is the conventional cover term for transitional dialects connecting late Middle Indo-Aryan or Prakrits (300 BCE-1500 CE) with early Modern Indo-Aryan, spanning roughly the 6th to 13th centuries. The next major milestone occurred with the Muslim conquests on the Indian subcontinent in the 13th–16th centuries. Under the flourishing Mughal Empire, Persian became very influential as the language of prestige of the Islamic courts due to adaptation of the foreign language by the emperors. However, Persian was soon displaced by Hindustani. This Indo-Aryan language is a combination with Persian, Arabic and Turkic elements in its vocabulary with the grammar of the local dialects. The two largest languages that formed from Apabhramsa were Bengali and Hindustani.

New Indo-Aryan

The Indo-Aryan languages of Northern India and Pakistan form a dialect continuum. Hindi the Sanskritized version of the colloquial Hindustani in India is frequently Standard Hindi, spoken in the Delhi area since the Mughals. However the term Hindi is also used for most of the central Indic dialects from Bihar to Rajasthan. The Indo-Aryan prakrits also gave rise to languages like Gujarati, Assamese, Bengali, Oriya, Nepali, Marathi and Punjabi which are not considered to be Hindi despite being part of the same dialect continuum.

In the Hindi speaking areas the prestige dialect was long Braj Bhasha. But this was replaced in the 19th century by Khari Boli based Hindustani. This state of affairs continued until the Partition of India in 1947, when Hindi continued as an official language of India and Urdu in Pakistan. In contemporary times there is a continuum of Hindi–Urdu with heavily Persianised Urdu at one end and Sanskritised Hindi at the other, although the basic grammar

remains identical. Most people in India and Pakistan speak something in the middle and this is what the term Hindustani is frequently used to mean today.

Dravidian

The second largest language family comprises 70 languages spoken by more than 215 million people in South Asia. In terms of population figures the major languages of the family may be listed in the following order: Telugu, 52,986,000; Tamil, 44,400,000; Kannada, 27,900,000; Malayalam, 27,500,000; Gondi, 2,460,100; Tulu, 1,427,000; and Kurukh, 1,358,000. The Dravidian languages are spoken in the Republic of India (mainly in its southern, eastern, and central parts), in Sri Lanka, and by settlers in areas of Southeastern Asia, southern and eastern Africa. Commerce and emigration have also spread Dravidian languages, particularly Tamil, to Burma, Indonesia, Malaysia, Fiji, Madagascar, Mauritius, Guyana, Martinique and Trinidad. Brahui (Brahui), with 750,000 speakers in Pakistan, is isolated from all of the other members of the family. The four major languages viz., Tamil, Telugu, Kannada, and Malayalam possess independent scripts and literary histories dating from the pre-Christian Era. The four major literary languages are recognized by the constitution of India and also official languages of the states Andhra Pradesh, Tamil Nadu, Kerala and Karnataka. As an independent family the Dravidian languages were first recognized in 1816 by Francis W. Ellis. The actual term Dravidian was first employed by Robert A. Caldwell who introduced the Sanskrit word *dravida* (which, in a 7th-century text, obviously meant Tamil) into his book entitled 'A Comparative Grammar of the Dravidian or South Indian Family of Languages (1856).'

The Dravidian languages are divided into South, South-Central, Central, and North groups. These groups are further organized into 24 subgroups. According to P.S. Subramaniam Proto – Dravidian gave rise to twenty-one languages. They can be broadly classified into three groups, viz., South Dravidian, Central Dravidian and North Dravidian languages. The South Dravidian consists of seven languages. They are Tamil, Malayalam, Kodagu, Koda, Toda, Kannada and Tulu. The central Dravidian consists of eleven languages, viz., Telugu, Gondi, Gonda, Pengo, Manda, Kui, Kuvi, Golami, Naiki, Parji, and Gadaba. The North Dravidian consists of three languages, viz., Kurukh, Malto and Brahui. Brahui is spoken in Baluchistan, Malto spoken in West Bengal and Orissa, while Kurukh is spoken in West Bengal, Orissa, Bihar and Madhya Pradesh. Among the Dravidian language group, the Telugu is numerically the biggest of the Dravidian language. Then Tamil is the oldest and purest language of this family. Kannada and Malayalam are the youngest languages of this group.

Tibeto – Burman

The Tibeto – Burman language family has a considerable huge area in India and stretches all over the Sub – Himalayan tracts covering North Bihar, North Bengal, Assam up to the North eastern frontiers of the country. These languages are considered to be older than the Indo – Aryan languages and are referred to in the oldest Sanskrit literature as *Kiratas*. Most of these languages belong to the Tibeto – Buran sub family with the exception of Ahom which

belonged to the Siamese – Chinese branch. The Tibeto – Burman languages are divided into four broad groups' viz., Tibetan, Himalayan, North Assam and Assam – Burmese. The important Indian languages of Tibetan group include Sikkimese, Bhotia, Balti, Sherpa, Lahuli and Ladakhi which are all dialects of Tibetan. The important languages of Himalayan group are Kanauri and Limbu. The North – Assam group includes a number of languages like Abor (Adi), Miri, Aka, Dafla and ishmi. The Assam – Burmese group is numerically and culturally the most important of the Tibeto – Burman sub family. It is again sub divided into four main groups, viz., Kuki – Chin, Mikir, Bodo and Naga. This group also includes other languages like Singhpho of Assam and Mogh of Tripura which are offshoots of the languages spoken in Myanmar. Manipuri or Meithi is the most important language of the Kuki – Chin sub group. The Bodo sub – group includes such dialects as Bodo, Rajbangsi, Koch, Mech, Rabha, Dimasa, Kachari, Chutiya, Garo, Hajjong and Tripuri. Mikir has strong affinities to the Bodo and is spoken in the Mikir hills and parts of Sibsagar district in Assam. The principal languages of the Naga sub – group are Angmi, Sema, Ao, Lotha, Mao, Konyak, Kabui and Lepcha.

Austro – Asiatic

The Austric languages of India belong to the Austro – Asiatic sub family which are represented by languages of the Munda or Kol group spoken in the central, eastern and north – eastern India and languages of the Mon – Khmer group like Khasi and Nicobarese. These are very ancient languages which have been in existence much before the ancient of Aryans and were referred in Sanskrit literature as Nisadas. The most important language of the Austric group is Santhali which is spoken by over five million santhals and is the largest spoken languages among Adivasi. Mundari spoken by about a million speakers is another important language of this group.

Andamanese:

There are ten languages in the Great Andamanese family which can be grouped into three varieties viz. Southern, central and northern. These are Aka-Bea, Aka Bale the southern variety, Aka-pucikwar, Aka-kol, Aka-kede, Aka-jowoi as the central variety and Aka-jeru, Aka-Bo, Aka-Kora and Aka-Cari are a northern variety. Except for and Sare all great Andamanese languages are now extinct. Not all languages were mutually intelligible with each other as the languages of the great Andamanese tribe formed a dialect continuum. So that each language was closely related to its neighbour on each side but those at the extreme ends of the geographic continuum were mutually unintelligible. Hence Aka-Cari a north great Andamanese language was mutually unintelligible with Aka-Bea. The present day great Andamanese language is a mixture of four northern varieties with sporadic interferences from the central variety such as Aka-Pucikwar.

Onges are one of the most primitive tribes in India. They belong to the Negrito racial stock and they have been mainly seen near the Dugong creek in Little Andaman. They are dependent on the food provided by nature and are a semi-nomadic tribe. The onge population fell post British colonization from 672 in 1986 to 92 in 1901 but has remained stable since. At present the Onge populations have opened up to the locals in the island. The population of this tribe is stable and is at present 110. The Jarawa tribe is one of the largest tribes in Andaman Islands with a population more or less 400. For centuries this tribe has avoided all interaction with outsiders and therefore their name means “The hostile ones” or “people of the earth”. The Jarawa are still at the primitive stage of life on earth. However some ornaments made with

shells and palm leaves are worn by them but these are not in the sense to cover their nudity. This tribe has lived in the southeast part of Andaman but after the British regime they shifted to the western region of the island. The Sentinels people are said to be so hostile that their home has been named the 'hardest place to visit' in the world.

They inhabit the North Sentinel Island, and are the only remaining tribe in the Andamans to still maintain their isolation from the rest of the world. Nobody knows exactly how they look, the population or how they live. Since 1967, the Indian governments with the help of anthropologists have tried to make contact with the tribe. They tried giving gifts of food, coconuts, etc but they were always met with opposition. The tribe showers arrows and stones at whoever comes near the island.

Language Families of Europe

Indo-European

The hypothesis reappeared in 1786 when Sir William Jones first lectured on similarities between four of the oldest languages known in his time: Latin, Greek, Sanskrit and Persian. It was **Thomas Young who first used the term Indo-European in 1813**, which became the standard scientific term (except in Germany) through the work of Franz Bopp, whose systematic comparison of these and other old languages supported the theory. Bopp's Comparative Grammar appearing between 1833 and 1852 counts as the starting point of Indo-European studies as an academic discipline.

The Proto-Indo-European language (PIE) is the common ancestor of the Indo-European languages spoken by the Proto-Indo-Europeans. The classical phase of Indo-European comparative linguistics leads from **Franz Bopp's Comparative Grammar (1833) to August Schleicher's 1861 Compendium and up to Karl Brugmann's work published from the 1880s**. Brugmann's re-evaluation of the field and Ferdinand de Saussure's development of the laryngeal theory may be considered the beginning of "contemporary" Indo-European studies. The generation of Indo-Europeanists active in the last third of the 20th century developed a better understanding of morphology. Using the method of internal reconstruction an earlier stage called Pre-Proto-Indo-European has been proposed.

. The Indo-European comprises roughly 12 major groups and hundreds of languages. The major groups or subfamilies are Celtic, Italic (including Romance), Baltic, Slavic, Germanic, Anatolian, Greek, Indic, Iranian, Tocharian, Albanian, and Armenian. In addition, it appears that Baltic and Slavic should form a larger Balto-Slavic group and Indic and Iranian should be placed in an Indo-Iranian group. Brugmann reconstructed Indo – European by using data from the following eight branched or sub families.

- Indo – Iranian
- Armenian
- Greek
- Albanian
- Italic
- Celtic
- Germanic

➤ Balto – Slavic

Besides these branches the end of 19th century Meillet used data from two additional Indo –European sub family whose written records were discovered in the early 20th century.

- Anatolian and
- Tocharian.

Indo - Iranian

The Indo-Iranian languages also called Indo-Iranic languages and known in older literature as Aryan languages. It constitutes the easternmost extant branch of the Indo-European language family. It has more than 1 billion speakers extending from the Caucasus and Europe eastward to Xinjiang and Assam (Assamese) and south to Maldives and Fiji, forming the majority of all Indo-European speakers.

The common ancestor of all of the languages in this family is called Proto-Indo-Iranian also known as Common Aryan which was spoken in approximately the late 3rd millennium BC. The three branches of modern Indo-Iranian languages are Indo-Aryan, Iranian and Nuristani. Additionally sometimes a fourth independent branch, Dardic is posited, but recent scholarship in general places Dardic languages as archaic members of the Indo-Aryan branch. The Indo-Iranian languages derive from a reconstructed common proto-language called Proto-Indo-Iranian.

Armenian

Armenian is an isolate among Indo-European languages. Its sole genetic relations among known languages are at the reconstructed Proto-Indo-European (PIE) level. Some have attempted to connect Armenian more closely with other Indo-European (IE) language stock such as Greek, but the outstanding hypotheses are questionable. Historical and linguistic evidence suggests that the proto-language ancestral to Armenian was spoken in the Armenian Highland by 600 BC. A singularly opaque problem lies in determining the rate at which this early Armenian had changed from PIE, if we are to assign an absolute chronology to a reconstructed Proto-Armenian.

Greek

Greek in its ancient form has been spoken since the late 3rd century BC, although its oldest surviving text was written in 1400 BC. Despite this later date Greek is still the oldest recorded living language and longest documented Indo-European language spanning thirty-four centuries. Many important works of Western literature including the *Iliad*, the *Odyssey*, the philosophies of Plato and Aristotle and the New Testament were originally written in Greek. This ancient language may possibly be related to ancient Macedonian (spoken before AD) and Phrygian (extinct by the 5th century), but there is not enough documentation of either language to make a strong argument.

This language evolved to Mycenaean Greek as it is found on tablets written in Linear B and on some lines of the Homeric Epics. However, the Epics were mostly written in Classical

Greek and are considered the oldest texts in that language. One of the most important dialects of Classical Greece was the **Attic dialect**, mainly used in Athens and the language of philosophers and scientists. The Attic dialect was recognized as the official language of Greece by Philip of Macedon and the official language of the Hellenistic World by Alexander the Great.

Due to the adoption of the Attic dialect by many non Greek speaking populations many alterations were made and ultimately the Hellenistic **Koine language** emerged. This is mainly used in religious scripts such as the New Testament. The Koine language evolved into Medieval Greek which was mainly used in traditional folk songs and the final evolutionary step was the Modern Greek language which has been used since the last years of the Byzantine Era.

Albanian

The Albanian language is an Indo-European language in a branch by itself sharing its branch with no other existing language. Albanian is a descendant of ancient Paleo-Balkan languages, Illyrian, Messapic and Thracian. Historically two main dialectal groups exist. Gheg and Tosk both understand each other. Albanian is genetically connected with Illyrian and Messapic languages. There are also Thracian elements in it. In the Middle Ages Albanian was situated within the Balkan language agreement and generated significant characteristic features for all tongues of the peninsula. Most contacts took place with Bulgarian, Greek, Turkish, and Romanian. The oldest evidence of Albanian is from 1285 and its oldest text is from 1462.

Albanian is divided into three major dialects viz., Gheg, Tosk and a middle dialect zone between them. The Shkumbin River is roughly the dividing line with Gheg spoken north of the Shkumbin and Tosk south of it. There are also other dialects like Arberesh and Arvanitika exist in some zones of Italy and Greece. Standard Albanian based on the Tosk dialect of southern Albania is the official language of Albania and Kosova.

Italic

The Italic languages are a subfamily of the Indo-European language family originally spoken by Italic peoples. They include the Romance languages such as Italian, Spanish, Catalan, Portuguese, French, Romanian, Occitan, etc., which are derived from Latin. At present Latin and its daughter Romance languages are the only surviving languages of the Italic language family. Italic includes the Latin subgroup as well as the ancient Italic languages such as Faliscan, Osco-Umbrian and two unclassified Italic languages, Aequian and Vestinian. Venetic as exposed by its inscriptions was also closely related to the Italic languages and is sometimes classified as Italic.

Italic languages are one of the ten major subgroups of the Indo-European language family and might therefore have had an ancestor. The daughter languages of this family which are descend from common Italic or Proto-Italic. Among ten major groups linguist suggest a four-way division of East, West, North and South Indo-European.

Proto-Italic as a “chronological stage” without an independent development of its own, but extending over late PIE and the initial stages of Proto-Latin dates from 4000 BC to 1800

BC well before Mycenaean Greek. The Italic family has two known branches viz., Latino-Faliscan and Osco-Umbrian or Sabellian. Faliscan was spoken in the area around Falerii Veteres, north of the city of Rome and Sardinia. Latin was spoken in west central Italy.

Celtic

Celtic languages are descended from Proto-Celtic or Common Celtic a branch of the Indo-European family. The term “Celtic” was first used to describe this language group by Edward Lhuys in 1707. Celtic languages are most commonly spoken on the north-western edge of Europe notably in Ireland, Scotland and Wales and can be found spoken on Cape Breton Island. There are also a substantial number of Welsh speakers in the Patagonia area of Argentina. Some people speak Celtic languages in the other Celtic diaspora areas of the United States, Canada, Australia and New Zealand. In all these areas the Celtic languages are now only spoken by minorities though there are continuing efforts at revitalization.

Proto-Celtic divided into four sub-families viz., Gaulish, Galatian, Lepontic, and Noric. Lepontic, the oldest attested Celtic language from the 6th century BC, is treated as a primary branch possibly the first language to diverge from Proto-Celtic. These languages were once spoken from France to Turkey and from Belgium to northern Italy. They are now all extinct.

The Celtic languages have been rather challenging owing to lack of primary source data. Some scholars distinguish Continental Celtic and Insular Celtic. They argue that the differences between the Goidelic and Brittonic languages arose after these split off from the Continental Celtic languages. Some other scholars distinguish between P-Celtic and Q-Celtic, putting most of the Gaulish and Brittonic languages in the former group and the Goidelic and Celtiberian languages in the latter. The P-Celtic languages also called Gallo-Brittonic are sometimes seen as a central innovating area as opposed to the more conservative marginal Q-Celtic languages. In the P/Q classification schema, the first language to split off from Proto-Celtic was Gaelic.

Germanic

The common form that the languages of the Germanic branch had before they became differentiated is known as Germanic or Proto-Germanic. It antedates the earliest written records of the family and is reconstructed by linguists in the same way as is the parent Indo-European. The languages descended from it fall into three groups: East Germanic, North Germanic, and West Germanic.

The principal language of East Germanic is Gothic. By the 3rd century the Goths had spread from the Vistula to the shore of the Black Sea and in the following century they were Christianized by a missionary named Ulfilas (311–383) whose father seems to have been a Goth and his mother a Greek. Our knowledge of Gothic is almost wholly due to a translation of the Gospels and other parts of the New Testament made by Ulfilas. Except for some runic inscriptions in Scandinavia it is the earliest record of a Germanic language we possess.

North Germanic is found in Scandinavia, Denmark, Iceland, and the Faroe Islands. Runic inscriptions from the 3rd century preserve our earliest traces of the language. The later runic inscriptions in Scandinavia reflect changes specific to Old Norse in its earliest attestations, when it derived the modern North Germanic languages of the region. By the High Middle Ages a dialect division had appeared splitting Old Norse into East Norse, spoken in

Sweden and Denmark and West Norse spoken in Norway and Iceland after it was colonized in the 9th century.

West Germanic is of chief interest to us as the group to which English belongs. It is divided into two branches, High and Low German by the operation of a Second (or High German) Sound-Shift analogous to that described above as Grimm's Law. In early times we distinguish as Low German tongues Old Saxon, Old Low Franconian, Old Frisian, and Old English. The last two are closely related and constitute a special or Anglo-Frisian subgroup.

Old Saxon has become the essential constituent of modern Low German; Old Low Franconia, with some mixture of Frisian and Saxon elements. It is the source of modern Dutch in the Netherlands and Flemish in northern Belgium. Frisian survives in the Netherland province of Friesland in a small part of Schleswig in the islands along the coast and other places. High German comprises a number of dialects. It is divided chronologically into Old High German (before 1100), Middle High German (1100–1500), and Modern High German (since 1500). High German, especially as spoken in the midlands and was popularized by Luther's translation of the Bible. Since the sixteenth century it has gradually established itself as the literary language of Germany.

Balto-Slavic

The Baltic languages are part of the Balto-Slavic branch of the Indo-European language family spoken by the Balts. The family is usually divided into two groups. They are Western Baltic and Eastern Baltic. However, these are sometimes classified as separate branches of Balto-Slavic. The Western Baltic containing only extinct languages and Eastern Baltic containing both extinct and the two living Baltic languages viz., Lithuanian and Latvian. The Lithuanian, the Latvian, and the Old Prussian vocabularies differ substantially from one another and are not mutually intelligible. The now extinct Old Prussian language is considered the most archaic of the Baltic languages.

The various Baltic tribes were mentioned by ancient historians as early as 98 B.C. The first attestation of a Baltic language was in about 1350 with the creation of the Elbing Prussian Vocabulary, a German to Prussian translation dictionary. It is also believed that Baltic languages are among the most archaic of the remaining Indo-European languages even though their late attestation. Lithuanian was first attested in a hymnal translation in 1545; the first printed book in Lithuanian, a Catechism by Martynas Mazvydas was published in 1547 in Königsberg, Prussia. Latvian appeared in a hymnal in 1530 and in a printed Catechism in 1585.

The traditional view is that the Balto-Slavic languages split into two branches, Baltic and Slavic with each branch developing as a single common language (Proto-Baltic and Proto-Slavic) for some time afterwards. Proto-Baltic is then thought to have split into East Baltic and West Baltic branches. However more recent scholarship has suggested that there was no unified Proto-Baltic stage, but that Proto-Balto-Slavic split directly into three groups. They are Slavic, East Baltic and West Baltic. Under this view the Baltic family is paraphyletic and consists of all Balto-Slavic languages that are not Slavic. This would imply that Proto-Baltic, the last common ancestor of all Baltic languages would be identical to Proto-Balto-Slavic itself, rather than distinct from it.

Anatolia

The oldest known family of Indo-European languages is Anatolian. Hittite the oldest and most richly attested of these languages was spoken in Central Anatolia. The first Hittite clay tablets were found at the end of the 19th century interpreted starting in 1915. The people that we call Hittites did not in fact call themselves Hittite, nor their language Hittite, their kings called their land the “Hatti Land” and Hittite scribes called their language “Nes(h)ite”. Archaeological excavations are still uncovering artifacts including new clay tablets.

By the late 20th century the term was most commonly used to designate the so called Anatolian group of Indo-European languages, Hittite, Palaic, Cuneiform Luwian, Hieroglyphic Luwian, Lycian, Lydian, Carian, and possibly Pisidian and Sidetic. Hittite, Palaic, and Cuneiform Luwian are known from 2nd millennium. Cuneiform texts found mainly in the ancient capital of the Hittite empire. Hieroglyphic Luwian is found on seals and inscriptions from approximately 1400 to about 700 BC. Lydian, Lycian and Carian are known from texts in alphabetic script from approximately 600 to perhaps 300 BC. Although there is evidence enough to suggest that they belong to the Anatolian group Sidetic (300 – 100 BC) and Pisidian (1–200CE) are very poorly attested languages.

Tocharian

Tocharian is an extinct branch of the Indo-European language family formerly spoken by Tocharian peoples in oases on the northern edge of the Tarim Basin (now part of the Xinjiang Uyghur independent Region of China). Documents dating from the 6th to the 8th century AD record two closely related languages called Tocharian A (“East Tocharian”, or Turfanian) and Tocharian B (“West Tocharian” or Kuchean). The subject matter of the texts suggests that Tocharian A was more archaic and used as a Buddhist language while Tocharian B was more actively spoken in the entire area from Turfan in the east to Tumshuq in the west.

A body of loanwords and names found in Prakrit documents has been dubbed Tocharian C (Kroranian). These languages became extinct after Turkic Uyghur tribes expanded into the Tarim Basin in the 9th century AD. Prakrit documents from 3rd century Kroran contain loanwords and names that appear to come from Tocharian C. The existence of the Tocharian languages and alphabet was not even suspected until archaeological exploration of the Tarim basin in the early 20th century brought to light remains of manuscripts in an unknown language dating from the 6th to 8th centuries AD. It soon became clear that these remains were actually written in two distinct but related languages belonging to a previously unknown branch of Indo-European now known as Tocharian.

Sound Change

Written records of earlier speech, resemblances between languages and the varieties of local dialects all show that languages change in the course of time. Sound changes are regular and they come under some rules. Sound change arise a language for so many reasons. Some changes come under physical character, social activities, simplicity and laziness. If we refer to sangam literature and middle Tamil / modern Tamil text we may find lot of changes. For instance in sangam Tamil we have initial ya- changed in to a: in modern Tamil.

ya:Ru → a:Ru - six
ya:r → a:r - who

ya:nai → a:nai - elephant
ya:Du → a:Du - goat
ya:NDu → a:NDu- year

Such types of some other changes are also occurring when we compare Old Tamil and Modern Tamil. In the pronunciation level there is no difference between the following sounds. For example l – L; L – l; R – r; n – N; ŋ - n; and n - n etc. These two types of sounds or phonemes are merging into one sound.

Example

(l) (l)
palam → palam - fruit
e:lu → e:lu - seven
va:lai → va:lai - plantain

(R) (r)
kaRi → kari - charcoal
kuRitta:n → kuritta:n -
kaRai → karai

Similarly in English the medial *u:* in old English changes into *au* in modern English. For instance

(u) (au)
mu:s → mouse
u:t → out
su:o → south

From above these examples we came to know that sound changes are regular. Changing sounds which have some similarities. The sound or phoneme of a language is changes on the basis of the manner of articulation / place of articulation. i.e., similar sounds are changed into another. When we compare two different languages in a language family we can find number of changes. For instances ai ending Tamil words changed in to / a / in Malayalam.

Tami		Malayalam
cilai	→	cila - statue
talai	→	tala - head
malai	→	mala - mountain
varai	→	vara - till

These types of changes commonly occur in these languages and it comes under some rules. If we compare above these examples (l) changed into (l) and (R) changed into (r) which are not comes any condition or influence of the neighboring sound. These types of sounds changed in all the places without any specific reason. We could not give condition of the following or proceeding sound features. Therefore there is no matter when it occurs or what sounds are nearby. Such a sound change is known as unconditioned change. So in this type we

would not give surrounding situation at the time rule frame. Opposed this concept a sound changed in to another with some reason or similar feature is called conditioned sound change.

Phonetic change

Any phonological change which affects only the phonetic realization of one or more segments and has no consequences for the phonological system. Phonetic change is one which merely affects the pronunciation of a given phoneme without altering the phonemic system of the language. For instance the Tamil consonant /k/ has undergone various types of changes, but this is purely a phonetic change so long as /k/ remains distinct from all other consonants and retains its distribution on words. For instance,

k >	{	g	}	makan >	magan	}	- son
		h			mahan		
		y			maxan		
					mayan		

Phonemic change

In this change the rule merely adds another allophone when a sound changed into another. It does not change the status of the phoneme. However when the numbers of contrasts are either increased or decreased the phonemic system is changed. The outcome of a phonological change a segment changes its phonetic nature in such a way as to shift from one phoneme to another. In Tamil due to social setup the phoneme /c/ has three allophonic variations.

/ C /	(i)	[c]
	(ii)	[s]
	(iii)	[j]

Example:

cey	>	sey	- to do
celavu	>	selavu	- expenditure
cempu	>	sembu	- copper
ceTi	>	seDi	- plant
ca:r	>	sa:r	- sir
santarppam	>	santarppam	- chance
cankam	>	sangam	- association
co:Ru	>	so:Ru	- rice
civappu	>	sevappu	- red
ca:ti	>	ja:t i	- caste
ca:n	>	ja:n	- measurement
ca:kkiratai	>	ja:kkiratai	- careful

After palatal sound the /c/ becomes /j/ in modern Tamil.

Example

nencu > nenju - chest

pincu > pinju - tender

kencu > kenju - beg

Major types of sound change

Split and Merger:

Any phonological change a single phoneme gives rise to two distinct phonemes. i.e., bifurcation of phonemes is known as split. For instance in Tamil the phoneme /c/ split into two distinct phonemes. i.e. /c/ split into /s/ and /j/. Split classified in to two types viz., primary split and secondary split. In primary split on outcome of them split immediately merges with another existing phoneme, so that the total number of phonemes remains and changed. For example early Latin \s\ changed to (r) between vowels and merged there with the existing \r\, so that only the distribution of \s\ and \r\ changed. In Tamil the phoneme /l/ split into /L/ and /y/ and both /L/ and /y/ are in the system of Tamil. The /l/ split in to two and immediately merge with the existing system of the language. In secondary split, both outcome of the split finds anything to merge with and hence, the total number of phonemes increases. For example, Tamil /c/ split in to two which are not in the system. So the outcome increase the sound system of the language is known as secondary split. Frequently, though not invariable, secondary split results from loss of conditioning environment. For instance in Tamil,

nencu	>	nenju	- Chest
tankam	>	tangam	- gold

Merger:

The loss of a contrast formerly existed between two or more phoneme. We understand that two phonemes are joined together or combined together. For instance Old Indo Aryan /S/ merge with /s/ in Modern Indo Aryan

śatya	>	sacca	- Truth
śatam	>	sata	- Hundred
da śa	>	dasa	- Ten

/c/ & /ch/ merge into /s/ in Assamese

Candrah	>	sa:nd	- Moon
Chattram	>	sa:ti	- Sunshad

/l/ & /L/ merge into /l/ and /R/ & /r/ merge into /r/ in Tamil.

In an unconditioned merger, the contrast disappears in every case and the number of phonemes is reduced. For example, the historical Basque contrast between laminal /s/ and apical /s' / has been lost in western varieties in favour of /ś / - that is the laminal has become

apical in every case. In a conditioned merger, the contrast only disappears in specified environments. For example, in southern varieties of American English, the contrast between /e/ (as in bed) and /I/ (as in bid) has been lost in favour of / I / before a nasal. Trudgill (1978) distinguish two mechanisms for merger. In merger by transfer, lexical items shift individually and abruptly from the class defined by one phoneme to the class defined by the other phoneme, until the first class is empty; this is a kind of lexical diffusion. In merger by approximation the phonetic realization of the two phonemes moves steadily closer together until they coincide. Labov (1994: 321-323) adds a third type i.e., merger by expansion, in which the contrast is abruptly abandoned and the phonological space formerly occupied by two phonemes is reassigned to the single new phoneme.

Conditioned and unconditioned sound changes

A conditioned sound change is one which is caused by some segment in the environment of another. A clear instance of this is *i*-umlaut which is caused in a given syllable by the high vowel or / j / in a following syllable. It is important to note that this change leads to morphological irregularity as in the following cases in German.

Jung > jünger
 Gut > güte
 Hohn > höhnisch

In English old: older ~ elder shows the semantic exploitation of the analogically regularised form and the original umlaut form. The latter is found when referring to siblings and in one or two set phrases: my elder sister, an elder statesman; otherwise older is used. It is the case that sound changes are influenced or conditioned by neighboring phonemes. In this change we should give conditioned environment. For instance in Tamil

k → **g / n-**

tankam → tangam – Gold

cankam → cangam - Association

panku → pangu - Share

tt → **cc /** $\left\{ \begin{array}{c} \mathbf{y} \\ \mathbf{i} \\ \mathbf{ai} \end{array} \right\}$

y → **∅ / -cc**

pa:y+tt+u → pa:yccu

aTi+tt+u → aTiccu

aTai+tt+u → aTaccu

vai+tt+u → vaccu

ka:y+tt+u → ka:ccatu

Above example we identify the two sound changes. One is /tt/ → /cc/ and /ya/ is lost. Which one is changed first? /cc/ changed first. /tt/ changed into /cc/ when it proceeded with /ya/, /i/, and /ai/. Otherwise it is not possible. *ka:ttatu*, *pu:ttatu*, *akattatu* and *ottatu* are the words in Tamil in which the (tt) never changed into (cc), because the sound not proceeding with /ya/, /i/, and /ai/. This type of sound change is called conditioned sound change.

Unconditioned sound change affects every possible segment which matches its input, i.e. it is not dependent on conditioned by its environment. An example of this would be the diphthongisation of Middle English /i:/ and /u:/, which does not cause any grammatical irregularity; the loss of /x/ in Middle English is another instance of unconditioned sound change.

Assimilation

Any syntagmatic change in which some segment becomes more similar in nature to another segment in the sequence, usually within a single phonological slot or phrase. The most common type of conditioned sound change is assimilation, where by one should becomes more like a neighboring one. This can be considered a simplification of the muscular movements needed to pronounce a given word. Assimilation of consonants usually involves one consonant becoming more like another. For example in Tamil

cempaTTai	>	cemmaTTai	- Brown colored hair
captam	>	cattam	- Sound
camarttu	>	camattu	- Smartness
senbagam	>	sembagam	- Name of a woman
sambantham	>	sammantham	- Name of a man
vilu	>	vulu	- Fell down
no:nbu	>	no:mbu	- Ritualistic observance of fasting
anRa:Tanka:ycci	>	anna:Tanka:cci	- Daily labour
anRa:Tam	>	anna:Tam	- Daily

There are two types of assimilation viz., progressive assimilation and regressive assimilation. A right to left assimilation is Regressive; a left to right assimilation is progressive. Regressive assimilation is that a consonant becomes more like one the follows; in the words the force of the change proceeds backwards from a phoneme to the one which precedes it. Example:

enbadu	>	embadu	- Eighty
nancey	>	nañcey	- Land

Progressive assimilation takes place when the first phoneme is dominant and in some way makes the second more like itself.

Example:

kattikku > kattikki - For knife
ve:lvikku > ve:vikki -

Assimilation in both directions at the same time is mutual assimilation. Assimilation between adjacent segments is contact assimilation, while that between non adjacent segments is distant assimilation. Assimilation in some phonetic features only is partial assimilation; assimilation in all phonetic features is total assimilation. For example, the pronunciation of ten pence te[m] pence is partial anticipatory contact assimilation; that of bacon as /beIkŋ/ is partial progressive contact assimilation. The change of earlier Basque alte ‘side’ into alde in most dialects is partial progressive contact assimilation, while the change of Pre-Basque bini ‘tongue’ to later mini is partial anticipatory distant assimilation. The development of Latin eclipse (m) ‘eclipse’ into Italian eclisse is total anticipatory contact assimilation, while the development of Proto – Germanic mu:siz ‘mice’ to Old English my:s is partial anticipatory distant assimilation.

Types of Assimilation

Partial versus complete assimilation: Does the assimilating sound take on every feature of the phoneme that activated assimilation (complete) or only some of or even just one of the features (partial). For instance the complete assimilation in Tamil:

cempaTTai > cemmaTTai ‘brown colored hair’
captam > cattam ‘sound’

On the other hand, the no:nbu > no:mbu ‘ritualistic observance of fasting’ undergoes a partial assimilation.

Contiguous versus non-contiguous assimilation: Are the assimilating sounds are closest to one another is contiguous assimilation. Opposed to this the assimilating sounds more distant are known as non-contiguous assimilation. Two sounds side by side (contiguous) or are they more distant (non-contiguous)? For instance the words are given below involved in contiguous phonemes.

campantam > cammandam
campantappaTuttu > cammandapaTuttu
campanti > cammandi

Non-contiguous assimilation

anRa:Tanka:ycci > anna:Tanka:cci - daily labour

anRa:Tam	>	anna:Tam	- daily
veTkam	>	vekkam	- shame

Assimilation by voicing and devoicing

Nearly all vowels in human languages are voiced and voiceless contrast in consonants. In some languages voiceless consonants occur / comes between vowels (voiced sound) in a word changed into voiced. Tamil speakers from Dravidian family pronounce the phoneme / p /, / k / the voiceless stop sound between vowels as / b /, / k / like **ampu** – **ambu** ‘arrow’, **ulagam** – **ulagam** ‘world’. The changes / p / to / b /, / k / to / g / is an instance of intervocalic voicing. Opposed to this situation voiced changed into voiceless when the speaker pronounce some words is devoicing. For example the word *has* ends in the voiced phoneme /v/. The word *to* begin with voiceless /t/, some speakers devoice /v/ to /f/ when they pronounce the expression *have to* as *hafta*.

Palatalization

Due to assimilation some sound moves its place of articulation closer to the palate. This happens when the sound it’s assimilating to already have a palatal or near-palatal place of articulation. In Tamil / T / the retroflex stop changed in to palatal in favour of following another palatal. In some situation the dental / t / occur in cluster / tt / changed in to palatal / cc / when it proceeded with high front vowel.

Example:

aravaNaittu	>	aravaNacci	- Embraces
tanittu	>	tanicci	-Solitary
tiruva:Tci	>	tiruva:cci	– medicinal plant

Velarization

A phoneme can assimilate to any place or manner of articulation. Assimilation to specific places of articulation includes the assimilating sound becomes more velar is known as velarization. Simplification of the muscular moment and laziness of pronunciation the Tamil words like **veTkam** ‘Shame’ and **veTkku** changed into **vekkam** and **vekku** respectively. The retroflex /T/ favourably changed into velar /k/.

Dissimilation

Any syntagmatic change in which one segment changes so as to become less similar to another segment in the same form. Two sounds are involved and one becomes less like the other. The dissimilating phoneme loses one or more of the features it shares with another nearby phoneme. The two phonemes become more dissimilar as a result. For example in Tamil the following words are changed with dissimilar sounds.

piyttu udaru	>	piccu odaru
pa:kaRka:y	>	pa:vakka:

va:nam	>	ma:nam
talaiyaNai	>	talava:Ni
to:cai tiruppi	>	to:cai tiluppi
tu:pakka:l	>	tu:vakka:

The English phoneme /p/ is a voiceless bilabial stop. If you say Peter Piper picked a peck of pickled peppers as fast as you can, chances are you dissimilate one or more /p/'s by pronouncing them as a /f/. The expected Ancient Greek verb /thethne:ke/ '(s)he has died' was actually pronounced /tethne:ke/. The first aspirated /th/ dissimilated from the following /th/. The Latin root peregrin- comes into English as pilgrim rather than pirgrim. The first /r/ dissimilates to a /l/. Latin arbore (m) 'tree' yields Spanish arbol with dissimilation of the second /r/ to /l/. Latin libellum 'level' yields French niveau (with dissimilation of the first /l/ to n., plus other changes). Latin anima 'soul' is borrowed into Basque as arima with dissimilation of the first nasal. The extreme case is dissimilation loss, when Basque santso 'Sancho' (Name) became Antso, with dissimilatory loss of the first sibilant.

Metathesis

Any syntagmatic change in which the order of segments in a word is altered is known as metathesis. In the simplest case, a single segment changes its position as in the development of Latin **crocodiles** 'Crocodile' to Spanish **cocodrilo**, with metathesis of the /r/. In more elaborate cases, two segments exchange their position; for example, Latin **miraculum** 'Miracle' should have yield Spanish **miraglo**, but the form is **milagro**, with metathesis of the two liquids. The label hypothesis is sometimes applied to an instance of metathesis between two segments very far apart. An instance of quantitative metathesis is provided by the development of earlier Greek /ne:os/ 'new' into Attic /neo:s/, with metathesis of the length. A still more complex example is provided by the Basque word for 'swallow' (the bird): common **ainara-enara** must derive from earlier **ain:ala**, but western **elae** points equally clearly to earlier **ail:ana**, with metathesis of /n/ and /l/ plus transfer of the length from one to the other.

Epenthesis

Any phonological change in which a sound is added in the position of consonant cluster which was not formerly present in a word is known as epenthesis. That is addition a vowel or consonant is employed in the position of consonant cluster. The new phoneme is called an epenthetic sound.

Example:

Old Indo-Aryan	Modern Indo –Aryan
prthvi	prithivi - earth
sneha	saneha - friend
Latin	Spanish
spina	espina - thorn

Middle English	Modern English
thuner	thunder

Loss of sounds: Syncope

Loss of vowel between consonants in a polysyllabic word is called syncope. i.e., the loss of a vowel from interior of a word in many varieties of English like fam(i)ly, mem(o)ry. Example:

Latin	Old Irish	
apostolus	apstal	- Postal
episcopus	epscop	- Bishop
na:ta:licia	notlaic	- X – mas

Latin	Spanish	
manicam	manga	- Sleeve
septimana	semana	- Week
viriden	verde	- Green'
temporanum	temprano	- early'
stabilem	estable	- Steadily

Apocope

Loss of vowel is especially common in languages with strong stress on one syllable of a word. As a result of the emphasis on the stressed syllable other syllable in the word tends to become reduced and may be lost. Vowel is lost at the end of the word; we refer to this loss as apocope. Narrowly, the loss of a word final vowel, as in the development of Old Spanish **mercede** and **pane** into modern Spanish **merced** 'mercy' and **pan** 'bread'. Loss of vowels in final position is also there in English. For instane,

Old English	Modern English
sticca	stick
so:nu	son
mo:na	moon

Apheresis

Loss of vowel in word initial position is called Apheresis. The loss of a word –initial vowel, as when **opossum** is reduced to **possum**, or when early Italian **istoria** 'history' was reduced to **storia** are under this category of change.

Haplology

Haplology is a type of phonological change in which one of two adjacent syllables of identical or similar form is lost or fails to appear in the first place. For example Latin **nutri** ‘nourish’ plus **-trix** ‘female agent’ should have yielded **nutritrix**, but the form is **nutrix** ‘nurse’; Basque **sagar** ‘apple’ plus **ardo** ‘wine’ should yield **sagar-ardo**, but the form is **sagardo** ‘cider’; English **library** is pronounced in England as though it were **libry**; Greek **amphi** – ‘on both sides’ plus **phoreus** ‘bearer’ yields Homeric Greek **amphiphoreus** ‘**amphora**’, reduced in classical Greek to **amphoreus**; Old English **Anglaland** ‘land of the Angels’ yields modern English **England**.

Addition of sounds

Prothesis

The introduction of an extra initial sound i.e., The addition of a segment to the beginning of a word, as in the development of Latin *stannum* ‘tin’ into Spanish *estano*, the borrowing of Latin *regem* ‘king’ into Basque as *errege*, or the development of pre – Motu **au* ‘me’ to Motu *lau*. Some linguists restrict this term to the addition of an initial vowel. Vowels are lost due to heavy stress on one particular syllable. Weakening of articulation may lead to the introduction of vowels. So vowel is thus introduced in the beginning of a word especially before [s]. The opposite is aphaeresis.

Example:

Latin	Spanish	
spiritus	espíritu	‘Spirit’
Schola	escuela	‘School’

Neogrammarian theory of gradualness

A label Neogrammarian is originally applied to a group of young historical linguist at Leipzig University, Germany who first declared the Neogrammarian hypothesis. It was seen by most established linguists. The most outstanding Neogrammarians were Karl Brugmann (1849 – 1919), Berthold, Delbruck, August Leskien and Hermann Osthoff. Neogrammarians originate the hypothesis to regulate the sound change in the given level.

- Sound changes are gradual
- Sound change is regular
- Sound change is mechanical
- Sound change is change in performance

There are seven interdependent variables associated with any sound change.

- ✓ Time
- ✓ Space
- ✓ Speakers and hearers

- ✓ Replaced and Replacing sounds
- ✓ Structural conditions for replacement
- ✓ Social factors governing replacement
- ✓ Lexical items that fulfill the structural and social conditions of replacement.

The hypothesis that all phonological change is regular often formulated as follows. Every sound change takes place according to laws that admit no exception. In this view, a phonological change must apply absolutely and simultaneously to every linguistic form in the language which exhibits the relevant phonological form. Traditional historical linguists have considered variables a, b, d and e as central to the study of sound change, while the remaining have been under invariant factors. This regularity implied that at a given time and place, all lexical items that fulfilled the structural conditions of a sound change in the speech of all speakers in all social settings. Gradualness has been studied in terms of the same four variables viz., Temporal, Spatial, Phonetic and Gradualness of sound change along social perceptual and lexical dimensions.

There are some words in any language which stand out as exceptions. Among these are the imitative or expressive words in which speakers believe to imitate various natural sounds. For instance the word **cuckoo** in Greek is **kokku-ks** and Latin **cucu-lus**. According to Grimm's Law we would expect something like **huhu** in English. English must not have a **k** in the word matching Greek and Latin **k**. The imitative and expressive words are not as completely arbitrary as most of the words in the language.

Social factor is the social attitude of the speech community toward particular linguistic forms. These may be comes under the headings of taboo which provides a motivation or cause for the appearance disappearance and change of linguistic forms. Taboo is defined as avoidance of particular words for various social reasons. A word may be avoided for reasons connected with religion, superstition, personal respect or social attitudes towards bodily functions or other matters. When words are so avoided, they are substituted by euphemism. In our society, one does not ordinarily speak of one's parents, teachers, by name. The wife does not take the name of her husband but uses like Mohan's father etc. In connection with certain things we never use the negative expression. We do not like to say that we do not have that thing, though that is the right expression. We use instead phrases of the meaning it is in excess. Here belong 'rice' 'salt' and objects such as turmeric and saffron.

Comparative philology is a term for the study of the historical development of the languages of the Indo-European language family. This technique was evolved by the group of linguists working in the second half of the 19th century, mostly in Germany. Going on a German term **Junggrammatiker** used first in disrespect for young linguists, the English translation Neogrammarians was devised. What is here called the Neogrammarian view refers to the methods used for linguistic reconstruction in the 19th century. It first arose towards the end of the 18th century with the discovery that classical languages like Greek, Latin and Sanskrit are related to each other. Soon after this a number of scholars independently established interconnections between the languages of the Germanic, Romance, Celtic, Slavic and Baltic groups of languages. Notable among the authors of this earlier group is Jakob Grimm who established a series of sound laws which applied to Germanic in its earlier stages.

Well different Laws of sound change:

Sound change was first achieved in trying to point out relationship between Germanic sounds. The phonemes of Indo-European are in to three main groups.

1. Consonant
2. Vowel
3. Resonant

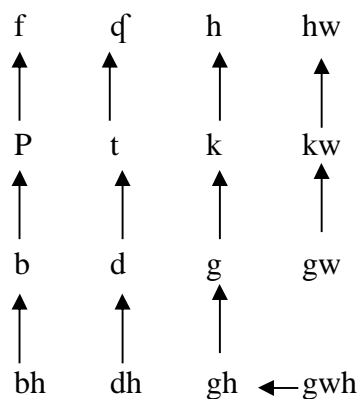
The term resonant are phonemes (e.g. r & l) which can function as either consonants or vowels and are involved in the system of alternation known as apophony. Indo-European had the following consonants:

Stops	Labial	Dental	Velar	Labio - velar
Voiceless	p	t	k	kw
Voiced	b	d	g	gh
Voiced aspirates	bh	dh	gh	gwh
Fricative		s		

The Indo-European consonant phonemes are given above together with their reflexes in Germanic languages. The Germanic consonant system looks like this

f	p	h	hw
P	t	k	kw
b	d	g	-

Putting these together with arrows showing the direction of the developments



We can summarize the Germanic consonant shift in three simple rules.

Voiceless stops - voiceless fricatives (p-f)

Voiced stops - voiceless stops (b-p)

Voiced aspirated stops - voiced unaspirated stops. (bh-b)

It was the identification of exactly these sound changes from Indo-European to Germanic which set in moving linguistics. It was first noted by the Danish scholar Rasmus Rask in his prize essay of 1818. Later further supporting examples from Sanskrit Jacob Grimm indicated Germanic representation or Indo-European sound in his 2nd edition of German grammar 1822. Grimm started with Latin, Greek & Sanskrit. As bases and looked upon the differences between these languages and Germanic as some kind or change. To use Grimm's term they were called sound shift. In later times this sound shift or the consonant system from Indo-European to Germanic came to be known as Grimm's Law.

Another major development in the consonants of Proto-Germanic was covered by Karl Verner (1876). The change or development addressed by Verner was a set of Germanic forms which appeared to be exceptions to the First Germanic Consonants Shift. Specifically the Proto Indo European / **p t k** / develop into voiced plosives / **b d g** /. For instance, the inherited / **p t k** / of Greek *hyper* 'over', *pater* 'father' and *hekyra* 'mother in law' are matched by / **b d g** / in old High German *ubar*, Old English *fæde* and old High German *swigur*. There appeared to be no possible conditioning environment for this irregular development in Germanic. However upon examining the forms in Sanskrit and Greek which largely preserve the Proto Indo European accent. Verner realized that a conditioning factor existed in the form of Proto Indo European accent. The consonant shifts occurred at a time when early Proto-Germanic still preserved the PIE accent. After the shifts, Germanic shifted the word accent to initial position in all cases destroying the evidence for the conditioning environment within Germanic itself. So that only a comparison with accentually conservative languages like Greek and Sanskrit could reveal what had happened.

Verner's formulation immediately accounted for the previously mysterious Germanic alternations helplessly dubbed grammatical change. Compare the Sanskrit verb *varṭate* 'turn' and its old English cognate *weorpan* 'become', in four cognate forms. Verner's Law succeeded in removing all remaining exceptions to Grimm's Law and its success provided the immediate inspiration for the neogrammarian hypothesis. In early and mid 19th century certain other exceptions were also accounted. In some Germanic words **b**, **d**, and **g** were found corresponding to Latin **p**, **t**, and **k**. From Grimm's Law the expected correspondences should have been f, ð and h. In the ancient Germanic languages there were also certain alternations were unexplained correspondences. For instance the medial consonants in the following data show the alternations.

Old High German

slahan 'to strike' gi-slagan 'struck'

Old English

snifan 'to cut' sniden 'brother'

Latin

frater 'brother' broþor 'brother'

Here we see

h → **g**
þ → **d**
f → **b**

In order to show Karl Verner solved this problem.

Old English	Latin	Sanskrit	
fæder	pater	pitar	'father'
bro:dor	fra:ter	bhra:tar	'brother'
mo:dor	ma:ter	ma:tar	'mother'

We will reconstruct the following I.E ancestor:

pəte:r, bhra:ter, ma:ter

Verner's Law which operates after Grimm's Law says that when f, þ and h are preceded by unaccented vowel, they become the corresponding voiced stops.

f	→	b	V [-accent]
þ	→	d	
h	→	g	

Sound law

This is a term which refers to a change or a series of related changes in the phonology of a language. The term law is used to stress the regularity of the change. The classic example of a sound law is the so-called Germanic sound shift or Grimm's Law. This states that in the early stage when Germanic was differentiating itself from the remaining dialects of Indo-Germanic, all voiceless stops were shifted to voiceless fricatives, i.e. /p/ became /f/, /t/ became /θ/ and /k/ became /x/.

Latin	English	Old English
Pes	foot (p ~ f)	fot
Tres	three (t ~ þ)	þreo
Collis	hill (k ~ x)	hyl

Quod

what (kw ~ xw)

hwæt

With *hill* and *what* one must consider the development of the /x/ sound in the history of English. Within the Old English period /x/ was weakened to /h/ in word position and represented as *h* in writing. From this fact one can conclude that those words which have been inherited from Germanic and which are written with initial *h* originally had /x-/. The sequence /xw/ in Germanic later developed analogically to /hw/. Through assimilation of the /w/ to the preceding voiceless glottal fricative the voiceless [x] which was voiced in the early modern period in southern English but which is retained in many conservative dialects.

Verner’s Law The sound law was the central theoretical concept in comparative philology so that exceptions to laws represented a considerable difficulty to linguistic interpretation within this model. The results of a sound law could be implied by later analogy which did away with irregularity in morphological paradigms. But still for Germanic, there were disturbing exceptions to the operation of established sound laws which could not be explained by an appeal to analogy. Scholars had noticed at an early stage that the Germanic consonant shift did not affect all instances where it might have applied. Grimm called these exceptional cases as ‘morphological alternation’ because they were to be found in verb paradigms.

In 1877 the Danish scholar Karl Verner published an article entitled “An exception to the Germanic sound shift”, in which he demonstrated that the apparent irregularity here constituted in fact a clear rule. Since then this phenomenon bears his name. In essence Verner’s Law runs as follows: If a voiceless consonant occurred in a syllable which was preceded by another which did not carry the main accent then this consonant was voiced unless the phonetic environment blocked this. Now because accent was relatively free in Indo-European there were many words which were affected by Verner’s Law. Later with the rise of initial accent in Germanic the environment for Verner’s Law was no longer available but there were reflexes of the earlier situation.

Indo-European	Old English	
pətér	fæder	‘father’

With the original word for ‘father’ the accent occurred after the /t/. This was changed as part of the Germanic sound shift to /θ/ and the accent was not on the preceding syllable, yielding /ð/ which was later fortified to a plosive. Another major sound shift in Germanic is the High German sound shift which explains many differences between English and German. The changes can be shown in a table with sample words as follows.

word initially		word medially	
/p/	/pf/	/p/	/f/
/t/	/ts/	/t/	/s/
/k/	/kx/	/k/	/x/
English	German	English	German
path	pfad	open	offen

tooth	zahn	eat	essen
church	kirche	make	machen

In the course of the 19th century examples for sound laws were described for many languages of the Indo-European family. Linguists stressed the exception less character of sound laws. According to this view the only exceptions to a valid sound law can occur due to the force of analogy. This is where an expected change is not found because another element in a grammatical paradigm has caused a change to be blocked a change at some later point. Thus the change of /s/ to /r/ in the past tense of the verb *sein* in German has been masked by the analogical spread of /r/ to all elements of the verbal paradigm, e.g. *ich war, du warst*, etc. In English we can see that an alternation originally existed: *I was, you were*, etc.

Lexical diffusion

In 1969, Wang proposed the hypothesis that the mechanism of sound change may be lexical diffusion, a notion related to but distinct from, Labov's variable rules. Under this hypothesis sound change application in many cases is specified individual lexical items.

Lexical diffusion refers to the way a sound change affects the lexicon. If sound change is lexically abrupt, all the words of a language are affected by the sound change at the same rate. If a sound change is lexically gradual, individual words undergo the change at different rates or different times. Whether sound changes exhibit gradual or abrupt lexical diffusion is a topic that surfaces persistently in historical linguistics, but as yet has not reached resolution. William labov's view of lexical diffusion is that it has only a very limited role to play in change. He says 'there is no evidence that lexical diffusion is the fundamental mechanism of sound change.' it happens but is only a complement and a small one at that to regular sound change. The most important factors in linguistic change appear to be long standing trends in the language, internal variation and social forces among speakers.

Lexical diffusion is the analogical generalization of lexical phonological rules. In the early articles by william wang and his collaborators, it was seen as a process of phonemic redistribution spreading rapidly through the vocabulary. Subsequent studies of lexical diffusion have supported a more constrained view of the process. They have typically shown a systematic pattern of generalization from a categorical or near categorical core through extension to new phonological contexts, which are then implemented in the vocabulary on a word by word basis.

The term 'lexical diffusion' is frequently employed in the context of phonology; there has been an increasing awareness in recent studies that the same concept is often applicable to syntactic changes as well. It maintains that less attention seems to have been paid to the problem of regularity versus lexical diffusion in syntax. At the same time wang argues that both morphology and syntax, lexical diffusion seems to have been implicitly taken for granted by many writers.' likewise, points out in the context of syntactic developments the fact that 'the incoming form does not spread to all contexts at once but some acquire it earlier than others,' and says that the phenomenon is called 'lexical diffusion.' in this manner the concept of lexical diffusion is extendable to various linguistic changes, including syntactic ones.

Exceptions to sound change

Analogy

Any linguistic change in which results from an attempt to make some linguistic forms more similar to other linguistic forms. Analogy is the basic form to make form in any language. (Pen+s - pens, tablet+s tablets, walk+ed, talk+ed etc). Most of the changes in a language can't be explained by operating sound laws. Sound laws which are operate only with regard to phonetic factors. Certain other important types of change have been grouped under the heading of analogy.

Analogy "it is a process whereby one form of a language becomes more like another with which it is somehow associated". Analogy does not create any new forms or categories in a language. Analogy only causes the spread or extension of already existing items. Arlotto (1972) has recognized the problem and offers a very general definition for it ... 'analogy is a process whereby one form of a language becomes more like another with which it has some association.' In other words, analogy is mostly everything that is not sound change or borrowing. Analogy thus becomes the default category of changes. In analogical change some instances of the language change become more like another in the language where speakers perceive the changed part as similar to the items which bring the change.

Analogy is sometimes referred as 'internal borrowing' with an idea that in analogical change a language may 'borrow' some patterns in the same language itself. Because analogy is usually not conditioned by regular phonological factors, rather it depends on aspects of the grammar especially morphology. In order to make little better sense of what we have been saying let's see some example: sorry and sorrow were quite separate words in English, but in its history sorry has changed under influence from sorrow to become more similar to sorrow. Sorry is from the adjective form of 'sore', Old English sarig 'sore, painful, which has cognates in other Germanic languages. The original /a:/ of sarig changed to /o:/ and then was shorted to /o/ under influence from sorrow which had no historical connection to sorry. This is an analogical change, where the form of sorry changed on analogy with that of sorrow.

Kinds of analogical changes

Proportional Analogy

Proportional Analogical changes are those which can be represented in the following principle, i.e. $a : b = c : x$ i.e. for example ride : rode = dive : x, where in this example x is solved with dove. The original past tense of dive was dived, but it changed to dove under analogy with the class of verbs which behave like drive: drove, ride: rode, write: wrote and so on. In English the pattern of the verb speak / spoke / spoken developed through remodeling on analogy with verbs of the pattern break / broke / broken.

Analogical levelling

Analogical leveling reduces the number of allomorphs in form i.e. it makes paradigms more uniform. In analogical leveling forms which formerly underwent alternations no longer

do so after the change. For instance, some English ‘strong’ verbs have been levelled to the ‘weak’ verb pattern, as for instance in dialects where throw / threw / thrown has become throw / throwed / throwed. There are numerous cases throughout the history of English in which strong verbs with stem alternations as in sing / sung / sung or write / wrote / written have been levelled to weak verbs.

The comparative and superlative form of old has been levelled from the pattern old / elder / eldest to the non-alternating pattern old / older / oldest. Here /o/ had been fronted by umlaut due to the former presence of front vowels in the second syllable of elder and eldest. But the effects of umlaut were levelled out and now the words elder and eldest remain only restricted contexts not as the regular comparative and superlative of old.

Analogical Extension

Analogical extension has extended the already existing alternation of some pattern to new forms which did not formerly undergo any change. An example of analogical extension is seen in the case mentioned above of dived being replaced by dove on analogy with the strong verb pattern as in drive / drove, ride / rode and so on. In modern English wear / wore, which is now in the strong verb pattern. It was historically a weak verb which changed by extension of the strong verb pattern, as seen in old English werede ‘wore’ which would have become wearied if it had survived. The development of the non standard past tense forms which show extension to the strong verb pattern which creates alternations that formerly were not these as in arrive / arrove (standard English arrive / arrived) and squeeze / squoze (standard squeeze / squeezed).

Relationship between analogy and sound change

The relationship between sound change and analogy is captured reasonably well by the following slogan. Sound change is regular and causes irregularity; analogy is irregular and causes regularity. The slogan is that a regular sound change can create alternations or variant allomorphs. For example umlaut was a regular sound change in which back vowels were fronted due to the presence of a front vowel in a later syllable as in brother + -en > brethren. As a result of this regular sound change, the root for brother came to have two variants brother and brethren.

Previous English forms had many alternations of this nature. However an irregular analogical change later created brothers as the plural form on analogy with the non alternating singular / plural pattern in such nouns as sister / sisters. This analogical change is irregular in that it applied only now and then here and there to individual alternating for and not across the slot. This analogical change in the case of brethren in effect resulted in undoing the irregularity created by the sound change leaving only a single form brother as root in both the singular and plural forms, i.e. analogy leveled out the alternation left behind by the sound change.

The history of the verb to choose in English shows the interaction of analogy and sound change. In Old English we had the forms ceosan ‘infinitive’, ceas ‘past singular’, curon [kuron] ‘past plural’ and coren [koren] ‘participle.’ These come from the Proto – Indo – European root

*geus- ‘to choose.’ From this Indo-European root came Proto-Germanic *keus-an. The differences in the consonants among the Old English forms of ‘to choose’ come from two sound changes. The participle and past plural forms had undergone Verner’s Law, which changed the *s to *z when the stress followed and then intervocalic /z/ changed to /r/ by rhotacism.

A somewhat more complicated but more informative example is seen in the table given below. Latin rhotacism and the interaction of analogy with sound change.

Stage: 1- Latin before 400 BC

honos	‘honour’	
labos	‘labour’	Nominative singular
hono:sem	labo:sem	Accusative singular

Stage: 2 – rhotacism: s>r / v-v

honos	
labos	Nominative singular
hono:rem	
labo:rem	Accusative singular
hono:ris	
labo:ris	genitive singular

Stage: 3- after 200 BC analogical reformation of nominative singular

honor	labor	Nominative singular
hono:rem	labo:rem	Accusative singular
hono:ris	labo:ris	Genitive singular

In the above example, the regular sound change in stage -2 rhotacism (s > r / v-v) created allomorphy (honos / honor-) i.e. irregularity in the paradigm. Later irregular analogy changed honos and labos to honor and labor. Both were now ending in r, matching the r of the rest of the forms in the paradigm. Thus irregular analogy has regularized the form of the root, eliminating the allomorphic alternations involving the final consonant of the root.

Generative grammar and various types of language changes

Linguistic changes are classified into phonological, morphological, syntactic and semantic. Traditionally these changes are described in terms of particular structure and events, for instance, k > c or p > f; plural kaL > al etc. Historical linguistics is more interested in collection and classification of linguistic changes. But sharp contrast with this traditional approach, generative grammar theory will always approach such linguistic changes from the point of underlying rule change. According to the views of generative grammar, linguistic change means rule change indirectly. So comparison of two languages or two stages of one language means comparison of two grammars in terms of their respective rules. Therefore there

is a marked difference between generative and traditional approaches in dealing with linguistic changes.

Linguistic changes

Let us discuss now the description and interpretation of language change within the framework of transformational generative theory. We shall concentrate on two basic questions viz., the representation of phonological change within the framework of a non independent phonology i.e, a phonology whose rules take account of structure at the grammatical level and the representation of syntactic change in terms of deep structure and transformational rules. Transformational generative grammar is concerned primarily not with the output of grammar but with the rules which produce the output. It follows that if the language changes this is the result of change in the grammar not as such a new concept, of course, but a point worth stressing in view of structuralist preoccupation with forms rather than rules. Let us discuss change in the phonological and syntactic component in turn.

Phonological change

Let us now examine the way in which transformational generative theory deals with phonological change. The transformational generative theory assumes that phonological change may be as traced from a comparison of the successive synchronic grammars of a language. In a linguistic model which operates with a set of underlying representations and ordered series of phonological rules generating the appropriate surface forms from these successive synchronic grammars. It may potentially differ from one another in their rule inventories in the forms and ordering of their rules and in the forms of the underlying representations of corresponding lexical items. Generative theory dictates that any linguistic changes must be interpreted and described through rule changes. As such rules are classified in to four types which would account for what are described in terms of sound changes by the traditional historical linguistics. The primary sound changes are described in terms of four rules viz., (1) Rule addition (2) Rule loss (3) Rule reordering and (4) Rule simplification.

Rule Addition

Rule addition is a type of rule change in which a new rule is introduced into the grammar. So that it applies after all previously existing rules is the most familiar type of rule change. This kind of rule accounts for such sound changes like terminal devoicing in modern German. Descriptive evidence in contemporary German is the inclusion of their innovative rule in the grammar of German. The famous devoicing phenomenon in the Middle High German is a good example to illustrate the rule addition.

[+stop] > [-voiced] / - #

This rule which changes [+voiced] into [-voiced] in the case of stops occurring word finally had the effects of creating from /veg / 'way', /tag / 'day' alternatives like [vek] ~ [veg-], [tak] ~ [tag]. This is a rule addition because there are evidences for the absence of this kind

of sound change in the earlier stages of Germanic language. There are following some other evidences in support of this view.

[Velar stop] > [Labio-dental] / - # -

In standard Tamil the words /makan/ 'son' changed into /magan/ and /pukai/ 'smoke' changed into /pogai/ in the spoken form of the language is common. But in some words the velar /k/ changed into labio-dental /v/ is added to sound change which is not common. For instance nukattaDi > novattaDi 'yoke', nakam > nevam 'nail', pa:kaRka:y > pa:vakka:y 'bitter guard'.

Rule Loss

Rule loss is a type of rule change in which a rule that was formerly present disappears later from the language. For instance all the southern varieties of German historically acquired a rule devoicing final obstruent. As illustration, the rule which devoices word-final voiced stops in standard German has no counterpart in certain present day varieties of Yiddish and words such as tag 'day' lid 'song' have voiced stops throughout their paradigms. However unlike the noun /veg/ 'way' which belongs to their class, the etymologically related adverb [a'vek] 'away' has undergone devoicing. The devoicing rule must therefore have been present in earlier grammars of Yiddish dialects have been subsequently lost. Here linguist can explain the adverb has a voiceless stop whereas the noun from which it derives for the voiced one.

Rule reordering

The state of affairs in which several independent changes expressible as phonological rules and it is applying at different times in the history of a language interacts. So as to produce the complex results the change can be understood by placing those changes in a particular historical order. Depending upon the chronological order of sound changes which may differ from language to language, rules may be explained in different order. Thus language A contains in its grammar rules X and Y, which must be applied in the order X first and Y second. Language B contains the same two rules but in the opposite order: Y first and X second.

According to Newman (1996) Middle Chinese *a is regularly lost in Cantonese by nucleus deletion when preceded by an on glide and followed by another segment as in *miu > miu 'ticket'. If there is no on glide, *a undergoes lengthening to a: as in *nai > na:i 'ticket'. However, words in *-ia appear to be exceptions to nucleus deletion, since they yield as in *niai > nai 'mud'. Invoking an earlier rule of palatal dissimilation, by which *i is lost in the configuration *-iai would account for the non-application of nucleus deletion, but would also make the result subject to lengthening, wrongly predicting *na:i 'mud'. The solution is to order the rules as follows: 1) lengthening, 2) palatal dissimilation, and 3) nucleus deletion. This gives the right result in all cases and presumably reflects the relative chronology of the changes.

Rule simplification

Rule simplification is the lost type of rule change in which by the elimination of some of the conditions on its process. A rule comes to apply to a wider range of cases than formerly. For example all varieties of English have lost word-final /b/ after nasal as in climb and lamb; most varieties have generalized this to the loss of word-final non-coronal voiced plosives after

nasal, as in sing and long and some southern and black varieties of American English have generalized it. Further to the loss of all word-final voiced plosives including /d/ after nasals in /stæn/ for stand and /main/ for mind.

Rule simplification is more generality in its application and less amount of complexities. Let us illustrate through the rules of devoicing and umlauting.

1. [+st] → [-voiced] --# [All stops affected]
2. $\left[\begin{array}{l} +st \\ +cont \end{array} \right] \rightarrow [-voiced] --#$ [only fricatives affected]

The first rule is simpler than the second rule. The first rule has a feature count of three while the second has a feature count of four. The first rule is also more general of the two since it applies to the natural class of all stops and fricatives while the second rule applies to the natural class of fricatives. We will now turn to the umlaut rules when the umlaut rule was first added it applied only to the back rounded vowels in Middle High German:

U → ü; o → ö

Later on the vowel /a/ was also affected thus a → ä. It means the earlier umlaut rule was simplified later by becoming more general in that it is applied to all back vowels being umlauted. It is in these forms of rules that generative grammar describes and treats various types of linguistic changes. Changes on other levels like morphology and syntax are also similarly dealt with in generative grammars.

Reconstruction

Historical linguists normally use the comparative method to reconstruct a protolanguage. By examining several related languages for cognates, words bearing a similarity due to their common descent. Linguists can postulate the original forms from which the cognates arose. This method uses lexical terms such as pronouns, kinship terms, body parts and lower numbers, which are terms most resistant to change. These methods have provided significant insight into the genetic links between different languages. It is important to remember that similarities can also arise from borrowing, linguistic universals and or by chance. The main aim of comparative linguistics is the comparative study of languages. Comparative study of languages means comparison of languages. Comparison of languages means comparison of the linguistic structures of given languages. Linguistic structures are phonemes, morphemes, lexicons, sentences etc. So the languages can be compared phonemically, morphemically, lexically, syntactically and semantically.

Cognate collection

A requirement to the comparative study of languages is the identification of words which show partial similarity both in sound and meaning. The words similar both in sound and meaning than this similarity is called double similarity.

Example:

Language – A	Language - B	
ka:l	ka:l	‘leg’
va:l	va:l	‘tail’
mayil	mayil	‘peacock’

Here the languages A and B have words which show similarity both in sound and meaning. This double similarity may be either due to (i) accident or (ii) borrowing or (iii) direct inheritance.

Accidental similarity

Greek	Malay	
mati	mati	‘eye’
Tamil	English	
eTTu	eight	
kol	kill	
kuTi	hut	
tin	dine	
kaTavuL	God	
onRu	one	
ella:m	all	
peccu	speech	
ve:NTu	want	
vant	went	

Borrowing

English	Tamil	
pen	pe:na:	
Sanskrit	Tamil / Malayalam	
ka rma	kar(u)mam	- action
a:sti	a:sti	- property

Direct inheritance [Genetic relationship]

Tamil	Malayalam	Kannada	Telugu	
mi:nu	mi:nu	mi:nu	mi:nu	'fish'
ni:r	ni:r	ni:r	ni:LLu	'water'
ponnu	ponnu	ponnu	ponnu	'gold'
kiLi	kiLi			'parrot'
vi:Du	vi:Du			'house'
na:Du	na:Du			'country'

For comparative study of languages, words showing double similarity due to common inheritance should be preferred and used. Therefore, double similarity due to accident and borrowing must be identified and eliminated. Only double similarity due to direct inheritance, i.e. words inherited directly from the parent language must be established. How to establish direct inheritance? In order to test it the linguists have to verify the given set of words showing double similarity due to direct inheritance or not. We must look for "Recurring correspondence" for the sounds / phoneme of those words. What is meant by correspondence? What is meant by recurring correspondence?

Correspondence

Tamil	Malayalam	Kannada	Telugu	
mi:nu	mi:nu	mi:nu	mi:nu	'fish'
m-	m-	m-	m-	

/ m / occurs initially in all the related words. Such occurrence of a linguistic unit / structure in the same environment in the majority of related words is called **correspondence**. So now m- m- m- m- is one correspondence. Similarly -i- -i- -i- -i-; -n- -n- -n- -n- and -u -u -u are all other correspondences. Now we have four sets of successive correspondences. After this we must be whether or not these four sets of successive correspondences occur in other sets of words also, i.e. whether or not these correspondences are recurring in other sets of related words also. For instance,

Tamil	Malayalam	Kannada	Telugu	
ma:Ru	ma:Ruka	ma:Ru	ma:Ru	-to change
ni:r	ni:r	ni:r	ni:ru	-water
tin(nu)	tinnuka	tinnu	tinu	-to eat'

ku:Du ku:Duka ku:D(u) ku:Du -to meet'

Recurring correspondence

Recurring correspondence is explain with help of the about mentioned example mi:nu (fish). The correspondence for this word is in the following manner.

Correspondence:-

- 1) m- m- m- m-
- 2) -i- -i- -i- -i-
- 3) -n- -n- -n- -n-
- 4) -u -u -u -u

Recurring correspondence:

1)	m-	m-	m-	m-	
Tamil	Malayalam	Kannada	Telugu		
ma:Ru	ma:Ruka	maRu	ma:Ru	-to change'	
mu: <u>n</u> Ru	mu: <u>nnu</u>	mu:Ru	mu:Du	-three'	
me:y	me:yuka	me:y	me:yu	-to graze'	
2)	-i-	-i-	-i-	-i-	
i:nu	i:nuka	i:n	i:nu	-to give birth'	
ni:r	ni:r	ni:r	ni:ru	-water'	
3)	-n-	-n-	-n-	-n-	
ni:r	ni:r	ni:r	ni:ru	-water'	
i:n	i:nuka	i:n	i:n	-to give birth'	
4)	-u	-u	-u	-u	
uL	uL	uL	uNDu	-to be'	
cikku	cikkuka	cikku	cikku	-to be caught'	

We see thus, that the successive correspondences in the words **mi:nu** 'fish' all recur in many other sets of words. The great regularity with which the correspondences recur makes in very unlikely that the double similarity in the set of words (mi:nu) are due to chance, accident or borrowing. Such words which show double similarity and recurring correspondences also

are called cognates. The dictionary which consists of cognates is called etymological dictionary. Etymon (sg.) roots, Etyma (pl) meaning roots. Etymology means the study of roots and their developments through the history of words.

The words mi:nu, mi:nu, mi:nu, mi:nu look similar not because of accident or borrowing, but because the languages having these words are related. This relationship among languages is more and more confirmed if more and more sets of cognates are identified. Thus identification of cognates leads to discovery of relationship among languages. Now the question of parent language arises. Here the assumption is that if some languages are proved to be related they are assumed to be descendants of a parent language called proto language.

Tamil Malayalam Kannada Telugu

mi:nu	mi:nu	mi:nu	mi:nu	-fish
-------	-------	-------	-------	-------

Here all the languages show the same word with the same meaning. One of the assumptions of comparative linguistics is that any linguistic feature which is found in the majority of languages is assumed to have existed in the parent language also. Here mi:nu is found in all four languages and so it must have existed in the parent language also. Now a question may be raised that if all the daughter languages and the parent language show only similar linguistic structures, what is the necessity of talking a parent language and daughter language? Going a step further if all the daughter languages or the sister languages show one and the same linguistic features there are not many languages at all. There is only one language. It means comparative Dravidian has no business at all. Comparative Dravidian will get a business only when there are linguistic similarities found between a given groups of languages.

The most important assumption of comparative Dravidian is that any linguistic features/ elements which are found in the majority of languages are assumed to have existed in the parent language also. As per this assumption the following words are all belong to the parent language.

Tamil Malayalam Kannada Telugu

ciRu	ciRu	kiRu	ciRu	-small
keDu	keDu	keDu	ceDu	-spoil
ki:rai	cira	ki:re	ku:ra	-greens

Here we find both similarities as well as dissimilarities between the related languages. Now the problem of identifying the parent language arises.

Tamil Malayalam Kannada Telugu

keDu	keDu	keDu	ceDu	-spoil
kay	kay	kay	ce:yi	-hand

In the above sets of cognates, Tamil, Malayalam, Kannada have all the same word whereas Telugu has different words. Now we must decide to which of them belongs to the parent language. Recall the above-mentioned concept i.e. any feature found in the majority of the related languages should be assumed to have existed in their parent language also. According to this keDu ‘spoil’ has to assume to have existed in the parent language. Similarly kay ‘hand’ may be reconstructed or assumed to have existed in the parent language. Reconstruction does not stop with the reconstruction of proto forms alone. Reconstruction procedure should also explain the exceptional forms which differ from the reconstructed proto form.

Phonological Reconstruction

The method through which we can acquire knowledge about ancestral languages is reconstruction. It is an approach in which we apply our knowledge of linguistic change to reverse linguistic history. We can do or by looking for synchronic evidence, which points to earlier linguistic change. Such evidence will consist of phonetic variation between forms which can be assumed to originally have had invariant structure. In other words, reconstruction attempts to earliest invariance. Usually the first reconstruction work that is done with any language family is an attempt to reconstruct the sound system and a collection of phones of the parent language. The process is called phonological reconstruction, through its implication for other parts of the grammar are obvious.

The initial step in this process is to group the words together from languages which have suspected or we know to have the same origin. The words may quite have same meaning in the attested languages. But we must be aware of possible semantic shifts which may cause different meaning. These sets of words are called ‘**COGNATES**’. A word in one language is cognate to a word in another language. Both have the same ancestor and neither is the result or any borrowing. In the regard we can assume that the English father and Latin *pater* are cognates, derived from Indo- European. As an example of reconstruction, we will examine words from four Polynesian languages.

MAORI	HAWAIIAN	SAMOAN	FIJIAN	MEANING
pou	Pou	pou	Bou	‘post’
tāpu	kapu	tapu	tabu	‘forbidden’
tāŋi	kani	taŋi	taŋi	cry
takere	ka?ele	ta?ele	takele	keel
hono	hono	fono	vono	site
marama	malama	malama	malama	moon
kaho	ŋaho	ŋaso	kaso	thatch

Let us see the changes in vowels:

<u>Maori</u>	Hawaiian	<u>Samoan</u>	<u>Fijian</u>	
o	=	o	=	o = o [1]
u	=	u	=	u = u [1,2]
a	=	a	=	a = a [2,3,4]
i	=	i	=	i = i [3]
e	=	e	=	e = e [4]

In the case of the vowels given above, there is no question, since the same vowels appear in all the different languages. So there are five vowel phonemes are there in the above mentioned languages.

CONSONANTS:

<u>Maori</u>	<u>Hawaiian</u>	<u>Samoan</u>	<u>Fijian</u>	
p	p	p	b	- [1,2]
t	k	t	t	- [2,3,4]
ŋ	n	ŋ	ŋ	- 3
k	?	?	k	- 4
r	l	l	l	- 5
h	h	f	v	- 6
n	n	n	n	- 7
m	m	m	m	- 8
n	h	s	s	- 9

1. A change in one direction as opposed another as known by experience with a wide number of languages.
2. Consideration of the parent language will look like, our first correspond set one we assumed proto phoneme |p| this involves, and a sound rule which says that proto Polynesian |p| becomes |b| in Fijian.

STEP 2

1. p → b voice less → voiced
2. ŋ → n velar nasal → bilabial nasal
3. t → k palatalize → velar stop
4. k → ? velar → vd velar
5. l → r alveolar → palatial
6. f → v,b vIF → vd F
7. n → n dental – dental
8. m → m - bilabial nasal – bilabial nasal
9. s → h - palatial F- velar F

CONSONANTS

Hawaiian

k → ?

t → k

ŋ → n

$\left\{ \begin{array}{l} f \\ s \end{array} \right\} \rightarrow h$

Samoan – k → ?

Maori

l → r

f → $\left\{ \begin{array}{l} h \\ S \end{array} \right\}$

Fijian – [Voiceless labials] → [Voiced labials]

List of the words we used are transcribed in their proto form as follows.

1. pou
2. tapu
3. taji
4. takele
5. fono

6. malama

7. kaso

The above rules are necessary to account for the reflexes of these words in the daughter languages.

Internal Reconstruction

Construction is the manner of grouping and combining elements of speech. The reconstruction of diachronic phonological rules on the sole evidence of synchronic morphological alternation is known as internal reconstruction. It describes about a single language or single stage of its developments. Basically, it involves comparison of form gives any indication of earlier stage of the language.

Fundamental Assumption

1. Finding the traces in the history of language one can draw inferences about the earlier incidents of the language.
2. Phylogenetic change is a key importance in their connections.
3. Morphophonemic alternations will reflect an earlier regularity disrupted by phonemic reconstruction.
4. By careful examination of morphophonemic irregularities and the distributional aspects of phonological systems will yield reasonable deduction about its earlier history.

The procedure by which morphophonemic irregularities can be utilized for internal reconstruction. The morphophonemic irregularities can be explained with following change.

avaL → ava	Rule /L/ → ∅ / #
ka:r → kar	/v:/ → v / #

A development occurring in early Greek and Indic first pointed out by Hermann Grassmann (1863). With considerable oversimplification, it may be stated as follows: If two aspirated consonants occur in the same syllable, or in consecutive syllables, then the first loses its aspiration. Hermann Grassmann who was a Indo-Europeanist of the 19th century discovered Internal Reconstruction. Indo-European correspondence sets were causing difficulty in that they seemed not to fit in which previously established pattern. Grassmann examined the data in Greek and Sanskrit and was able to show the developments peculiar to those languages resulted in the apparent irregularities.

Example:

Greek:

1.a. trekh – o ‘I walk’ 1.b. threk-s-o ‘I will walk’

2.a. thrik-s ‘hair’ 2.b. trikh –os ‘of the hair’

Each of the above forms shows that only one aspirated consonant occurs, but this varies in each of the two sets.

1.a and 2.b we have **kh**

1.b and 2a we have **th**

In other words in these paradigms we find aspirated consonants (**th, kh**) alternating with non aspirated (**t, k**) which gives us different forms of the stem.

trekh, threk
thrik, trikh

But the pre Greek the original stems were like this:

threkh
thrikh

So the Greek underwent two sound changes. i.e., Removed the aspiration from the consonant when it was followed by – s; and the deaspirated a consonant when it was followed by another aspirated consonant in the same stem.

(a) → Ch c / - s
(b) → Ch c / - ...ch

The rule that two aspirated consonants may not occur in the stem also applied in Sanskrit and came to be known as **Grassmann's law**.

Relative chronology of changes

Of the several changes that a language might suffer, it may be possible sometimes to order these changes. We had seen earlier how old English hnecca [hneka] became modern English neck [nek]. This process involves two sound changes: The loss of the initial consonant [h] and the loss of the final vowel [a]

Given only these two forms, it is difficult to state which change took place first or whether there was any intermediate stage of development. In Middle English texts we find the same word spelled nekke [nekə] we can therefore see that between old and Middle English two sound changes had taken place.

(i) h → ø/#-n
(ii) a → ə /-#

And for the development from the middle to the New English period, we can posit a third rule which drops the last syllable.

$$(iii) \quad a \rightarrow \emptyset / - \#$$

The phonological shape of the Middle English form is concrete evidence that the initial consonant was dropped before the final vowel. It also tells us that there was an intermediate stage in the loss of the final vowel, i.e., there was not a simple $a \rightarrow \emptyset$; but rather a process such as, $a \rightarrow \text{ə} \rightarrow \emptyset$.

We may not always be fortunate in having textual evidence. Still, there may be some situation, which allows stating relative chronology of changes. Such a situation is when the conditioning environment itself changing after a conditioned change has taken place. Otherwise, we may find that a particular change much have taken place in an environment produced by another rule. By deferring what sounds constitute the environments necessary for specific sound changes, we can often tell which changes occurred before other chronological order. For instance take the following pre-Sanskrit forms and their descendants in the attested language.

Pre. Skt.	skt.	
koti	kati	‘how much’
ke	ca	‘and’

We can note the following changes. The vowels *o and *e both become [a], *k becomes [c] before *e, but not before *o. We see that require the vowel [e] as opposed to [a] in order to describe the conditioners environment of the consonant sound change. Therefore, we must assume that the changes took place before the vowel *e lost its distinction from *o. our rules must be ordered to reflect this relative timing. Thus, the two rules are in the order.

1. $k \rightarrow c / - e$
2. $\left. \begin{array}{c} e \\ o \end{array} \right\} \rightarrow a$

Language Contact

Language contact occurs when two or more languages or varieties interact. The study of language contact is called contact linguistics. Language contact can occur at language borders between adstratum languages or as the result of migration with an interfering language acting as either a superstratum or a substratum. Language contact occurs in a variety of phenomena, including language convergence and borrowing. The most common products are pidgins, creoles, code-switching and mixed languages.

Borrowing

As the lexicon of a language is an open class it is constantly expanding. The direct goal is gaining words for new phenomena, concepts, etc. in the society which uses the language in question. The side-effect is an increase in the size of the lexicon. There are various means of extending a language's word stock which can be broken down into two basic groups. The first creates compounds out of material from the language itself and the second resorts to borrowing material, integrating it into the system (phonology, morphology, semantics) of the language as it does so.

The most common way that languages influence each other is the exchange of words. That is the transfer of a word from one language into a second language as a result of some kind of contact between speakers of the two languages. Much is made about the contemporary borrowing of English words into other languages, but this phenomenon is not new, nor is it very large by historical standards. The large-scale introduction of words from Latin, French and other languages into English in the 16th and 17th centuries was more significant. There are internal and external evidences are there for Tamil contacting with some languages like Persian, Arabic, Portuguese, French and Dutch etc. First instance words are enter into other languages in terms of interference. After some time the words are considered as borrowing items from the source language. For instance, the languages like Tamil and English contacted each other number of words borrowed from Tamil to English and English to Tamil. T. P. Meenakshi sundaram (1965) listed some words borrowed from Tamil to English during 18th and 19th centuries. They are:

Tamil	English
aNaikkaTTu	anicut - dam
milakuracam	mulligatawny –pepper soup
cakkai	jack - jack fruit
ma:ñka:y	mango - mango
puNNa:kku	poonac - oil cake

Broadly the transfer of linguistic features of any kind borrowed from one language to another as a result of contact. This is a very common process which is attested for all periods of the history of English or any other language for that matter. The reasons for borrowing are basically twofold. On the one hand there may be a necessity or need for a foreign word, to fill a gap so to speak. This is the case with many scientific terms are used in Modern Tamil which are not available at the time of invention. For example, the words like dish, satellite, mobile etc. The adjectival formations in the Early Modern English period which were coined on the basis of classical stems and which provided a form not available in English at the time or not appropriate, e.g. *marine* as an adjective to *sea*; *pedestrian* to *walk*: *walker*; *aquatic* to *water*, etc.

The second reason for borrowing is because of the relative prestige (social standing) of the speakers using the donor language. This was the case with many English and other language loans often used by the native speaker of Tamilnadu even today. However, loans made for this reason will only survive in the language if there is a semantic justification for them, i.e. if the loanword is separate from the corresponding native word in some aspect of its meaning. For instance no one used *co:Ru* 'rice' when they are going to hotel. They used to ask *rice / ca:tam* instead of *co:ru* 'rice'. These words are considered as high level words which are fit for their communication to maintain the status.

The treatment of loans in a language depends on the structure of the lexicon in the borrowing language. For instance in English Bakery means 'a place where bread, biscuit etc., baked to sold' whereas in Tamil *aTumanai* which is not clear to the speaker of Modern Tamil. One importance of this situation is that English tends to be easy in its handling of loans from the language. Similarly a word waterfall is a normal loan from English but the word *aruvi* 'water falls' is a original Tamil form which does not match in English. But the original loan as water + falls represent in Tamil *ni:r + vi:lcci* which means water falls to replace the word *aruvi* 'water falls.'

Pidgin and Creole

Pidgin language is a simplified version of a language that develops as a means of communication between two or more groups that do not have a language in common. it is most commonly employed in situations such as trade, or where both groups speak languages different from the language of the country in which they inhabit. Fundamentally, a pidgin is a simplified means of linguistic communication, as it is constructed unplanned or by convention between individuals or groups of people. a pidgin is not the native language of any speech community, but is instead learned as a second language. a pidgin may be built from words, sounds, or body language from multiple other languages and cultures. They allow people who have no common language to communicate with each other. Pidgins usually have low prestige with respect to other languages.

Creole is a stable natural language that has developed from a pidgin, i.e. a pidgin language becomes the mother tongue of a population. When that happens, it is called a creole language. Creoles differ from pidgins because creoles have been nativized by children as their primary language. The result that they have features of natural languages that are normally missing from pidgins, which are not anyone's first language. The vocabulary of a creole language is largely supplied by the parent languages, particularly that of the most dominant group in the social context. While the construction of creole, there are often clear phonetic and semantic shifts. On the other hand, the grammar often has original features that may differ substantially from those of the parent languages.

Dialectal and sub-cultural change

Language and speech is not the same thing. Speech is a broad term simply referring to patterned verbal behavior. In contrast, a language is a set of rules for generating speech. The increasing assurance about the reliability of sound change led to a great attention in the study

of various level of speech, especially geographical dialects. In spite of the clarifications produced by Hermann Grassmann and Karl Verner, some elements in the standard languages under examination still showed irregularities. It assumed as a result that standard languages enclosed irregularities because they are mixed. To get pure language they thought that collecting data from speech of the everyday people commonly known as dialect is helpful for historical linguistic study. A dialect is a variant of a language. If it is associated with a geographically isolated speech community, it is referred to as a regional dialect. However, if it is spoken by a speech community that is merely socially isolated, it is called a social dialect. These latter dialects are mostly based on class, ethnicity, gender, age, and particular social situations. Dialects may be both regional and social. The term class dialect is also found for clearly defined varieties distinguishable along class lines. A distinction is sometimes made between rural dialects, spoken in rural areas and urban dialects spoken in cities. The rural dialects often showing maintenance of older dialect structures and urban dialects often characterized by dialect convergence and mixing as speakers of different dialects come to contact with one another.

Semantic change

In semantics and historical linguistics, any change in the meaning(s) of a word over the course of time. Common types of semantic change include amelioration, pejoration, broadening, semantic narrowing and metonymy. Semantic change may also occur when native speakers of another language adopt English expressions and apply them to activities or conditions in their own social and cultural environment.

Amelioration

Amelioration is a type of semantic change that occurs when a word's meaning get better or becomes more positive over time. The adjective "pretty" comes from the Old English "prættig" which meant "cunning, artful, and wily." In the change from Old English to Middle English, "pretty" had come to mean "manly, gallant," and then it shifted to "attractive, skilfully made," and then to "fine," and then by the mid-fifteenth century came to mean "beautiful in a slight way."

The word "nice" came into English in the 12th century from the Old French "nice" which meant "careless, clumsy; weak; poor, needy; simple, stupid, silly, foolish" and by the late 13th century it had come to mean "foolish, stupid, and senseless." With regard to semantic change, it then came to mean "timid" and then by the late 14th century it meant "fussy, fastidious" and by 1400 it meant "dainty, delicate." In the 1500s, "nice" came to mean "accurate, careful" and by 1700 it meant "agreeable, delightful" and by 1800 it had come to mean "kind, thoughtful." At the word nice this means 'pleasant, good natured and agreeable etc.

A more recent change can be seen in "geek" which in 1900s was listed as United States carnival and circus slang meaning "sideshow freak." The word appears to be a variant of "geck" which in the 1500s meant "a fool, dupe, and simpleton." By 1983, "geek" was used as a slang term referring to students who lacked social graces but were obsessed with computers and new

technology. In the twenty-first century, “geek” has expanded to refer to someone with special knowledge, not limited to computers.

Similarly in Tamil the word “*kalakam*” means gambling house in old Tamil. But in modern Tamil it refers to any statutory bodies.

Pejoration

Pejoration is another type of semantic change that occurs when a word’s meaning becomes more negative. In English the word notorious initially meant “publicly known”. It came from Medieval Latin “notorius” which meant “well-known, commonly known”. During the 17th century it began to acquire a negative connotation from its frequent association with derogatory nouns.

The word “hussy” in the 1500s simply meant “mistress of a household, housewife” and was an alteration of the Middle English “husewif.” Over time, the meaning broadened to mean “any woman or girl” and it was being used in a more negative manner to refer to any woman or girl who showed casual or improper behavior. By the late eighteenth century the word had a general derogatory sense.

The word “*kuppai*” in old Tamil was referred to a “heap of grain.” But in modern Tamil it refers to ‘garbage’. Similarly the words like “*na:RRam*” and “*te:vanaTiya:r*” changed its meaning when we compare with old Tamil. The word *na:RRam* in old Tamil simply meant ‘good smell’. But in modern Tamil it refers to ‘bad smell’. Another word *te:vanaTiya:r* in old Tamil any women who served in temple. In modern Tamil is being used in the sense of a woman of loose morals with sound change.

Another example of pejorative semantic change can be seen in “wench.” In the twelfth century, “wenchel” meant child and by the thirteenth century it had become “wenche” meaning “girl or young woman.” “Wench” then came to refer to “servant, particularly servant girls” and by the mid-fourteenth century being used in the sense of a “woman of loose morals, mistress.”

Broadening

Semantic change is the development of word usage and change in one of the meanings of a word. When features are dropped, this is called widening. Widening may result in either more homonymy or in more polysemy. Semantic widening broadens the meaning of a word. This process is called “generalization.” Some issues that affect semantic widening are linguistic factors, psychological factors, and sociocultural factors. For example in old Tamil the word *eNNai* used to refer *eL eNNai* [*nalleNNai*] ‘gingily oil’ only. But in modern Tamil it extended to refer any oil like, *kaTalai eNNani* ‘groundnut oil’, *te:nka:y eNNai* ‘coconut oil’, *su:ryaka:nti eNNai* ‘sun flower oil’ *maNNeNNai* ‘kerosine’ etc.

In English the word *guy* meant the leader of the plot to blow up the English houses of Parliament. At last it came to mean “a person of fantastic appearance.” Over time, the word came to mean “any man or any boy.” The word *business* that meant “busy, careworn, or anxious,” and now it means “a corporation or occupation.” Similarly the word *cool*, the

colloquial speech word that meant “a jargon of jazz musicians.” Over time, it began to mean a general word to describe something preferable.

Specialization of meaning

A word which was used to refer general sense relating to number of objects, get restriction over the period of time and refer to any one member of the objects is called specialization of meaning. In Tamil the word “pon” was used to refer to any metal in old Tamil i.e. *irumbu* ‘iron’ *pittaLai* ‘brass’ *cembu* ‘copper etc. But in modern Tamil it is used to refer only ‘gold’. The meaning of the word is get restriction when we compare two different periods of Tamil.

Metaphor

"Metaphor involves extensions in the meaning of a word that suggest a semantic similarity or connection between the new sense and the original one. Metaphor is considered a major factor in semantic change. The semantic change of *grasp* ‘seize’ to ‘understand,’ thus can be seen as such a leap across semantic domains, from the physical domain to the mental domain. In Tamil the word *ka:tu* is a part of the body which means ‘ear’ Due to metaphorical extension it is used in some other objects like *pai ka:tu* ‘bag handle’ *va:Nal ka:tu* ‘handle of big pan’ etc., Similarly the words like *ka:l* ‘leg’ *va:y* ‘mouth’ etc are also extended its level based on the similarity of sense with other objects.

Metonymy

Metonymy is another type of semantic change in which a word acquires the sense of something else with which its referent is commonly associated. That is the word acquires the sense on the basis of relation other than similarity. For instance the word Thirubuvanam in Tamil is name of the town in Tamilnadu and it is famous for production of silk saree. Normally we say ‘*kalya:Nattukku thirubuvanam eTutten:n*’. Literally it means “I took Thirubuvanam for my marriage”. But in this context due to metonymy the word Thirubuvanam refers to silk saree.

India as a Linguistic Area

The term linguistic area is used to denote a group of geographically contiguous languages characterized by a number of specific structural isoglosses. These shared areal features having been acquired as the result of contact and not inherited. The linguistic area can therefore be seen as the natural counterpart of the language family. This was certainly Trubetzkoy’s intention when he first proposed the concept in 1928, coining the term “*sprachbund*” or ‘Language association’ to describe it. Thus, while languages may be grouped in language families on the basis of shared inheritance from a common ancestor. They may also be grouped areally on the basis of shared features which they have acquired as a result of mutual contact. Linguistic areas have been postulated in various regions of the world and on the basis of variety criteria. In South-East Asia, for instance, Chinese, Tai, Vietnamese and number of other languages have in common the features of tone and this has been attributed to contacts.

India is the home of a very large number of languages. In fact, so many languages and dialects are spoken in India that it is often described as a ‘**museum of languages**’. The language

diversity is by all means mysterious. In popular idiom it is often described as 'linguistic pluralism'. But this may not be a correct description. The current situation in the country is not pluralistic but that of a continuum. One dialect merges into the other almost barely; one language replaces the other gradually. Moreover, along the line of contact between two languages, there is a zone of transition in which people are bilingual.

The give and take between the language groups has been very common, often resulting in systematic borrowings from one language to the other. According to the Linguistic Survey of India conducted by G. A. Grierson towards the end of the nineteenth century, there were 179 languages and as many as 544 dialects in the country. It may even be misleading in the sense that dialects and languages were enumerated separately, although they were taxonomically part of the same language. Out of the 179 languages as many as 116 were speech-forms of the Sino-Tibetan family, spoken by small tribal communities in the remote Himalayan and the north-eastern parts of the country.

Geographic Patterning of Languages:

The geographic patterning of languages in the South Asian subcontinent can perhaps be understood in the context of the space relations the region had with other parts of Asia. As already pointed out, the sub-continent marks a southward projection of the Asian landmass into the Indian Ocean. The overland connections with West and Central Asia, Tibet, China and other regions of Southeast Asia helped the process of penetration of linguistic influences into the South Asian region. This is evident from the fact that the languages spoken in the peripheral regions of South Asia, such as Baluchistan, Pak-Afghan borderlands, Kashmir and hilly parts of Himachal Pradesh and the regions in the Northeast have strong affinity with the languages spoken in the regions beyond the Hindu-Kush Himalayas. The remote Himalayan areas became the abode of Tibeto-Chinese languages. Similarly, the North-eastern region continued to receive influences from the neighbouring parts of Myanmar, Thailand and Indo-China. These regions are now the domain of the Tibeto-Burman languages.

The people in the plains of North India from Sind to Assam acquired different branches of the Indo-European family of languages. The peninsular region continued to retain the Dravidian speech forms even though the north was completely influenced over by the Indo-European languages. Between the Indo-European and the Dravidian one finds the Austric-speaking tribes nestled in the hills of the mid-Indian region. The linguistic heterogeneity of India can perhaps be brought to some order when one realizes that these speeches really belong to four language families Viz., Indo-Aryan, Dravidian, Tibeto-Burman and Austro-Asiatic. In the course of usage over millennia of years these language families have found for themselves niches in the Indian social space in different parts of the sub-continent.

According to Emeneau "Linguistic Area is an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of atleast one of the families. The quotation of Emeneau well applied to India. India is a museum of languages. These languages may be broadly classified into five major linguistic families namely Indo-Aryan, Dravidian, Tibeto-Burman, Munda and Andaman. The languages of these five major linguistic families show overtly diffusion of certain linguistic traits across their genetic boundaries. Due to geographical proximity for well over 3000 years linguistic

area where genetically unrelated languages possess certain common linguistic characterization which are not found in some of their genetically related languages. That is why India is called a linguistic area. Some of the common traits are described in the following paragraphs.

Vocabulary Borrowing

One of the common traits found among the Dravidian, Indo-Aryan and Munda languages is their mutual vocabulary borrowing. In the Dravidian languages have all borrowed many items at all periods from Sanskrit, Middle Indo-Aryan, and Modern Indo-Aryan. For example the following Tamil words are all said to be borrowed from Munda.

tavaLai - 'frog'
 iLani:r - 'tender coconut'
 a:L - 'person'

Similarly Sanskrit has also borrowed from both Dravidian and Munda. Dravidian words are found even in the earliest Sanskrit texts viz., Vedas. Burrow and others have identified so many Dravidian words in Vedic texts particularly the earliest text Rig Veda. For instance

Sanskrit		Dravidian	
e:la:	<	e:lam	'cardamomum'
palli	<	palli	'lizard'
mayu:ra	<	mayil	'peacock'
malaya	<	malai	'mountain'
ma:lai	<	ma:lai	'garland'

Munda		Sanskrit	
ondren	>	unduru	'rat'
jobo	>	jamba:la	'mud'
jim	>	jim	'to eat'
murid-sa:	>	marica	'pepper'

It has long been recognized that even the earliest Sanskrit texts show features which are historically non-Indo-European in their nature, but which resemble features of the Dravidian languages. In addition to that, as time went on more and more non Indo-European features appeared in Indo-Aryan. In late Jules Block collected all such non Indo-European features in his book 'Indo-Aryan'. Broadly speaking there emerge three general tendencies in connection with the common linguistic traits found in Dravidian, Indo-Aryan and Munda. They are as follows:-

1. Specifically, Indianization of Indo-Aryan.
2. Common trait in some contiguous languages of Dravidian, Indo-Aryan and Munda without the source of origin being known.
3. Common trait in some contiguous languages of Dravidian, Indo-Aryan and Munda with the source of origin being known.

Indianization of Indo-Aryan

Most of the languages of India belonging to all the three families Indo-Aryan Dravidian and Munda have a set of retroflex or cerebral consonants in contrast with dentals. The retroflexes include stops and nasals and also fricatives. Indo-Aryan, Munda and Dravidian languages form a practically solid bloc characterized by their phonological feature. This retroflex is really Dravidian in origin. As a matter of fact, retroflexes in contrast with dentals are Proto-Dravidian in origin but with regard to Munda and Indo-Aryan that is not the case. As for Munda there is one language so:ra: which does not have it. Retroflexes are unknown elsewhere in Indo-European field. From this it is well clear that this is a clear instance of Indianization of the Indo-European component in the Indian linguistic scene.

Common trait whose original source is unknown

In Marathi the palatals of Old-Indo-Aryan are represented by /tʃ/ and /dʒ/ affricate before front vowels but by /ts/ and /dz/ affricates before back vowels. A similar distribution is found in Southern Oriya, Telugu, Northern Kannada and Kurku a Munda language. These languages form a continuous band across central India and the trait must have no doubt originated in one language and spread to their languages from it; which was the originator is unknown now. Another Pan-Indian trait found in all three families of languages but origin of source is unknown in the presence of echo word construction. This is found in Dravidian like puli gili (tiger) and Indo-Aryan pa:ni va:ni (water) etc. The echo-morpheme in sora a Munda language is m- which can be found also in Brahui, Kolami, Parji, Telugu, Tamil, and Malayalam and also in various Indo-Aryan languages.

Common traits with origin of source known

The existence of what is called classifier or quantifier in all the three families can be cited as a good example to illustrate the above mentioned linguistic trait. The construction which has classifier is as follows:-

Numeral + classifier + noun

iraNT + a:m / a:vatu + paiyan - 'second boy'

Classifier morphemes are found in Indo-Aryan languages like Marathi, Eastern Indo-Aryan languages and Nepali also. Classificatory systems are found in some of the Dravidian

languages like Tamil, Malayalam and some other South Dravidian languages. This is found in North Dravidian languages also.

Though this is not an Indo-European feature, there are sufficient evidences to know that this linguistic feature might have originated in Indo-Aryan and spread to other languages. Here it must be mentioned in support of this view that some Indo-Aryan classifier morphemes are used in the languages involved and that these morphemes are used only with Indo-Aryan numerals in some of the non Indo-Aryan languages.

Historical Process of Language Diffusion

The history of Indian languages is not easy to reconstruct. As an over-view of the processes of peopling of India shows, Negroids were the first people to arrive. However, we do not exactly know about their language affiliation. The subsequent waves of migrations were so strong that the Negroids lost their identity completely, leaving behind little traces of either their racial or linguistic past.

The story of the four families of languages may be briefly recapitulated here, although it is not easy to establish the chronological sequence in which the speakers of the Austric, Tibeto-Burman and the Dravidian languages came to India. It is almost certain that these families were already there at the time of the advent of the Indo-Aryan. This is, however, an established fact that the Tibeto-Burman speech communities were Mongoloids racially. The original Tibeto-Burman, the parent of the early Chinese is supposed to have developed somewhere in western China around 400 B.C. It is also believed that the diffusion of this language eventually affected the regions lying to the south and the southwest of China-Tibet, Ladakh, northeastern India, Myanmar and Thailand. Perhaps, the Vedic Aryans were familiar with this group. They described the Tibeto-Burman-speaking Mongoloids of the Brahmaputra valley and the adjoining regions as Kiratas.

The speakers of the Kirata family of languages are distributed all along the Himalayan axis from Baluchistan and Ladakh to Arunachal Pradesh. They occupy the regions surrounding the Brahmaputra valley in the northeast from Nagaland to Tripura and Meghalaya. There are striking differences between the languages of the Kirata family distributed over such a vast geographical area. The speakers of the Tibeto-Himalayan branch of the Kirata languages occupy the Himalayan regions from Baltistan to Sikkim and beyond to Arunachal Pradesh.

The Bhotia group consists of the Balti, Ladakhi, Lahauli, Sherpa and the Sikkim Bhotia dialects. Linguists also identify a Himalayan group consisting of Lahuli of Chamba, Kanauri and Lepcha which is noticeable on the basis of certain linguistic behavior. In the east there is a North-Assam branch including the dialects of Arunachal Pradesh, such as Miri and Mishing. In other parts of the northeast the languages belong to the Assam-Burmese branch and are divided into Bodo, Naga, Kachin and Kuki-Chin groups. The speakers of the Kirata languages came to India in different streams at different points of time.

Similarly, the Kachin and the Kuki-Chin groups followed separate routes of migration. This is why there is a vast variety of dialects within the Kirata family and the roots of linguistic

heterogeneity go far beyond the Indian borders into the neighbouring parts of Tibet, Myanmar and Indo-China. Anthropologists as well as linguists believe that the Austric-speaking groups came earlier than the Dravidian speaking communities. The Austric speech communities were already there in the mid-Indian region before the arrival of the Dravidian. The present geography of the Austric dialect groups holds some clues to the historical processes of their diffusion into India.

Generally, the Austric family of languages is recognized as consisting of a Mon-Khmer and a Munda branch. The Mon-Khmer speakers belong to two separate groups, viz., Khasi and the Nicobarese, both separated by a distance of more than 1,500 kilometers which spans over an expanse of the Bay of Bengal. There is no clarity among the scholars about the routes taken by the speakers of the Mon-Khmer dialects. The Khasi speakers themselves are surrounded by other Kirata and Aryan dialects in the Meghalayan plateau.

The arrival of the Dravidian in India is generally associated with a branch of the Mediterranean racial stock which was already there in India before the rise of the Indus valley civilization. In fact, archaeologists believe that they were the builders of the Harappan civilization along with the Proto-Austroloids. The Dravidian speech communities were found over most of the northern and the northwestern region of India before the advent of the Indo-Aryan. However, following the rise of the Indo-Aryan in northwestern India, a linguistic change came and the Dravidian speaking area shrank in its geographical extent.

The present distributions of the Dravidian dialects in different parts of North India, such as Baluchistan, Chotanagpur plateau and eastern Madhya Pradesh, where Baruhi, Kurukh-Oraon and the Gondi are spoken respectively, suggest the earlier stage of distribution of this family of languages. In fact, Gondi is spoken in many parts of Central India from Madhya Pradesh and Maharashtra to Orissa and Andhra Pradesh. Dravidian speech forms were in use for many centuries in the pre-Christian era. The literary development in the Dravidian speech community could take place only in the first few centuries after Christ. It is believed that the old Tamil, old Kannada and the old Telugu had already come into being by 1000 A.D. Malayalam acquired its form a little later. With the Vedic Sanskrit, a branch of the Indo-European, the Indo-Aryan established itself in northwestern India. It had definite relations with the different Indo-European languages, such as Persian, Armenian, Greek, French, Spanish, German and English. An early form of Indo-European seems to have genetic relations with the Hittite speech of Asia Minor.

The form of Indo-European which was spoken in India came to be known as Indo-Aryan. Its arrival in India is seen with the rise of the Vedic Sanskrit. However, the old Sanskrit changed into Prakrit and several speech forms developed in different parts of northern and western India. The region lying between Saraswati and Ganga, encompassing the upper Ganga-Yamuna doab and adjoining parts of Haryana, to the west of the Yamuna, became the stage for the transformation of classical Sanskrit into a Prakrit form. From this early stage of development of Prakrit came the different Indo-Aryan vernaculars which are now spoken in north-western, north-central, central and eastern parts of India.

The Suraseni emerged in the core region of the midland as the popular language. Its core area extended over western Uttar Pradesh and the adjoining parts of Haryana. A developed form of this parent language is described by the linguists as Western Hindi. Around the core region of Suraseni other speech forms developed on the west, south and east. These languages formed an outer band around the core language. On the west and the northwest lay the Punjabi and the Pahari dialects. Rajasthani and Gujarati emerged on the southwest. On the east a form of language, now known as Eastern Hindi, emerged in Kosali. Linguists believe that these outer dialects were all more closely related to each other than any one of them was to the language of the midland.

“In fact, at an early period of the linguistic history of India, there must have been two sets of Indo-Aryan dialects - one the language of the midland and the other the group of dialects forming the outer band.” This first stage was followed by a subsequent phase of expansion. As the population of the midland region increased expansion became a necessity. Thus, on the periphery of the languages of the outer band developed new speech forms which were by and large not related to the language of the midland. For example, while Punjabi was closely related to the language of the upper doab it got transformed into Lahnda in southwestern Punjab. This language had little relationship with the language of the midland. The midland language occupies the Ganga-Yamuna doab and the regions to its north and south. This core region is encircled by different speech forms in eastern Punjab, Rajasthan and Gujarat.

Further beyond in the west and the northwest Kashmiri, Sindhi, Lahnda and Kohistani is a band of outer languages. The languages of this band may be described as constituting the northwestern group of the outer languages. On the southern periphery lies the Marathi. In the intermediate band are situated languages, such as Awadhi, Bagheli and Chattisgarhi. On the eastern periphery lie the three dialects of Bihari, viz., Bhojpuri, Maithili and Maghadi. The Bihari is surrounded by Oriya in the southeast and Bengali in the east. The languages of the eastern branch of the Indo-Aryan extend further in the east where Assamese occupies the Brahmaputra valley.

Linguists believe that the development of the Indo-Aryan languages completed itself through several phases. The Prakrits developed into two stages: Primary Prakrits and Secondary Prakrits. The Primary Prakrits which were the first to evolve out of the classical Sanskrit were synthetic languages with a complicated grammar. In the course of time they ‘decayed’ into Secondary Prakrits. “Here we find the languages still synthetic, but diphthongs and harsh combinations are avoided, till in the latest developments we find a condition of almost absolute fluidity. Each language is becoming an ineffectual collection of vowels hanging for support on an occasional consonant. The last stage of development of the Prakrits is known as literary Apabrahmsa. It is supposed that the modern vernaculars are the direct children of these Apabrahmsas. The sequence of change was like this. The Suraseni Apabrahmsa was the parent of Western Hindi and Punjabi. Closely connected with it were Avanti, the parent of Rajasthani, and Gaurjari the parent of Gujarati. The other intermediate language Kosali (Eastern Hindi) sprang from Ardha-Magadhi Apabrahmsa.

Reference

- Anttila, R. 1972 An introduction to Historical and Comparative Linguistics. New York.
- Arlotto, A. 1972 Introduction to Historical Linguistics. Boston: Houghton Mifflin Co.
- Bloomfield, L. 1933 Language. New York: Holt, Rinehart and Winston
- Bynon, T. 1977 Historical Linguistics. Cambridge, Cambridge University Press.
- Caldwell,
Robert. 1856 A comparative Grammar of the Dravidian or South Indian Family of Languages. London. Reprinted (1976) Madras. University of Madras Press.
- Campbell, Lyle 1998 Historical Linguistics. Edinburgh: Edinburgh University Press.
- Hock, H.H. 1986 Principles of Historical Linguistics. Amsterdam, Mouton.
- King, Robert D. 1960 Historical Linguistics and Generative Grammar. Englewood cliffs: Prentice – Hall Inc.
- Lehmann, W.P. 1962 Historical Linguistics: An Introduction. New York: Holt, Rinehart and Winston.
- Masica, C.P. 1976 Defining a Linguistic Area: South Asia. Chicago: The University of Chicago Press.
- Milroy, James 1988 Pidgin and Creole Languages. London: Longman.
- Trask, R.L. 1996 Historical Linguistics. London: Arnold.
- Weinreich, Uriel. 1953 Languages in Contact: Findings and Problems. The Hague: Mouton.

&&&&&&