

List of basic linguistics topics

Basic topics in linguistics include:

Contents

- [1 Nature of linguistics](#)
- [2 Branches of linguistics](#)
 - [2.1 Subfields of linguistics](#)
 - [2.2 Schools, movements, and approaches of linguistics](#)
- [3 History of linguistics](#)
 - [3.1 Timeline of discovery of basic linguistics concepts](#)
- [4 Basic questions in linguistics](#)
- [5 Basic concepts](#)
- [6 Linguistics scholars](#)

Nature of linguistics

Linguistics is the [scientific](#) study of human [language](#) and [speech](#), and someone who engages in this study is called a [linguist](#). Concerning the linguistic profession, it may be useful to distinguish linguistics as conducted along two major axes, namely:

- **Theoretical** vs. **Applied**: [Theoretical \(or general\) linguistics](#) is concerned with frameworks for describing individual languages and theories about universal aspects of language; [applied linguistics](#) applies these theories to practical problems such as [language teaching](#), [speech synthesis](#), or [speech therapy](#).
- **Autonomous** vs. **Contextual** (note: these terms are not well-established): Autonomous linguistics studies what [Saussure](#) called *langue* or what [Chomsky](#) calls *I[n]ternal-language*: the nature of language abstracting away from many aspects of its day-to-day usage. Contextual linguistics is concerned with what [Saussure](#) called *parole* or what [Chomsky](#) calls *E[x]ternal-language*, namely how language fits into the world: its social function, or its use in the broader context of human behavior.

Given these distinctions, scholars who call themselves simply *linguists* or *theoretical linguists*, with no further qualification, tend to be concerned with autonomous, theoretical synchronic linguistics, which is acknowledged as the core of the discipline. Linguistic [inquiry](#) is pursued by a wide variety of specialists, who may not all be in harmonious agreement; as [Russ Rymer](#) flamboyantly puts it:

"Linguistics is arguably the most hotly contested property in the academic realm. It is soaked with the blood of [poets](#), [theologians](#), [philosophers](#), [philologists](#), [psychologists](#), [biologists](#), [anthropologists](#), and [neurologists](#), along with whatever blood can be got out of [grammarians](#)."

Divisions, specialties, and subfields

The central concern of autonomous theoretical linguistics is to characterize the nature of human linguistic ability, or competence: to explain what it is that an individual knows when an individual knows a language, and to explain how it is that individuals come to know languages. All humans (setting aside extremely pathological cases) achieve competence in whatever language is spoken around them when they are growing up (or signed, in the case of sign language), without formal instruction being necessary. Non-humans do not. Therefore, there is some basic innate property of humans that causes them to learn language. Since children learn whatever language is spoken around them, there is no genetic basis for the differences between one language and another.

Linguists are engaged in a multiplicity of studies, some of which bear little direct relationship to each other. This is an incomplete list of the branches of linguistics, new ones continue to arise.

Linguistic structures are pairings of meaning and sound (or other externalization). Linguists may be engaged in some subpart of the linguistic structure, some of which bear little direct relationship to each other. This is an incomplete list of its branches, while new ones continue to arise.

- [Phonetics](#), the study of the sounds of human language
- [Phonology](#) (or phonemics), the study of patterns of a language's basic sounds
- [Morphology](#), the study of the internal structure of words
- [Syntax](#), the study of how words combine to form grammatical sentences
- [Semantics](#), the study of the meaning of words ([lexical semantics](#)) and fixed word combinations ([phraseology](#)), and how these combine to form the meanings of sentences
- [Pragmatics](#), the study of how utterances are used (literally, figuratively, or otherwise) in communicative acts
- [Discourse analysis](#), the study of sentences organized into texts
- [Lexicography](#), the compiling of dictionaries. Lexicography could be seen as a branch of applied linguistics.
- [Applied linguistics](#). The application of the methods and results of linguistics to such areas as language teaching; national language policies; translation; language in politics, advertising, classrooms and courts (forensic linguistics).

(Note: Whereas theoretical linguistics is concerned with finding and [describing](#) generalities both within particular languages and among all languages, [applied linguistics](#) takes the results of those findings and *applies* them to other areas. Often *applied linguistics* refers to the use of linguistic research in language teaching, but results of linguistic research are used in many other areas, as well.)

Many areas of applied linguistics today involve the explicit use of computers. [Speech synthesis](#) and [speech recognition](#) use phonetic and phonemic knowledge to provide voice interfaces to computers. Applications of [computational linguistics](#) in [machine translation](#),

[computer-assisted translation](#), and [natural language processing](#) are extremely fruitful areas of applied linguistics which have come to the forefront in recent years with increasing computing power.

- [Computational linguistics](#), Computational linguists study natural languages, such as English and Japanese, rather than computer languages, such as Fortran, Snobol, or Java. The field of computational linguistics has two aims: the technological aim to enable computers to be used as aids in analyzing and processing natural language and the psychological aim to understand, by analogy with computers, more about how people process natural languages. It also includes research on automatic translation, electronic production of artificial speech and the automatic recognition of human speech.

The independent significance of each of these areas is not universally acknowledged, however, and many linguists would agree that the divisions overlap considerably. Nevertheless, each area has core concepts that foster significant scholarly inquiry and research. Intersecting with these specialty domains are fields arranged around the kind of external factors that are considered. For example

- [Language acquisition](#), the study of how language is acquired
- [Historical linguistics](#) or [Diachronic linguistics](#), A branch of linguistics which studies the development of language and languages over time. Historical linguistics uses the methods of the various branches of linguistics (including sociolinguistics, especially in considering the reasons for language change). One thus encounters such subfields as “historical phonology/ morphology/ syntax” etc. It is thus the study of languages whose historical relations are recognizable through similarities in vocabulary, word formation, and syntax
- [Psycholinguistics](#), the study of the mental/cognitive processes and representations underlying language use, for example how memory limitations affect speech production and comprehension. The best developed branch of psycholinguistics is the study of language acquisition.
- [Sociolinguistics](#), the study of social patterns of linguistic variability
- [Stylistics](#), The study of that variation in language which is dependent on the situation in which the language is used and also on the effect the writer/speaker wishes to create on the reader/hearer. Stylistics tries to establish principles capable of explaining the particular choices made by individuals and social groups in their use of language.
- [Literary stylistics](#), The analysis of literary texts applying linguistic methods and theories (phonetics, morphology, syntax, discourse analysis, pragmatics, etc.) with the aim of providing retrievable interpretations which allow comparisons of different texts, genres (fiction, drama and poetry) etc.
- [Neurolinguistics](#), The study of the brain and how it functions in the production, perception, and acquisition of language as well as disorders like aphasia.

[Sociolinguistics](#), [anthropological linguistics](#), and [linguistic anthropology](#) are social sciences that consider the interactions between linguistics and society as a whole.

[Critical discourse analysis](#) is where [rhetoric](#) and [philosophy](#) interact with linguistics.

[Psycholinguistics](#) and [neurolinguistics](#) combine [medical science](#) and linguistics.

Other cross-disciplinary areas of linguistics include [language acquisition](#), [evolutionary linguistics](#), [computational linguistics](#) and [cognitive science](#).

Variation

A substantial part of linguistic investigation is into the nature of the differences among the languages of the world. The nature of variation is very important to an understanding of human linguistic ability in general: if human linguistic ability is very narrowly constrained by biological properties of the species, then languages must be very similar. If human linguistic ability is unconstrained, then languages might vary greatly.

But there are different ways to interpret similarities among languages. For example, the Latin language spoken by the Romans developed into Spanish in Spain and Italian in Italy. Similarities between Spanish and Italian are in many cases due to both being descended from Latin. So in principle, if two languages share some property, this property might either be due to common inheritance or due to some property of the human language faculty.

Often, the possibility of common inheritance can be essentially ruled out. Given the fact that learning language comes quite easily to humans, it can be assumed that languages have been spoken at least as long as there have been biologically modern humans, probably at least fifty thousand years. Independent measures of language change (for example, comparing the language of ancient texts to the daughter languages spoken today) suggest that change is rapid enough to make it impossible to reconstruct a language that was spoken so long ago; as a consequence of this, common features of languages spoken in different parts of the world are not normally taken as evidence for common ancestry.

Even more striking, there are documented cases of [sign languages](#) being developed in communities of congenitally deaf people who could not have been exposed to spoken language. The properties of these sign languages have been shown to conform generally to many of the properties of spoken languages, strengthening the hypothesis that those properties are not due to common ancestry but to more general characteristics of the way languages are learned.

Loosely speaking, the collection of properties which all languages share can be referred to as "[universal grammar](#)" (or UG). However, there is much debate around this topic and the term is used in several different ways.

Universal properties of language may be partly due to universal aspects of human experience; for example all humans experience water, and the fact that all human languages have a word for water is probably not unrelated to this fact. The challenging questions regarding universal grammar generally require one to control this factor. Clearly, experience is part of the process by which individuals learn languages. But experience by itself is not enough, since animals raised around people learn extremely little human language, if any at all.

A more interesting example is this: suppose that all human languages distinguish nouns from verbs (this is generally believed to be true). This would require a more sophisticated explanation, since nouns and verbs do not exist in the world, apart from languages that make use of them.

In general, a property of UG could be due to general properties of human cognition, or due to some property of human cognition that is specific to language. Too little is understood about human cognition in general to allow a meaningful distinction to be made. As a result, generalizations are often stated in theoretical linguistics without a stand being taken on whether the generalization could have some bearing on other aspects of cognition.

Properties of language

It has been understood since the time of the ancient Greeks that languages tend to be organized around grammatical categories such as noun and verb, nominative and accusative, or present and past. The vocabulary and grammar of a language are organized around these fundamental categories.

In addition to making substantial use of discrete categories, language has the important property that it organizes elements into recursive structures; this allows, for example, a noun phrase to contain another noun phrase (as in *the chimpanzee's lips*) or a clause to contain a clause (as in *I think that it's raining*). Though recursion in grammar was implicitly recognized much earlier (for example by [Jespersen](#)), the importance of this aspect of language was only fully realized after the 1957 publication of [Noam Chomsky's](#) book [Syntactic Structures](#),^[1] which presented a formal grammar of a fragment of English. Prior to this, the most detailed descriptions of linguistic systems were of phonological or morphological systems, which tend to be closed and admit little creativity.

Chomsky used a [context-free grammar](#) augmented with transformations. Since then, context-free grammars have been written for substantial fragments of various languages (for example [GPSG](#), for English), but it has been demonstrated that human languages include cross-serial dependencies, which cannot be handled adequately by Context-free grammars. This requires increased power, for example transformations.

An example of a natural-language clause involving a cross-serial dependency is the Dutch^{[2][3]}

Ik denk dat Jan Piet de kinderen zag helpen zwemmen
I think that Jan Piet the children saw help swim
'I think that Jan saw Piet help the children swim'

The important point is that the noun phrases before the verb cluster (*Jan, Piet, de kinderen*) are identified with the verbs in the verb cluster (*zag, helpen, zwemmen*) in left-right order.

This means that natural language formalisms must be relatively powerful in terms of generative capacity. The models currently used ([LFG](#), [HPSG](#), Minimalism) are very powerful, in general too powerful to be computationally tractable in principle. Implementations of them are scaled down.

Details on selected divisions and subfields

Contextual linguistics

Contextual linguistics may include the study of linguistics in interaction with other academic disciplines. Whereas in core theoretical linguistics language is studied for its own sake, the interdisciplinary areas of linguistics consider how language interacts with the rest of the world.

[Sociolinguistics](#), [anthropological linguistics](#), and [linguistic anthropology](#) are social sciences that consider the interactions between linguistics and society as a whole.

[Critical discourse analysis](#) is where [rhetoric](#) and [philosophy](#) interact with linguistics.

[Psycholinguistics](#) and [neurolinguistics](#) combine [medical science](#) and linguistics.

Other cross-disciplinary areas of linguistics include [language acquisition](#), [evolutionary linguistics](#), [computational linguistics](#) and [cognitive science](#).

Applied linguistics

Whereas theoretical linguistics is concerned with finding and [describing](#) generalities both within particular languages and among all languages, [applied linguistics](#) takes the results of those findings and *applies* them to other areas. Often *applied linguistics* refers to the use of linguistic research in language teaching, but results of linguistic research are used in many other areas, as well.

Many areas of applied linguistics today involve the explicit use of computers. [Speech synthesis](#) and [speech recognition](#) use phonetic and phonemic knowledge to provide voice interfaces to computers. Applications of [computational linguistics](#) in [machine translation](#), [computer-assisted translation](#), and [natural language processing](#) are extremely fruitful areas of applied linguistics which have come to the forefront in recent years with increasing computing power. Their influence has had a great effect on theories of syntax

and semantics, as modelling syntactic and semantic theories on computers constrains the theories to [computable](#) operations and provides a more rigorous mathematical basis.

Diachronic linguistics

Whereas the core of theoretical linguistics is concerned with studying languages at a particular point in time (usually the present), diachronic linguistics examines how language changes through time, sometimes over centuries. Historical linguistics enjoys both a rich history (the study of linguistics grew out of historical linguistics) and a strong theoretical foundation for the study of [language change](#).

In universities in the United States, the non-historic perspective seems to have the upper hand. Many introductory linguistics classes, for example, cover historical linguistics only cursorily. The shift in focus to a non-historic perspective started with [Saussure](#) and became predominant with [Noam Chomsky](#).

Explicitly historical perspectives include [historical-comparative linguistics](#) and [etymology](#).

Prescription and description

Main article: [Prescription and description](#).

Research currently performed under the name "linguistics" is purely *descriptive*; linguists seek to clarify the nature of language without passing value judgments or trying to chart future language directions. Nonetheless, there are many professionals and amateurs who also *prescribe* rules of language, holding a particular standard out for all to follow.

Prescriptivists tend to be found among the ranks of language educators and journalists, and not in the actual academic discipline of linguistics. They hold clear notions of what is right and wrong, and may assign themselves the responsibility of ensuring that the next generation use the variety of language that is most likely to lead to "success," often the [acrolect](#) of a particular language. The reasons for their intolerance of "incorrect usage" may include distrust of [neologisms](#), connections to socially-disapproved dialects (i.e., [basilects](#)), or simple conflicts with pet theories. An extreme version of prescriptivism can be found among censors, whose personal mission is to eradicate words and structures which they consider to be destructive to society.

Descriptivists, on the other hand, do not accept the prescriptivists' notion of "incorrect usage." They might describe the usages the other has in mind simply as "idiosyncratic," or they may discover a regularity (a *rule*) that the usage in question follows (in contrast to the common prescriptive assumption that "bad" usage is unsystematic). Within the context of [fieldwork](#), [descriptive linguistics](#) refers to the study of language using a descriptivist approach. Descriptivist methodology more closely resembles scientific methodology in other disciplines.

Speech versus writing

Most contemporary linguists work under the assumption that [spoken language](#) is more fundamental, and thus more important to study than [written language](#). Reasons for this perspective include:

- Speech appears to be a human universal, whereas there have been many [cultures](#) and speech communities that lack written communication;
- People learn to speak and process spoken languages more easily and much earlier than writing;
- A number of [cognitive scientists](#) argue that the [brain](#) has an innate "[language module](#)", [knowledge](#) of which is thought to come more from studying speech than writing, particularly since language as speech is held to be an evolutionary adaptation, whereas writing is a comparatively recent invention.

Of course, linguists agree that the study of written language can be worthwhile and valuable. For linguistic research that uses the methods of [corpus linguistics](#) and [computational linguistics](#), written language is often much more convenient for processing large amounts of linguistic data. Large corpora of spoken language are difficult to create and hard to find, and are typically [transcribed](#) and written.

The study of [writing systems](#) themselves is in any case considered a branch of linguistics.

History of linguistics

Main article: [History of linguistics](#)

Early Indian Vedic texts (Rig Veda 1:164:45; 4:58:3; 10:125) provide a reasonably convincing structure for Languages in general: Language is composed of sentences with four stages of evolution that are expressed in three tenses (past, present and future). The sentences are composed of words that have two distinct forms of existence (vocal form, the word, and perceptual form, the meaning). These words are recognized mainly as verbs that represent real world acts and nouns that take on seven* cases (depending on their mode of participation in real world acts). (* The number, seven, here is not very critical; the message is that the nouns are inflected into appropriate cases to indicate their mode of participation in concerned acts). In [ancient India](#), the [Sanskrit grammarian Pāṇini](#) (c. [520–460 BC](#)) is the earliest known linguist and is often acknowledged as the founder of linguistics. He is most famous for formulating the 3,959 rules of Sanskrit [morphology](#) in the text [Aṣṭādhyāyī](#), which is still in use today. Pāṇini's [grammar](#) of Sanskrit is highly systematised and technical. Inherent in its analytic approach are the concepts of the [phoneme](#), the [morpheme](#) and the [root](#), only recognized by Western linguists some two millennia later. His rules fully describe Sanskrit morphology without any redundancy. A consequence of his grammar's focus on brevity is its highly unintuitive structure, reminiscent of contemporary "machine language" (as opposed to

"human readable" programming languages). His sophisticated logical rules and technique have been widely influential in ancient and modern linguistics.

The [South Indian](#) linguist [Tolkāppiyar](#) (c. [3rd century BC](#)) wrote the *Tolkāppiyam*, the grammar of [Tamil](#), which is also still in use today. [Bhartrihari](#) (c. [450–510](#)) was another important author on [Indic](#) linguistic theory. He theorized the act of speech as being made up of four stages: first, conceptualization of an idea, second, its verbalization and sequencing and third, delivery of speech into atmospheric air, all these by the speaker and last, the comprehension of speech by the listener, the interpreter. The work of Pāṇini, and the later Indian linguist Bhartrihari, had a significant influence on many of the foundational ideas proposed by [Ferdinand de Saussure](#), professor of Sanskrit, who is widely considered the father of modern structural linguistics.

Other early [scholars](#) of linguistics include [Jakob Grimm](#), who devised the principle of consonantal shifts in pronunciation known as [Grimm's Law](#) in 1822, [Karl Verner](#), who discovered [Verner's Law](#), [August Schleicher](#) who created the "Stammbaumtheorie" and [Johannes Schmidt](#) who developed the "Wellentheorie" ("wave model") in 1872. [Ferdinand de Saussure](#) was the founder of modern structural linguistics. [Edward Sapir](#), a leader in American structural linguistics, was one of the first who explored the relations between language studies and anthropology. His methodology had strong influence on all his successors. [Noam Chomsky's](#) formal model of language, [transformational-generative grammar](#), developed under the influence of his teacher [Zellig Harris](#), who was in turn strongly influenced by [Leonard Bloomfield](#), has been the dominant one from the [1960s](#).

Chomsky remains by far the most influential linguist in the world today. Linguists working in frameworks such as [Head-Driven Phrase Structure Grammar](#) (HPSG) or [Lexical Functional Grammar](#) (LFG) stress the importance of formalization and formal rigor in linguistic description, and may distance themselves somewhat from Chomsky's more recent work (the "Minimalist" program for [Transformational grammar](#)), connecting more closely to earlier work of Chomsky's. Linguists working in [Optimality Theory](#) state generalizations in terms of violable rules, which is a greater departure from mainstream linguistics, and linguists working in various kinds of [functional grammar](#) and [Cognitive Linguistics](#) tend to stress the non-autonomy of linguistic knowledge and the non-universality of linguistic structures, thus departing importantly from the Chomskian paradigm.

Branches of linguistics

Subfields of linguistics

- [Computational linguistics](#)

Computational linguistics is an [interdisciplinary](#) field dealing with the [statistical](#) and [logical](#) modeling of [natural language](#) from a computational perspective. This modeling is not limited to any particular field of [linguistics](#). Computational linguistics was formerly usually done by [computer scientists](#) who had specialized in the application of computers

to the processing of a [natural language](#). Recent research has shown that language is much more complex than previously thought, so computational linguistics work teams are now sometimes interdisciplinary, including linguists (specifically trained in linguistics). Computational linguistics draws upon the involvement of linguists, [computer scientists](#), experts in [artificial intelligence](#), [cognitive psychologists](#) and [logicians](#), amongst others.

- [Comparative linguistics](#)

Comparative linguistics (originally **comparative philology**) is a branch of [historical linguistics](#) that is concerned with comparing languages in order to establish their historical relatedness.

Relatedness implies a common origin or [proto-language](#), and comparative linguistics aims to reconstruct proto-languages and specify the changes that have resulted in the documented languages. In order to maintain a clear distinction between attested and reconstructed forms, comparative linguists prefix an asterisk to any form that is not found in surviving texts.

The fundamental technique of comparative linguistics is the [comparative method](#), which aims to compare phonological systems, morphological systems, syntax and the lexicon. In principle, every difference between two related languages should be explicable to a high degree of plausibility, and systematic changes, for example in phonological or morphological systems, are expected to be highly regular. Although the proto-languages reconstructed by the comparative methods are hypothetical, a reconstruction may have predictive power. The most notable example of this is [Saussure's](#) proposal that the [Indo-European consonant](#) system contained [laryngeals](#), a type of consonant attested in no Indo-European language known at the time. The hypothesis was vindicated with the discovery of [Hittite](#), which proved to have exactly the consonants Saussure had hypothesized in the environments he had predicted.

- [Dialectology](#)

Dialectology is a sub-field of [linguistics](#). It studies variations in [language](#) based primarily on geographic distribution (as opposed to variations based on social factors, which are studied in sociolinguistics, or variations based on time, which are studied in [historical linguistics](#)) and the features associated with it. Dialectology treats such topics as divergence of two local [dialects](#) from a common ancestor and synchronic variation. [William Labov](#) is one of the most prominent researchers in this field.

Dialectologists are ultimately concerned with grammatical and syntactical features which correspond to regional areas. Thus they are usually dealing with populations living in their areas for generations without moving, but also with immigrant groups bringing their languages to new settlements.

- [Etymology](#)

Etymology is the study of the origins of [words](#). Through old texts and comparison with other [languages](#), [etymologists](#) reconstruct the [history](#) of words — when they entered a language, from what source, and how their form and meaning changed.

In languages with a long written history etymology makes use of [philology](#), the study of old texts. However, etymologists also apply the methods of [comparative linguistics](#) to reconstruct information about languages that are too old for any direct information (such as writing) to be known. By comparing related languages with a technique known as the [comparative method](#), linguists can make inferences about their shared parent language and its vocabulary. In this way, [word roots](#) have been found which can be traced all the way back to the origin of, for instance, the [Indo-European language family](#).

- [Historical linguistics](#)

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

- to describe and account for observed changes in particular languages
- to describe the history of [speech communities](#)
- 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.

Modern historical linguistics dates from the late [18th century](#) and grew out of the earlier discipline of [philology](#), the study of ancient texts and documents, which goes back to antiquity.

At first historical linguistics was [comparative linguistics](#) and mainly concerned with establishing language families and the reconstruction of prehistoric languages, using the [comparative method](#) and [internal reconstruction](#). The focus was on the well-known [Indo-European languages](#), many of 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 now, however, only a part of a more broadly conceived discipline of historical linguistics. For the [Indo-European languages](#) comparative study is now a highly specialised field and most research is being carried out on the subsequent development of these languages, particularly the development of the modern standard varieties.

- [Grammar](#)

Grammar is the study of [rules](#) governing the use of [language](#). The set of rules governing a particular language's use is also called the **grammar** of the language; thus, each language can be said to have its own distinct grammar. Grammar is part of the general study of language called [linguistics](#).

The subfields of contemporary grammar are [phonetics](#), [phonology](#), [morphology](#), [syntax](#), [semantics](#), and [pragmatics](#). Traditional grammars include only [morphology](#) and [syntax](#).

- [Lexicology](#)

Lexicology is a speciality in [linguistics](#) dealing with the study of the [lexicon](#). The term first appeared in the 1820s, though obviously there were lexicologists before that.

A good example of lexicology at work, and one everyone is familiar with, is the [dictionary](#) or [thesaurus](#). As there are many different types of dictionaries, there are many different types of lexicologists. There are difficulties in working out what simple words such as 'the' mean, and what complicated words mean. Also which words to keep in and which not to include.

- [Linguistic Typology](#)

Linguistic typology aims to answer the question why linguistic diversity is the way it is. It includes three subdisciplines: Qualitative typology deals with the issue of comparing languages and within-language variance, Quantitative typology deals with the distribution of structural patterns in the world's languages, and Theoretical typology explains these distributions.

- [Mathematical linguistics](#)

In [mathematics](#), [logic](#), and [computer science](#), a **formal language** L is a [set](#) of finite-length sequences of elements drawn from a specified finite set A of [symbols](#). Among the more common options that are found in applications, a formal language may be viewed as being analogous to (1) a collection of words or (2) a collection of sentences. In Case 1, the set A is called the [alphabet](#) of L , whose elements are called *words*. In Case 2, the set A is called the *lexicon* or the *vocabulary* of L , whose elements are then called *sentences*. In any case, the mathematical theory that treats formal languages in general is known as *formal language theory*.

Although it is common to hear the term *formal language* used in other contexts to refer to a mode of expression that is more disciplined or more precise than everyday speech, the sense of *formal language* discussed in this article is restricted to its meaning in formal language theory.

An alphabet might be $\{a, b\}$, and a string over that alphabet might be *ababba*.

A typical language over that alphabet, containing that string, would be the set of all strings which contain the same number of symbols *a* and *b*.

The **empty word** (that is, length-zero string) is allowed and is often denoted by e , ε or Λ . While the alphabet is a finite set and every string has finite length, a language may vary

well have infinitely many member strings (because the length of words in it may be unbounded).

- [Morphology](#)

Morphology is a sub-discipline of [linguistics](#) that studies word structure. While words are generally accepted as being the smallest units of [syntax](#), it is clear that in most (if not all) languages, words can be related to other words by rules. For example, [English](#) speakers recognize that the words *dog*, *dogs* and *dog-catcher* are closely related. English speakers recognize these relations by virtue of the unconscious linguistic knowledge they have of the rules of word-formation processes in English. Therefore, these speakers intuit that *dog* is to *dogs* just as *cat* is to *cats*, or *encyclopaedia* is to *encyclopaedias*; similarly, *dog* is to *dog-catcher* as *dish* is to *dishwasher*. The rules comprehended by the speaker in each case reflect specific patterns (or regularities) in the way words are formed from smaller units and how those smaller units interact in speech. In this way, morphology is the branch of linguistics that studies such patterns of word-formation across and within languages, and attempts to explicate formal rules reflective of the knowledge of the speakers of those languages.

- [Phonetics](#)

Phonetics (from the [Greek](#) word φωνή, *phone* = sound/voice) is the study of [sounds](#) and the [human voice](#). It is concerned with the actual properties of speech sounds ([phones](#)) as well as those of non-speech sounds, and their production, audition and perception, as opposed to [phonology](#), which is the study of sound systems and abstract sound units (such as [phonemes](#) and [distinctive features](#)). Phonetics deals with the sounds themselves rather than the contexts in which they are used in languages. Discussions of meaning ([semantics](#)) do not enter at this level of [linguistic analysis](#), therefore.

While [writing systems](#) and [alphabets](#) often attempt to represent the sounds of speech, phoneticians are more concerned with the sounds themselves than the [symbols](#) used to represent them. So close is the relationship between them, however, that many dictionaries list the study of the symbols (more accurately [semiotics](#)) as a part of phonetic studies ^{[[citation needed](#)]}. [Logographic](#) writing systems typically give much less phonetic information, although it is not necessarily non-existent. For instance, in [Chinese characters](#), a *phonetic* is a portion of the character that hints at its pronunciation, while the [radical](#) gives [semantic](#) information. Characters featuring the same phonetic typically have similar pronunciations, but by no means are the pronunciations predictably determined by the phonetic; this is because pronunciations diverged over many centuries while the characters remained the same. Not all Chinese characters are *radical-phonetic compounds*, but a good majority of them are.

Phonetics has three main branches:

- [articulatory phonetics](#), concerned with the positions and movements of the lips, tongue, vocal tract and folds and other [speech organs](#) in producing speech;
- [acoustic phonetics](#), concerned with the properties of the sound waves and how they are received by the inner ear;
- [auditory phonetics](#), concerned with speech perception, principally how the brain forms perceptual representations of the input it receives.

There are over a hundred different phones recognized as distinctive by the [International Phonetic Association](#) (IPA) and transcribed in their [International Phonetic Alphabet](#).

Phonetics was studied as early as 2500 years ago in [ancient India](#), with [Pāṇini](#)'s account of the place and manner of articulation of consonants in his [5th century BCE](#) treatise on [Sanskrit](#). The major [Indic alphabets](#) today, except [Tamil script](#), order their consonants according to Pāṇini's classification.

- [Phonology](#)

Phonology ([Greek](#) *phonē* = voice/sound and *logos* = word/speech), is a subfield of [linguistics](#) which studies the [sound](#) system of a specific [language](#) (or languages). Whereas [phonetics](#) is about the physical production and [perception](#) of the sounds of speech, phonology describes the way sounds function within a given language or across languages.

An important part of phonology is studying which sounds are distinctive units within a language. In [English](#), for example, /p/ and /b/ are distinctive units of sound, (i.e., they are [phonemes](#) / the difference is *phonemic*, or *phonematic*). This can be seen from [minimal pairs](#) such as "pin" and "bin", which mean different things, but differ only in one sound. On the other hand, /p/ is often [pronounced](#) differently depending on its position relative to other sounds, yet these different pronunciations are still considered by [native speakers](#) to be the same "sound". For example, the /p/ in "pin" is [aspirated](#) while the same phoneme in "spin" is not. In some other languages, eg [Thai](#) and [Quechua](#), this same difference of aspiration or non-aspiration does differentiate phonemes.

In addition to the minimal meaningful sounds (the phonemes), phonology studies how sounds alternate, such as the /p/ in English described above, and topics such as [syllable](#) structure, [stress](#), [accent](#), and [intonation](#).

The principles of phonological theory have also been applied to the analysis of [sign languages](#), in which it is argued that the same or a similar phonological system underlies both signed and spoken languages. (Signs are distinguished from [gestures](#) in that the latter are non-linguistic or supply extra meaning alongside the linguistic message.)

- [Pragmatics](#)

In [linguistics](#) and [semiotics](#), **pragmatics** is concerned with bridging the explanatory gap between [sentence meaning](#) and speaker's meaning. The study of how [context](#) influences

the interpretation is then crucial. In this setting, *context* refers to any factor — linguistic, objective, or subjective — that affects the actual [interpretation](#) of signs and expressions.

- [Psycholinguistics](#)

Psycholinguistics or **psychology of language** is the study of the [psychological](#) and [neurobiological](#) factors that enable [humans](#) to acquire, use, and understand [language](#). Initial forays into psycholinguistics were largely philosophical ventures, due mainly to a lack of cohesive data on how the human brain functioned. Modern research makes use of [biology](#), [neuroscience](#), [cognitive science](#), and [information theory](#) to study how the brain processes language. There are a number of subdisciplines; for example, as non-invasive techniques for studying the neurological workings of the brain become more and more widespread, [neurolinguistics](#) has become a field in its own right.

Psycholinguistics covers the cognitive processes that make it possible to generate a grammatical and meaningful [sentence](#) out of [vocabulary](#) and [grammatical structures](#), as well as the processes that make it possible to understand utterances, words, [text](#), etc. Developmental psycholinguistics studies infants' and children's ability to learn language, usually with experimental or at least quantitative methods (as opposed to naturalistic observations such as those made by [Jean Piaget](#) in his research on the development of children).

- [Semantics](#)

Semantics ([Greek](#) *semantikos*, giving signs, significant, symptomatic, from *sema*, [sign](#)) refers to the aspects of [meaning](#) that are expressed in a [language](#), [code](#), or other form of representation. Semantics is contrasted with two other aspects of meaningful expression, namely, [syntax](#), the construction of complex signs from simpler signs, and [pragmatics](#), the practical use of signs by agents or communities of interpretation in particular circumstances and contexts. By the usual convention that calls a study or a theory by the name of its subject matter, **semantics** may also denote the theoretical study of meaning in systems of signs.

Though terminology varies, writers on the subject of meaning generally recognize two sorts of meaning that a significant expression may have: (1) the relation that a sign has to objects and objective situations, actual or possible, and (2) the relation that a sign has to other signs, most especially the sorts of mental signs that are conceived of as [concepts](#).

Most theorists refer to the relation between a sign and its objects, as always including any manner of objective reference, as its [denotation](#). Some theorists refer to the relation between a sign and the signs that serve in its practical interpretation as its [connotation](#), but there are many more differences of opinion and distinctions of theory that are made in this case. Many theorists, especially in the [formal semantic](#), [pragmatic](#), and [semiotic](#) traditions, restrict the application of *semantics* to the denotative aspect, using other terms or altogether ignoring the connotative aspect.

- [Sociolinguistics](#)

Sociolinguistics is the study of the effect of any and all aspects of [society](#), including cultural norms, expectations, and context, on the way [language](#) is used.

It also studies how [lects](#) differ between groups separated by certain [social variables](#), e.g., [ethnicity](#), [religion](#), [status](#), [gender](#), level of [education](#), etc., and how creation and adherence to these rules is used to categorize individuals in [social class](#) or [socio-economic classes](#). As the usage of a language varies from place to place ([dialect](#)), language usage varies among social classes, and it is these [sociolects](#) that sociolinguistics studies.

For example, a sociolinguist might determine through study of social attitudes that [Black English Vernacular](#) would not be considered appropriate language use in a business or professional setting; he or she might also study the [grammar](#), [phonetics](#), [vocabulary](#), and other aspects of this sociolect much as a dialectologist would study the same for a regional dialect.

The study of language variation is concerned with social [constraints](#) determining language in its [contextual environment](#). [Code-switching](#) is the term given to the use of different varieties of language in different social situations.

[William Labov](#) is often regarded as the founder of the study of sociolinguistics.

Sociolinguistics differs from [sociology of language](#) in that the focus of sociolinguistics is the effect of the society on the language, while the latter's focus is on the language's effect on the society.

Schools, movements, and approaches of linguistics

- [Cognitive linguistics](#)

In [linguistics](#) and [cognitive science](#), **cognitive linguistics** (CL) refers to the school of linguistics that views the important essence of language as innately based in [evolutionarily](#)-developed and [speciated](#) faculties, and seeks explanations that advance or fit well into the current understandings of the human [mind](#).

The guiding principle behind this area of linguistics is that language *creation, learning, and usage* must be explained by reference to human [cognition](#) in general —the basic underlying mental processes that apply not only to language, but to all other areas of human [intelligence](#).

It assumes that language is both [situated](#) in a specific [bioregion](#) and that it is [embodied](#). This can be considered a more developed form of the [Sapir-Whorf hypothesis](#), in that not only are language and cognition mutually influential, but also embodied experience and environmental factors of the bioregion.

- [Functionalism](#)

Functional grammar is the name given to any of a range of [functionally](#)-based approaches to the scientific study of language, such as the [grammar](#) model developed by [Simon Dik](#) or [Michael Halliday](#)'s [Systemic functional grammar](#); another important figure in recent linguistic functionalism is [Talmy Givón](#).

Dik characterises functional grammar as follows:

In the functional paradigm a language is in the first place conceptualized as an instrument of social interaction among human beings, used with the intention of establishing communicative relationships. Within this paradigm one attempts to reveal the instrumentality of language with respect to what people do and achieve with it in social interaction. A natural language, in other words, is seen as an integrated part of the [communicative competence](#) of the natural language user.(2, p. 3)

Because of its emphasis on usage, [communicative function](#), and the social context of language, functional grammar differs significantly from other linguistic theories which stress purely [formal](#) approaches to grammar, for instance Chomskyan [generative grammar](#). Functional grammar is strongly associated with the school of [linguistic typology](#) that takes its lead from the work of [Joseph Greenberg](#).

- [Generative linguistics](#)

Generative linguistics is a school of thought within [linguistics](#) that makes use of the concept of a [generative grammar](#). The term "generative grammar" is used in different ways by different people, and the term "generative linguistics" therefore has a range of different, though overlapping, meanings.

Formally, a generative grammar is defined as one that is fully explicit. It is a finite set of rules that can be applied to *generate* exactly those sentences (often, but not necessarily, infinite in number) that are grammatical in a given language (or, of course, particular [dialect](#) or otherwise [sociolinguistically](#) defined way of using a language), and no others. This is the definition that is offered by [Noam Chomsky](#), who popularised the term, and by most dictionaries of linguistics. It is important to note that *generate* is being used as a technical term with a slightly obscure sense. To say that a grammar generates a sentence means that the grammar "assigns a structural description" to the sentence.^[1]

More popularly, but somewhat to the apparent distaste of certain professional linguists including Chomsky, the term is used to define the approach to linguistics taken by Chomsky and his followers. Chomsky's approach is characterised by the use of [transformational grammar](#) - a theory that has changed greatly since it was first promulgated by Chomsky in his [1957](#) book [Syntactic Structures](#) - and by the assertion of a strong linguistic [nativism](#) (and therefore an assertion that some set of fundamental characteristics of all human languages must be the same). The term "generative linguistics" is often applied to the earliest version of Chomsky's transformational

grammar, which was associated with a distinction between "[Deep Structure](#)" and "[Surface Structure](#)" of sentences.

Chomsky also launched his approach to linguistics with a virulent attack on alternative approaches, in particular the [behaviorist](#) view then popular, in the form in which it had been put forward by [B. F. Skinner](#) in a book also published in 1957, *Verbal Behavior*. A final, and still looser, meaning of "generative linguistics", therefore, might be summarised as "anti-Skinnerian linguistics" - or just generalised anti-behaviorism.

[Psycholinguistics](#), which in the early [1960s](#) was developing rapidly as part of the general movement towards [cognitive psychology](#), found this anti-behaviorist emphasis congenial, and rapidly absorbed many Chomskian ideas including the notion of generative grammar. However, as both cognitive psychology and psycholinguistics have matured, they have found less and less use for generative linguistics, not least because Chomsky has repeatedly emphasised that he never intended to specify the mental processes by which people actually generate sentences, or parse sentences that they hear or read.

[Cognitive linguistics](#) emerged in the latter years of the twentieth century as an alternative linguistic paradigm to generative linguistics. Cognitive linguistics seeks to unify the understanding of language with the understanding of how specific neural structures function biologically. This is more a difference in practical research strategy than in philosophy: in principle, neurological evidence has always been considered relevant by generative linguists, but in practice it has usually been regarded as too inconclusive and open to interpretation to be of much use. However, some researchers within generative linguistics (e.g. [Alec Marantz](#)) publish in neurolinguistics.

- [Geneva School](#)

Geneva School refers to the group of linguists based in Geneva who pioneered modern [structural linguistics](#). The most prominent figure of this school was [Ferdinand de Saussure](#). Other important colleagues and students of Saussure who comprise this school include [Albert Sechehaye](#), [Albert Riedlinger](#), Sergej Karcevski and [Charles Bally](#).

The two successors of de Ferdinand de Saussure

One significant linguistic book connected with this school is 'Cours de linguistique générale', the main work of de Saussure, which was published by his students Charles Bally and Albert Sechehaye. The book was based on lectures with this title that de Saussure gave three times in Geneva from 1906 to 1912. Sechehaye and Bally took not part themselves in these lecture classes. They used notes from other students, the most important of these students was Albert Riedlinger, who provided them the most material. Furthermore Bally and Sechehaye continued to develop de Saussure's theories, mainly focusing on the linguistic research of speech. Sechehaye also concentrated on syntactic problems.

- [Neo-Grammarians](#)

The **Neogrammarians** (also **Young Grammarians**, German *Junggrammatiker*) were a [German](#) school of [linguists](#), originally at the [University of Leipzig](#), in the late [19th century](#) who proposed the **Neogrammarian hypothesis** of the regularity of [sound change](#). According to this hypothesis, a [diachronic](#) sound change affects simultaneously all words in which its environment is met, without exception. [Verner's law](#) is a famous example of the Neogrammarian hypothesis, as it resolved an apparent exception to [Grimm's law](#). The Neogrammarian hypothesis was the first hypothesis of sound change to attempt to follow the principle of [falsifiability](#) according to [scientific method](#). Today this hypothesis is considered more of a guiding principle than an exceptionless fact, as numerous examples of [lexical diffusion](#) (where a sound change affects only a few words at first and then gradually spreads to other words) have been attested.

Other contributions of the Neogrammarians to general linguistics were:

The object of linguistic investigation is not the language system, but rather the [idiolect](#), that is, language as it is localized in the individual, and therefore is directly observable.

Autonomy of the sound level: being the most observable aspect of language, the sound level is seen as the most important level of description, and absolute autonomy of the sound level from syntax and semantics is assumed.

Historicism: the chief goal of linguistic investigation is the description of the [historical change](#) of a language.

Analogy: if the premise of the inviolability of sound laws fails, [analogy](#) can be applied as an explanation if plausible. Thus, exceptions are understood to be a (regular) adaptation to a related form.

- [Prague School](#)

The **Prague linguistic circle** or "**Prague school**" ([French](#) *Cercle linguistique de Prague*, [Czech](#) *Pražský lingvistický kroužek*) was an influential group of [literary critics](#) and [linguists](#) in [Prague](#). Its proponents developed methods of [structuralist literary analysis](#) during the years 1928–1939. It has had significant continuing influence on [linguistics](#) and [semiotics](#). After [WWII](#), the circle was disbanded but the Prague School continued as a major force in linguistic functionalism (distinct from the [Copenhagen school](#) or English Firthian — later [Hallidean](#) — linguistics).

The Prague linguistic circle included Russian emigrés such as [Roman Jakobson](#), [Nikolay Trubetzkoy](#), and [Sergei Karcevskiy](#), as well as the famous Czech literary scholars [René Wellek](#) and [Jan Mukařovský](#). Among its founders was the eminent [Czech](#) linguist [Vilém Mathesius](#) (President of PLC until his death in 1945).

The group's work before World War II was published in the *Travaux Linguistique* and its theses outlined in a collective contribution to the World's Congress of Slavists. The

Travaux were briefly resurrected in the 1960s with a special issue on the concept of [center and periphery](#) and are now being published again by John Benjamins. The group's Czech work is published in *Slovo a slovesnost*. English translations of the Circle's seminal works were published by the Czech linguist [Josef Vachek](#) in several collections.

- [Prescription and description](#)

In [linguistics](#), **prescription** is the laying down or *prescribing* of normative rules for the use of a language, or the making of recommendations for effective language usage. It includes the mechanisms for establishing and maintaining an interregional language or a standardised spelling system. It can also include arbitrary declarations of what particular individuals consider to be good taste, and if these tastes are conservative, prescription may be (or appear to be) resistant to natural language evolution.

The **description** of language, which observes and records how language is used in practice, is the basis of all linguistic research. Serious scholarly descriptive work is usually based on text or corpus analysis, or on field studies, but the term "description" includes each individual's observations of their own language usage.

Prescription and description are often seen as opposites, in the sense that one declares how language *should be* while the other declares how language *is*. But they can also be complementary, and usually exist in a dynamic tension to each other. Most commentators on language show elements of both prescription and description in their thinking, and popular debate on language issues frequently revolves around the question of how to balance these.

- [Stratificational linguistics](#)

Stratificational Linguistics is a view of linguistics advocated by [Sydney Lamb](#). His theories advocate that language usage and production is stratificational in nature.

Specifically, that there are separate 'strata' or levels in the brain used for language. Each level provides actualization or 'realization' for the next higher level, and the elements on its level are similar to each other. Several strata are involved in the production of a sound from an initial idea.

Some strata include:

[Phoneme](#) as the unit on the [Phonemic](#) strata.

[Lexeme](#) as a unit on the [Lexical](#) strata.

[Morpheme](#) as the unit on the [Morphemic](#) strata

[Sememe](#) as the unit on the [Semantic](#) strata.

- [Structuralism](#)

Structuralism is an approach in [academic disciplines](#) that explores the relationships between fundamental elements of some kind, upon which some higher mental, linguistic, social, cultural etc. "structures" are built, through which then [meaning](#) is produced within a particular person, system, or culture.

Structuralism appeared in academic psychology for the first time in the [19th century](#) and then reappeared in the second half of the [20th century](#), when it grew to become one of the most popular approaches in the academic fields that are concerned with analyzing language, culture, and society. The work of [Ferdinand de Saussure](#) concerning [linguistics](#) is generally considered to be a starting point of the [20th century](#) structuralism. The term of "structuralism" itself appeared in relation to [French anthropologist Claude Lévi-Strauss](#)' works, and gave rise, in [France](#), to the "structuralist movement," which gathered thinkers such as [psychoanalyst Lacan](#), [Foucault](#) or [Althusser](#) and [Poulantzas](#)' [structural Marxism](#). One should note that almost all members of this so-called movement denied that they were part of it. [Post-structuralism](#) attempted to distinguish itself from the use of the structural [method](#). It is also worth noting that it has been influential in some part of social sciences particularly in the field of [sociology](#) while in the field of [economics](#) it is barely mentioned.

- [Systemic linguistics](#)

Systemic linguistics is an approach to [linguistics](#) that considers [language](#) as a [system](#). It was developed by [M.A.K. Halliday](#). It is considered "functional" rather than "[generative](#)" in linguistic orientation.

- [SIL International](#)

SIL International (*Summer Institute of [Linguistics](#)*.) is a worldwide [non-profit evangelical Christian](#) organization whose main purpose is to study, develop and document lesser-known [languages](#) in order to expand [linguistic](#) knowledge, promote [literacy](#) and aid [minority language](#) development.

- [Tagmemics](#)

Tagmemic discourse theory, as developed from the work of linguist, Kenneth L. Pike, is a theory of discourse founded upon certain axioms about human behavior and language use that foreground the situatedness of all communication and the necessity of viewing every act of discourse as a form-meaning composite inseparable from communicators, their audiences, and the varied worlds they may construct and inhabit through the use of language.

Timeline of discovery of basic linguistics concepts

When were the basic concepts first described and by whom?

- [Ancient Sanskrit grammarians](#)
- Ancient Greek study of language
- Roman elaborations of Greek study
- Medieval philosophical work in Latin
- Beginnings of modern linguistics in the 19th century
- Behaviorism and mental *tabula rasa* hypothesis
- Chomsky and functionalism
- Generative grammar leads to generative phonology and semantics
- Birth of Nicaraguan Sign Language
- Alternate syntactic systems develop in 80s
- Computational linguistics becomes feasible the late 80s
- Neurolinguistics and the biological basis of cognition
- Controversy over [Pirahã](#) number conception

Basic questions in linguistics

What are the basic questions asked in linguistics?

1. What is language?
2. How did it/does it evolve?
3. How does language serve as a medium of communication?
4. How does language serve as a medium of thinking?
5. What is common to all languages?
6. How do languages differ?
7. Is X a language or a dialect?

(The answers are not necessarily basic or easy to understand.)

Basic concepts

- [Morphology](#)
 - [morpheme](#), [inflection](#), [paradigm](#), [derivation](#), [compound](#)
- [Phonology](#)
 - [phoneme](#), [segment](#) [syllable](#), [foot](#), [stress](#)
- [Grammar](#)
 - [tense](#), [aspect](#), [grammatical number](#), [grammatical gender](#), [declension](#)
- [Syntax](#)
 - [phrase](#), [clause](#), [grammatical function](#), [grammatical voice](#)
- [Semantics](#)
 - [entailment](#), [truth condition](#), [compositionality](#)
- [Pragmatics](#)
 - [presupposition](#), [implicature](#), [deixis](#)

Linguistics scholars

People who had a significant influence on the development of the field

- [Benjamin Lee Whorf](#)
- [Claude Levi-Strauss](#)
- [Edward Sapir](#)
- [Ferdinand de Saussure](#)
- [Franz Bopp](#)
- [August Schleicher](#)
- [John Langshaw Austin](#)
- [John R. Searle](#)
- [Louis Hjelmslev](#)
- [Kenneth L. Pike](#)
- [M.A.K. Halliday](#)
- [Noam Chomsky](#)
- [Pāṇini](#)
- [Rasmus Rask](#)
- [Roman Jakobson](#)
- [Sir William Jones](#)
- [Noah Webster](#)