Language

Language is a structured system of communication that consists of grammar and vocabulary. It is the primary means by which humans convey meaning, both in spoken and written forms, and may also be conveyed through sign languages. The vast majority of human languages have developed writing systems that allow for the recording and preservation of the sounds or signs of language. Human language is characterized by its cultural and historical diversity, with significant variations observed between cultures and across time.[1] Human languages possess the properties of productivity and displacement, which enable the creation of an infinite number of sentences, and the ability to refer to objects, events, and ideas that are not immediately present in the discourse. The use of human language relies on social convention and is acquired through learning.

Estimates of the number of human languages in the world vary between 5,000 and 7,000. Precise estimates depend on an arbitrary distinction (dichotomy) established between languages and dialects.[2] Natural languages are spoken, signed, or both; however, any language can be encoded into secondary media using auditory, visual, or tactile stimuli – for example, writing, whistling, signing, or braille. In other words, human language is modality-independent, but written or signed language is the way to inscribe or encode the natural human speech or gestures.

Depending on philosophical perspectives regarding the definition of language and meaning, when used as a general concept, "language" may refer to the cognitive ability to learn and use systems of complex communication, or to describe the set of rules that makes up these systems, or the set of utterances that can be produced from those rules. All languages rely on the process of semiosis to relate signs to particular meanings. Oral, manual and tactile languages contain a phonological system that governs how symbols are used to form sequences known as words or morphemes, and a syntactic system that governs how words and morphemes are combined to form phrases and utterances.

The scientific study of language is called linguistics. Critical examinations of languages, such as philosophy of language, the relationships between language and thought, how words represent experience, etc., have been debated at least since Gorgias and Plato in ancient Greek civilization. Thinkers such as Jean-Jacques Rousseau (1712–1778) have argued that language originated from emotions, while others like Immanuel Kant (1724–1804) have argued that languages originated from rational and logical thought. Twentieth century philosophers such as Ludwig Wittgenstein (1889–1951) argued that philosophy is really the study of language itself. Major figures in contemporary linguistics of these times include Ferdinand de Saussure and Noam Chomsky.

Language is thought to have gradually diverged from earlier primate communication systems when early hominins acquired the ability to form a theory of mind and shared intentionality.[3][4] This development is sometimes thought to have coincided with an increase in brain volume, and many linguists see the structures of language as having evolved to serve specific communicative and social functions. Language is processed in many different locations in the human brain, but especially in Broca's and Wernicke's areas. Humans acquire language through social interaction in early childhood, and children generally speak fluently by approximately three years old. Language and culture are codependent. Therefore, in addition to its strictly communicative uses, language has social uses such as signifying group identity, social stratification, as well as use for social grooming and entertainment.

Languages evolve and diversify over time, and the history of their evolution can be reconstructed by comparing modern languages to determine which traits their ancestral languages must have had in order for the later developmental stages to occur. A group of languages that descend from a common ancestor is known as a language family; in contrast, a language that has been demonstrated to not have any living or non-living relationship with another language is called a language isolate. There are also many unclassified languages whose relationships have not been established, and spurious languages may have not existed at all. Academic consensus holds that between 50% and 90% of languages spoken at the beginning of the 21st century will probably have become extinct by the year 2100.[5][6][7]

Definitions

The English word *language* derives ultimately from Proto-Indo-European *$dn̥ǵʰwéh_2s$* "tongue, speech, language" through Latin *lingua*, "language; tongue", and Old French *language*.[8] The word is sometimes used to refer to codes, ciphers, and other kinds of artificially constructed communication systems such as formally defined computer languages used for computer programming. Unlike conventional human languages, a formal language in this sense is a system of signs for encoding and decoding information. This article specifically concerns the properties of natural human language as it is studied in the discipline of linguistics.

As an object of linguistic study, "language" has two primary meanings: an abstract concept, and a specific linguistic system, e.g. "French". The Swiss linguist Ferdinand de Saussure, who defined the modern discipline of linguistics, first explicitly formulated the distinction using the French word *language* for language as a concept, *langue* as a specific instance of a language system, and *parole* for the concrete usage of speech in a particular language.[9]

When speaking of language as a general concept, definitions can be used which stress different aspects of the phenomenon.[10] These definitions also entail different approaches and understandings of language, and they also inform different and often incompatible schools of linguistic theory.[11] Debates about the nature and origin of language go back to the ancient world. Greek philosophers such as Gorgias and Plato debated the relation between words, concepts and reality. Gorgias argued that language could represent neither the objective experience nor human experience, and that communication and truth were therefore impossible. Plato maintained that communication is possible because language represents ideas and concepts that exist independently of, and prior to, language.[12]
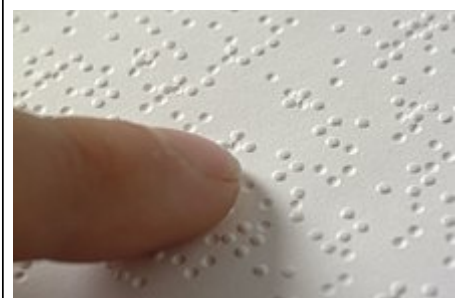

A mural in Teotihuacan, Mexico (c. 2nd century) depicting a person emitting a speech scroll from his mouth, symbolizing speech

Cuneiform is the first known form of written language, but spoken language predates writing by at least many tens of thousands of years.

Children of deaf adults using American Sign Language

During the Enlightenment and its debates about human origins, it became fashionable to speculate about the origin of language. Thinkers such as Rousseau and Johann Gottfried Herder argued

that language had originated in the instinctive expression of emotions, and that it was originally closer to music and poetry than to the logical expression of rational thought. Rationalist philosophers such as Kant and René Descartes held the opposite view. Around the turn of the 20th century, thinkers began to wonder about the role of language in shaping our experiences of the world – asking whether language simply reflects the objective structure of the world, or whether it creates concepts that in turn impose structure on our experience of the objective world. This led to the question of whether philosophical problems are really firstly linguistic problems. The resurgence of the view that language plays a significant role in the creation and circulation of concepts, and that the study of philosophy is essentially the study of language, is associated with what has been called the linguistic turn and philosophers such as Wittgenstein in 20th-century philosophy. These debates about language in relation to meaning and reference, cognition and consciousness remain active today.[13]

Mental faculty, organ or instinct

One definition sees language primarily as the mental faculty that allows humans to undertake linguistic behaviour: to learn languages and to produce and understand utterances. This definition stresses the universality of language to all humans, and it emphasizes the biological basis for the human capacity for language as a unique development of the human brain. Proponents of the view that the drive to language acquisition is innate in humans argue that this is supported by the fact that all cognitively normal children raised in an environment where language is accessible will acquire language without formal instruction. Languages may even develop spontaneously in environments where people live or grow up together without a common language; for example, creole languages and spontaneously developed sign languages such as Nicaraguan Sign Language. This view, which can be traced back to the philosophers Kant and Descartes, understands language to be largely innate, for example, in Chomsky's theory of Universal Grammar, or American philosopher Jerry Fodor's extreme innatist theory. These kinds of definitions are often applied in studies of language within a cognitive science framework and in neurolinguistics.[14][15]

Formal symbolic system

Another definition sees language as a formal system of signs governed by grammatical rules of combination to communicate meaning. This definition stresses that human languages can be described as closed structural systems consisting of rules that relate particular signs to particular meanings.[16] This structuralist view of language was first introduced by Ferdinand de Saussure,[17] and his structuralism remains foundational for many approaches to language.[18]

Some proponents of Saussure's view of language have advocated a formal approach which studies language structure by identifying its basic elements and then by presenting a formal account of the rules according to which the elements combine in order to form words and sentences. The main proponent of such a theory is Noam Chomsky, the originator of the generative theory of grammar, who has defined language as the construction of sentences that can be generated using transformational grammars.[19] Chomsky considers these rules to be an innate feature of the human mind and to constitute the rudiments of what language is.[20] By way of contrast, such transformational grammars are also commonly used in

formal logic, in formal linguistics, and in applied computational linguistics.[21][22] In the philosophy of language, the view of linguistic meaning as residing in the logical relations between propositions and reality was developed by philosophers such as Alfred Tarski, Bertrand Russell, and other formal logicians.

Tool for communication

Yet another definition sees language as a system of communication that enables humans to exchange verbal or symbolic utterances. This definition stresses the social functions of language and the fact that humans use it to express themselves and to manipulate objects in their environment. Functional theories of grammar explain grammatical structures by their communicative functions, and understand the grammatical structures of language to be the result of an adaptive process by which grammar was "tailored" to serve the communicative needs of its users.[23][24]

A conversation in American Sign Language

This view of language is associated with the study of language in pragmatic, cognitive, and interactive frameworks, as well as in sociolinguistics and linguistic anthropology. Functionalist theories tend to study grammar as dynamic phenomena, as structures that are always in the process of changing as they are employed by their speakers. This view places importance on the study of linguistic typology, or the classification of languages according to structural features, as it can be shown that processes of grammaticalization tend to follow trajectories that are partly dependent on typology.[22] In the philosophy of language, the view of pragmatics as being central to language and meaning is often associated with Wittgenstein's later works and with ordinary language philosophers such as J.L. Austin, Paul Grice, John Searle, and W.O. Quine.[25]

Distinctive features of human language

A number of features, many of which were described by Charles Hockett and called design features[26] set human language apart from communication used by non-human animals.

Communication systems used by other animals such as bees or apes are closed systems that consist of a finite, usually very limited, number of possible ideas that can be expressed.[27] In contrast, human language is open-ended and productive, meaning that it allows humans to produce a vast range of utterances from a finite set of elements, and to create new words and sentences. This is possible because human language is based on a dual code, in which a finite number of elements which are meaningless in themselves (e.g. sounds, letters or gestures) can be combined to form an infinite number of larger units of meaning (words and sentences).[28] However, one study has demonstrated that an Australian bird, the chestnut-crowned babbler, is capable of using the same acoustic elements in different arrangements to create two functionally distinct vocalizations.[29] Additionally, pied babblers have demonstrated the ability to generate two functionally distinct vocalisations composed of the same sound type, which can only be distinguished by the number of repeated elements.[30]

Several species of animals have proved to be able to acquire forms of communication through social learning: for instance a bonobo named Kanzi learned to express itself using a set of symbolic lexigrams. Similarly, many species of birds and whales learn their songs by imitating other members of their species. However, while some animals may acquire large numbers of words and symbols,[note 1] none have been able to learn as many different signs as are generally known by an average 4 year old human, nor have any acquired anything resembling the complex grammar of human language.[32]

Human languages differ from animal communication systems in that they employ grammatical and semantic categories, such as noun and verb, present and past, which may be used to express exceedingly complex meanings.[32] It is distinguished by the property of recursivity: for example, a noun phrase can contain another noun phrase (as in "[[the chimpanzee]'s lips]") or a clause can contain another clause (as in "[I see [the dog is running]]").[4] Human language is the only known natural communication system whose adaptability may be referred to as *modality independent*. This means that it can be used not only for communication through one channel or medium, but through several. For example, spoken language uses the auditive modality, whereas sign languages and writing use the visual modality, and braille writing uses the tactile modality.[33]

Human language is unusual in being able to refer to abstract concepts and to imagined or hypothetical events as well as events that took place in the past or may happen in the future. This ability to refer to events that are not at the same time or place as the speech event is called *displacement*, and while some animal communication systems can use displacement (such as the communication of bees that can communicate the location of sources of nectar that are out of sight), the degree to which it is used in human language is also considered unique.[28]

Origin

Theories about the origin of language differ in regard to their basic assumptions about what language is.[35] Some theories are based on the idea that language is so complex that one cannot imagine it simply appearing from nothing in its final form, but that it must have evolved from earlier pre-linguistic systems among our pre-human ancestors. These theories can be called continuity-based theories. The opposite viewpoint is that language is such a unique human trait that it cannot be compared to anything found among non-humans and that it must therefore have appeared suddenly in the transition from pre-hominids to early man. These theories can be defined as discontinuity-based. Similarly, theories based on the generative view of language pioneered by Noam Chomsky see language mostly as an innate faculty that is largely genetically encoded, whereas functionalist theories see it as a system that is largely cultural, learned through social interaction.[36]

The Tower of Babel by Pieter Bruegel the Elder. Oil on board, 1563. Humans have speculated about the origins of language throughout history. The Biblical myth of the Tower of Babel is one such account; other cultures have different stories of how language arose.[34]

Continuity-based theories are held by a majority of scholars, but they vary in how they envision this development. Those who see language as being mostly innate, such as psychologist Steven Pinker, hold the precedents to be animal cognition,[15] whereas those who see language as a socially learned tool of communication, such as psychologist Michael Tomasello, see it as having developed from animal communication in primates: either gestural or vocal communication to assist in cooperation.[37] Other continuity-based models see language as having developed from music, a view already espoused by Rousseau, Herder, Humboldt, and Charles Darwin. A prominent proponent of this view is archaeologist Steven Mithen.[38] Stephen Anderson states that the age of spoken languages is estimated at 60,000 to 100,000 years[39] and that:

> Researchers on the evolutionary origin of language generally find it plausible to suggest that language was invented only once, and that all modern spoken languages are thus in some way related, even if that relation can no longer be recovered ... because of limitations on the methods available for reconstruction.[40]

Because language emerged in the early prehistory of man, before the existence of any written records, its early development has left no historical traces, and it is believed that no comparable processes can be observed today. Theories that stress continuity often look at animals to see if, for example, primates display any traits that can be seen as analogous to what pre-human language must have been like. Early human fossils can be inspected for traces of physical adaptation to language use or pre-linguistic forms of symbolic behaviour. Among the signs in human fossils that may suggest linguistic abilities are: the size of the brain relative to body mass, the presence of a larynx capable of advanced sound production and the nature of tools and other manufactured artifacts.[41]

It was mostly undisputed that pre-human australopithecines did not have communication systems significantly different from those found in great apes in general. However, a 2017 study on *Ardipithecus ramidus* challenges this belief.[42] Scholarly opinions vary as to the developments since the appearance of the genus *Homo* some 2.5 million years ago. Some scholars assume the development of primitive language-like systems (proto-language) as early as *Homo habilis* (2.3 million years ago) while others place the development of primitive symbolic communication only with *Homo erectus* (1.8 million years ago) or *Homo heidelbergensis* (0.6 million years ago), and the development of language proper with anatomically modern *Homo sapiens* with the Upper Paleolithic revolution less than 100,000 years ago.[43][44]
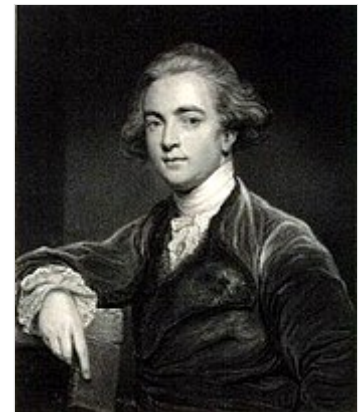
Chomsky is one prominent proponent of a discontinuity-based theory of human language origins.[36] He suggests that for scholars interested in the nature of language, "talk about the evolution of the language capacity is beside the point."[45] Chomsky proposes that perhaps "some random mutation took place [...] and it reorganized the brain, implanting a language organ in an otherwise primate brain."[46] Though cautioning against taking this story literally, Chomsky insists that "it may be closer to reality than many other fairy tales that are told about evolutionary processes, including language."[46]

Study

The study of language, linguistics, has been developing into a science since the first grammatical descriptions of particular languages in India more than 2000 years ago, after the development of the Brahmi script. Modern linguistics is a science that concerns itself with all aspects of language, examining it from all of the theoretical viewpoints described above.[47]

Subdisciplines

The academic study of language is conducted within many different disciplinary areas and from different theoretical angles, all of which inform modern approaches to linguistics. For example, descriptive linguistics examines the grammar of single languages, theoretical linguistics develops theories on how best to conceptualize and define the nature of language based on data from the various extant human languages, sociolinguistics studies how languages are used for social purposes informing in turn the study of the social functions of language and grammatical description, neurolinguistics studies how language is processed in the human brain and

William Jones discovered the family relation between Latin and Sanskrit, laying the ground for the discipline of historical linguistics.

allows the experimental testing of theories, computational linguistics builds on theoretical and descriptive linguistics to construct computational models of language often aimed at processing natural language or at

testing linguistic hypotheses, and historical linguistics relies on grammatical and lexical descriptions of languages to trace their individual histories and reconstruct trees of language families by using the comparative method.[48]

Early history

The formal study of language is often considered to have started in India with Pāṇini, the 5th century BC grammarian who formulated 3,959 rules of Sanskrit morphology. However, Sumerian scribes already studied the differences between Sumerian and Akkadian grammar around 1900 BC. Subsequent grammatical traditions developed in all of the ancient cultures that adopted writing.[49]
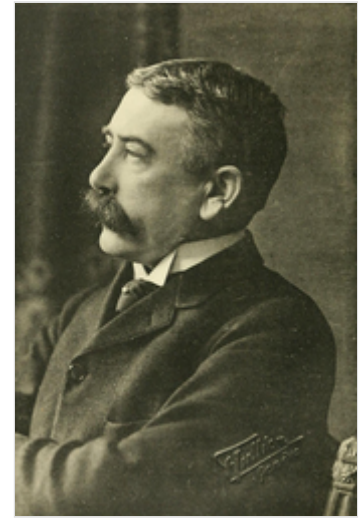
In the 17th century AD, the French Port-Royal Grammarians developed the idea that the grammars of all languages were a reflection of the universal basics of thought, and therefore that grammar was universal. In the 18th century, the first use of the comparative method by British philologist and expert on ancient India William Jones sparked the rise of comparative linguistics.[50] The scientific study of language was broadened from Indo-European to language in general by Wilhelm von Humboldt. Early in the 20th century, Ferdinand de Saussure introduced the idea of language as a static system of interconnected units, defined through the oppositions between them.[17]

Ferdinand de Saussure developed the structuralist approach to studying language.

By introducing a distinction between diachronic and synchronic analyses of language, he laid the foundation of the modern discipline of linguistics. Saussure also introduced several basic dimensions of linguistic analysis that are still fundamental in many contemporary linguistic theories, such as the distinctions between syntagm and paradigm, and the Langue-parole distinction, distinguishing language as an abstract system (*langue*), from language as a concrete manifestation of this system (*parole*).[51]
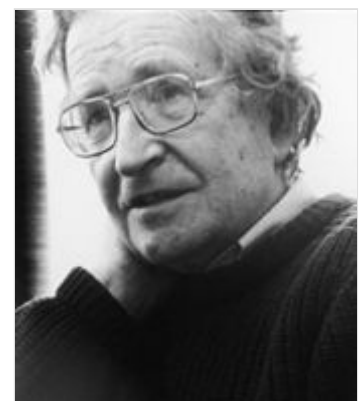
Modern linguistics

In the 1960s, Noam Chomsky formulated the generative theory of language. According to this theory, the most basic form of language is a set of syntactic rules that is universal for all humans and which underlies the grammars of all human languages. This set of rules is called Universal Grammar; for Chomsky, describing it is the primary objective of the discipline of linguistics. Thus, he considered that the grammars of individual languages are only of importance to linguistics insofar as they allow us to deduce the universal underlying rules from which the observable linguistic variability is generated.[52]

Noam Chomsky is one of the most important linguistic theorists of the 20th century.

In opposition to the formal theories of the generative school, functional theories of language propose that since language is fundamentally a tool, its structures are best analyzed and understood by reference to their functions. Formal theories of grammar seek to define the different elements of language and describe the way they relate to each other as systems of formal rules or operations, while functional theories seek to define the functions performed by language and then relate them to the linguistic elements that carry them out.[22][note 2] The framework of cognitive

linguistics interprets language in terms of the concepts (which are sometimes universal, and sometimes specific to a particular language) which underlie its forms. Cognitive linguistics is primarily concerned with how the mind creates meaning through language.[54]

Physiological and neural architecture of language and speech

Speaking is the default modality for language in all cultures. The production of spoken language depends on sophisticated capacities for controlling the lips, tongue and other components of the vocal apparatus, the ability to acoustically decode speech sounds, and the neurological apparatus required for acquiring and producing language.[55] The study of the genetic bases for human language is at an early stage: the only gene that has definitely been implicated in language production is FOXP2, which may cause a kind of congenital language disorder if affected by mutations.[56]

The brain

The brain is the coordinating center of all linguistic activity; it controls both the production of linguistic cognition and of meaning and the mechanics of speech production. Nonetheless, our knowledge of the neurological bases for language is quite limited, though it has advanced considerably with the use of modern imaging techniques. The discipline of linguistics dedicated to studying the neurological aspects of language is called neurolinguistics.[57]

Language Areas of the brain.
- 🟧 Angular gyrus
- 🟨 Supramarginal gyrus
- 🟦 Broca's area
- 🟩 Wernicke's area
- 🟥 Primary auditory cortex

Early work in neurolinguistics involved the study of language in people with brain lesions, to see how lesions in specific areas affect language and speech. In this way, neuroscientists in the 19th century discovered that two areas in the brain are crucially implicated in language processing. The first area is Wernicke's area, which is in the posterior section of the superior temporal gyrus in the dominant cerebral hemisphere. People with a lesion in this area of the brain develop receptive aphasia, a condition in which there is a major impairment of language comprehension, while speech retains a natural-sounding rhythm and a relatively normal sentence structure. The second area is Broca's area, in the posterior inferior frontal gyrus of the dominant hemisphere. People with a lesion to this area develop expressive aphasia, meaning that they know what they want to say, they just cannot get it out.[58] They are typically able to understand what is being said to them, but unable to speak fluently. Other symptoms that may be present in expressive aphasia include problems with word repetition. The condition affects both spoken and written language. Those with this aphasia also exhibit ungrammatical speech and show inability to use syntactic information to determine the meaning of sentences. Both expressive and receptive aphasia also affect the use of sign language, in analogous ways to how they affect speech, with expressive aphasia causing signers to sign slowly and with incorrect grammar, whereas a signer with receptive aphasia will sign fluently, but make little sense to others and have difficulties comprehending others' signs. This shows that the impairment is specific to the ability to use language, not to the physiology used for speech production.[59][60]
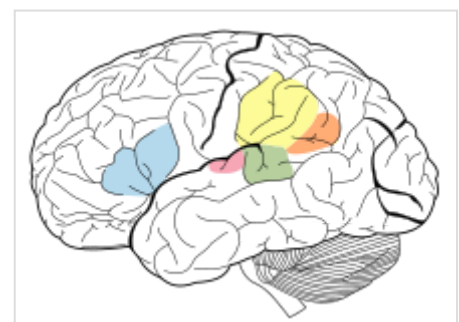
With technological advances in the late 20th century, neurolinguists have also incorporated non-invasive techniques such as functional magnetic resonance imaging (fMRI) and electrophysiology to study language processing in individuals without impairments.[57]

Anatomy of speech

Spoken language relies on human physical ability to produce sound, which is a longitudinal wave propagated through the air at a frequency capable of vibrating the ear drum. This ability depends on the physiology of the human speech organs. These organs consist of the lungs, the voice box (larynx), and the upper vocal tract – the throat, the mouth, and the nose. By controlling the different parts of the speech apparatus, the airstream can be manipulated to produce different speech sounds.[61]

The sound of speech can be analyzed into a combination of segmental and suprasegmental elements. The segmental elements are those that follow each other in sequences, which are usually represented by distinct letters in alphabetic scripts, such as the Roman script. In free flowing speech, there are no clear boundaries between one segment and the next, nor usually are there any audible pauses between them. Segments therefore are distinguished by their distinct sounds which are a result of their different articulations, and can be either vowels or consonants. Suprasegmental phenomena encompass such elements as stress, phonation type, voice timbre, and prosody or intonation, all of which may have effects across multiple segments.[62]

Consonants and vowel segments combine to form syllables, which in turn combine to form utterances; these can be distinguished phonetically as the space between two inhalations. Acoustically, these different segments are characterized by different formant structures, that are visible in a spectrogram of the recorded sound wave. Formants are the amplitude peaks in the frequency spectrum of a specific sound.[62][63]

The human vocal tract

Spectrogram of American English vowels [i, u, ɑ] showing the formants f_1 and f_2

Real time MRI scan of a person speaking in Mandarin Chinese

Vowels are those sounds that have no audible friction caused by the narrowing or obstruction of some part of the upper vocal tract. They vary in quality according to the degree of lip aperture and the placement of the tongue within the oral cavity.[62] Vowels are called
close when the lips are relatively closed, as in the pronunciation of the vowel [i] (English "ee"), or *open* when the lips are relatively open, as in the vowel [a] (English "ah"). If the tongue is located towards the back of the mouth, the quality changes, creating vowels such as [u] (English "oo"). The quality also changes depending on whether the lips are rounded as opposed to unrounded, creating distinctions such as that between [i] (unrounded front vowel such as English "ee") and [y] (rounded front vowel such as German "ü").[64]

Consonants are those sounds that have audible friction or closure at some point within the upper vocal tract. Consonant sounds vary by place of articulation, i.e. the place in the vocal tract where the airflow is obstructed, commonly at the lips, teeth, alveolar ridge, palate, velum, uvula, or glottis. Each place of articulation produces a different set of consonant sounds, which are further distinguished by manner of articulation, or the kind of friction, whether full closure, in which case the consonant is called *occlusive* or
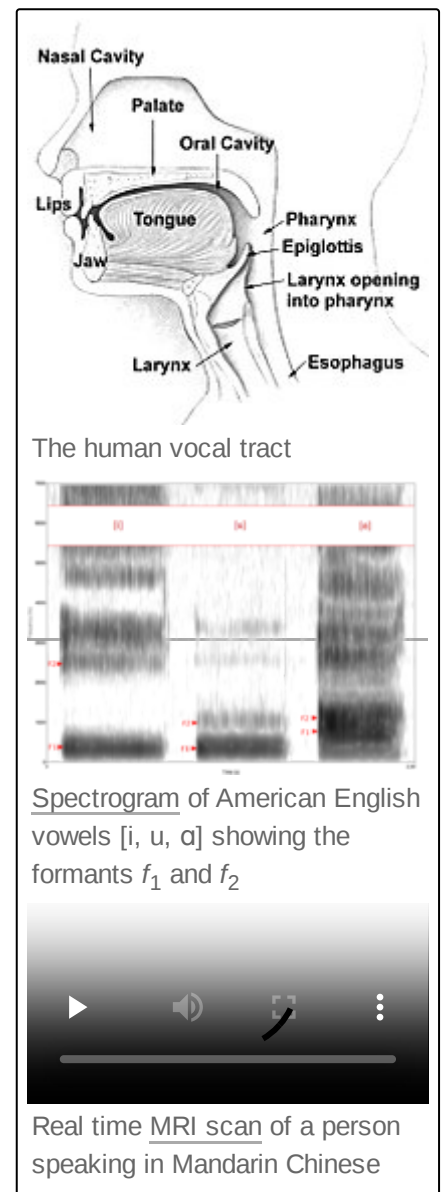
stop, or different degrees of aperture creating *fricatives* and *approximants*. Consonants can also be either *voiced or unvoiced*, depending on whether the vocal cords are set in vibration by airflow during the production of the sound. Voicing is what separates English [s] in *bus* (unvoiced sibilant) from [z] in *buzz* (voiced sibilant).[65]

Some speech sounds, both vowels and consonants, involve release of air flow through the nasal cavity, and these are called *nasals* or *nasalized* sounds. Other sounds are defined by the way the tongue moves within the mouth such as the l-sounds (called *laterals*, because the air flows along both sides of the tongue), and the r-sounds (called *rhotics*).[63]

By using these speech organs, humans can produce hundreds of distinct sounds: some appear very often in the world's languages, whereas others are much more common in certain language families, language areas, or even specific to a single language.[66]

Modality

Human languages display considerable plasticity[1] in their deployment of two fundamental modes: oral (speech and mouthing) and manual (sign and gesture).[note 3] For example, it is common for oral language to be accompanied by gesture, and for sign language to be accompanied by mouthing. In addition, some language communities use both modes to convey lexical or grammatical meaning, each mode complementing the other. Such bimodal use of language is especially common in genres such as story-telling (with Plains Indian Sign Language and Australian Aboriginal sign languages used alongside oral language, for example), but also occurs in mundane conversation. For instance, many Australian languages have a rich set of case suffixes that provide details about the instrument used to perform an action. Others lack such grammatical precision in the oral mode, but supplement it with gesture to convey that information in the sign mode. In Iwaidja, for example, 'he went out for fish using a torch' is spoken as simply "he-hunted fish torch", but the word for 'torch' is accompanied by a gesture indicating that it was held. In another example, the ritual language Damin had a heavily reduced oral vocabulary of only a few hundred words, each of which was very general in meaning, but which were supplemented by gesture for greater precision (e.g., the single word for fish, *l*i*, was accompanied by a gesture to indicate the kind of fish).[67]

Secondary modes of language, by which a fundamental mode is conveyed in a different medium, include writing (including braille), sign (in manually coded language), whistling and drumming. Tertiary modes – such as semaphore, Morse code and spelling alphabets – convey the secondary mode of writing in a different medium. For some extinct languages that are maintained for ritual or liturgical purposes, writing may be the primary mode, with speech secondary.

Structure

When described as a system of symbolic communication, language is traditionally seen as consisting of three parts: signs, meanings, and a code connecting signs with their meanings. The study of the process of semiosis, how signs and meanings are combined, used, and interpreted is called semiotics. Signs can be composed of sounds, gestures, letters, or symbols, depending on whether the language is spoken, signed, or written, and they can be combined into complex signs, such as words and phrases. When used in communication, a sign is encoded and transmitted by a sender through a channel to a receiver who decodes it.[68]

Some of the properties that define human language as opposed to other communication systems are: the arbitrariness of the linguistic sign, meaning that there is no predictable connection between a linguistic sign and its meaning; the duality of the linguistic system, meaning that linguistic structures are built by combining elements into larger structures that can be seen as layered, e.g. how sounds build words and

words build phrases; the discreteness of the elements of language, meaning that the elements out of which linguistic signs are constructed are discrete units, e.g. sounds and words, that can be distinguished from each other and rearranged in different patterns; and the productivity of the linguistic system, meaning that the finite number of linguistic elements can be combined into a theoretically infinite number of combinations.[68]

The rules by which signs can be combined to form words and phrases are called syntax or grammar. The meaning that is connected to individual signs, morphemes, words, phrases, and texts is called semantics.[69] The division of language into separate but connected systems of sign and meaning goes back to the first linguistic studies of de Saussure and is now used in almost all branches of linguistics.[70]

Ancient Tamil inscription at Thanjavur

Semantics

Languages express meaning by relating a sign form to a meaning, or its content. Sign forms must be something that can be perceived, for example, in sounds, images, or gestures, and then related to a specific meaning by social convention. Because the basic relation of meaning for most linguistic signs is based on social convention, linguistic signs can be considered arbitrary, in the sense that the convention is established socially and historically, rather than by means of a natural relation between a specific sign form and its meaning.[17]

Thus, languages must have a vocabulary of signs related to specific meaning. The English sign "dog" denotes, for example, a member of the species *Canis familiaris*. In a language, the array of arbitrary signs connected to specific meanings is called the lexicon, and a single sign connected to a meaning is called a lexeme. Not all meanings in a language are represented by single words. Often, semantic concepts are embedded in the morphology or syntax of the language in the form of grammatical categories.[71]

All languages contain the semantic structure of predication: a structure that predicates a property, state, or action. Traditionally, semantics has been understood to be the study of how speakers and interpreters assign truth values to statements, so that meaning is understood to be the process by which a predicate can be said to be true or false about an entity, e.g. "[x [is y]]" or "[x [does y]]". Recently, this model of semantics has been complemented with more dynamic models of meaning that incorporate shared knowledge about the context in which a sign is interpreted into the production of meaning. Such models of meaning are explored in the field of pragmatics.[71]

Sounds and symbols

Depending on modality, language structure can be based on systems of sounds (speech), gestures (sign languages), or graphic or tactile symbols (writing). The ways in which languages use sounds or signs to construct meaning are studied in phonology.[72]

Sounds as part of a linguistic system are called phonemes.[73] Phonemes are abstract units of sound, defined as the smallest units in a language that can serve to distinguish between the meaning of a pair of minimally different words, a so-called minimal pair. In English, for example, the words *bat* [bæt] and *pat* [pʰæt] form a minimal pair, in which the distinction between /b/ and /p/ differentiates the two words, which have different meanings. However, each language contrasts sounds in different ways. For example, in a language that does not distinguish between voiced and unvoiced consonants, the sounds [p] and [b] (if they

both occur) could be considered a single phoneme, and consequently, the two pronunciations would have the same meaning. Similarly, the English language does not distinguish phonemically between aspirated and non-aspirated pronunciations of consonants, as many other languages like Korean and Hindi do: the unaspirated /p/ in *spin* [spɪn] and the aspirated /p/ in *pin* [pʰɪn] are considered to be merely different ways of pronouncing the same phoneme (such variants of a single phoneme are called allophones), whereas in Mandarin Chinese, the same difference in pronunciation distinguishes between the words [pʰá] 'crouch' and [pá] 'eight' (the accent above the á means that the vowel is pronounced with a high tone).[74]

All spoken languages have phonemes of at least two different categories, vowels and consonants, that can be combined to form syllables.[62] As well as segments such as consonants and vowels, some languages also use sound in other ways to convey meaning. Many languages, for example, use stress, pitch, duration, and tone to distinguish meaning. Because these phenomena operate outside of the level of single segments, they are called suprasegmental.[75] Some languages have only a few phonemes, for example, Rotokas and Pirahã language with 11 and 10 phonemes respectively, whereas languages like Taa may have as many as 141 phonemes.[74] In sign languages, the equivalent to phonemes (formerly called cheremes) 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.[76][77][78]

Writing systems represent language using visual symbols, which may or may not correspond to the sounds of spoken language. The Latin alphabet (and those on which it is based or that have been derived from it) 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, such as the Inuktitut syllabary, each sign represents a whole syllable. In logographic scripts, each sign represents an entire word,[79] and will generally bear no relation to the sound of that word in spoken language.

Because all languages have a very large number of words, no purely 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 as the Latin script or right to left as the Arabic script), 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.[80]
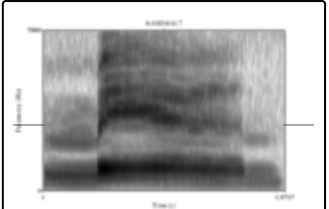
A spectrogram showing the sound of the spoken English word "man", which is written phonetically as [mæn]. In flowing speech, there is no clear division between segments, only a smooth transition as the vocal apparatus moves.

The syllable "wi" in the Hangul script

The sign for "wi" in Korean Sign Language (see Korean manual alphabet)

In order to represent the sounds of the world's languages in writing, linguists have developed the International Phonetic Alphabet, designed to represent all of the discrete sounds that are known to contribute to meaning in human languages.[81]

Grammar

Grammar is the study of how meaningful elements called *morphemes* within a language can be combined into utterances. Morphemes can either be *free* or *bound*. If they are free to be moved around within an utterance, they are usually called *words*, and if they are bound to other words or morphemes, they are called affixes. The way in which meaningful elements can be combined within a language is governed by rules. The study of the rules for the internal structure of words are called morphology. The rules of the internal structure of phrases and sentences are called *syntax*.[82]

Grammatical categories

Grammar can be described as a system of categories and a set of rules that determine how categories combine to form different aspects of meaning.[83] Languages differ widely in whether they are encoded through the use of categories or lexical units. However, several categories are so common as to be nearly universal. Such universal categories include the encoding of the grammatical relations of participants and predicates by grammatically distinguishing between their relations to a predicate, the encoding of temporal and spatial relations on predicates, and a system of grammatical person governing reference to and distinction between speakers and addressees and those about whom they are speaking.[84]

Word classes

Languages organize their parts of speech into classes according to their functions and positions relative to other parts. All languages, for instance, make a basic distinction between a group of words that prototypically denotes things and concepts and a group of words that prototypically denotes actions and events. The first group, which includes English words such as "dog" and "song", are usually called nouns. The second, which includes "think" and "sing", are called verbs. Another common category is the adjective: words that describe properties or qualities of nouns, such as "red" or "big". Word classes can be "open" if new words can continuously be added to the class, or relatively "closed" if there is a fixed number of words in a class. In English, the class of pronouns is closed, whereas the class of adjectives is open, since an infinite number of adjectives can be constructed from verbs (e.g. "saddened") or nouns (e.g. with the -like suffix, as in "noun-like"). In other languages such as Korean, the situation is the opposite, and new pronouns can be constructed, whereas the number of adjectives is fixed.[85]

Word classes also carry out differing functions in grammar. Prototypically, verbs are used to construct predicates, while nouns are used as arguments of predicates. In a sentence such as "Sally runs", the predicate is "runs", because it is the word that predicates a specific state about its argument "Sally". Some verbs such as "curse" can take two arguments, e.g. "Sally cursed John". A predicate that can only take a single argument is called *intransitive*, while a predicate that can take two arguments is called *transitive*.[86]

Many other word classes exist in different languages, such as conjunctions like "and" that serve to join two sentences, articles that introduce a noun, interjections such as "wow!", or ideophones like "splash" that mimic the sound of some event. Some languages have positionals that describe the spatial position of an event or entity. Many languages have classifiers that identify countable nouns as belonging to a particular type or having a particular shape. For instance, in Japanese, the general noun classifier for humans is *nin* (人), and it is used for counting humans, whatever they are called:[87]

san-nin no gakusei (三人の学生) lit. "3 human-classifier of student" – three students

For trees, it would be:

san-bon no ki (三本の木) lit. "3 classifier-for-long-objects of tree" – three trees

Morphology

In linguistics, the study of the internal structure of complex words and the processes by which words are formed is called morphology. In most languages, it is possible to construct complex words that are built of several morphemes. For instance, the English word "unexpected" can be analyzed as being composed of the three morphemes "un-", "expect" and "-ed".[88]

Morphemes can be classified according to whether they are independent morphemes, so-called roots, or whether they can only co-occur attached to other morphemes. These bound morphemes or affixes can be classified according to their position in relation to the root: *prefixes* precede the root, suffixes follow the root, and infixes are inserted in the middle of a root. Affixes serve to modify or elaborate the meaning of the root. Some languages change the meaning of words by changing the phonological structure of a word, for example, the English word "run", which in the past tense is "ran". This process is called *ablaut*. Furthermore, morphology distinguishes between the process of inflection, which modifies or elaborates on a word, and the process of derivation, which creates a new word from an existing one. In English, the verb "sing" has the inflectional forms "singing" and "sung", which are both verbs, and the derivational form "singer", which is a noun derived from the verb with the agentive suffix "-er".[89]

Languages differ widely in how much they rely on morphological processes of word formation. In some languages, for example, Chinese, there are no morphological processes, and all grammatical information is encoded syntactically by forming strings of single words. This type of morpho-syntax is often called isolating, or analytic, because there is almost a full correspondence between a single word and a single aspect of meaning. Most languages have words consisting of several morphemes, but they vary in the degree to which morphemes are discrete units. In many languages, notably in most Indo-European languages, single morphemes may have several distinct meanings that cannot be analyzed into smaller segments. For example, in Latin, the word *bonus*, or "good", consists of the root *bon-*, meaning "good", and the suffix *-us*, which indicates masculine gender, singular number, and nominative case. These languages are called *fusional languages*, because several meanings may be fused into a single morpheme. The opposite of fusional languages are agglutinative languages which construct words by stringing morphemes together in chains, but with each morpheme as a discrete semantic unit. An example of such a language is Turkish, where for example, the word *evlerinizden*, or "from your houses", consists of the morphemes, *ev-ler-iniz-den* with the meanings *house-plural-your-from*. The languages that rely on morphology to the greatest extent are traditionally called polysynthetic languages. They may express the equivalent of an entire English sentence in a single word. For example, in Persian the single word *nafahmidamesh* means *I didn't understand it* consisting of morphemes *na-fahm-id-am-esh* with the meanings, "negation.understand.past.I.it". As another example with more complexity, in the Yupik word *tuntussuqatarniksatengqiggtuq*, which means "He had not yet said again that he was going to hunt reindeer", the word consists of the morphemes *tuntu-ssur-qatar-ni-ksaite-ngqiggte-uq* with the meanings, "reindeer-hunt-future-say-negation-again-third.person.singular.indicative", and except for the morpheme *tuntu* ("reindeer") none of the other morphemes can appear in isolation.[90]

Many languages use morphology to cross-reference words within a sentence. This is sometimes called *agreement*. For example, in many Indo-European languages, adjectives must cross-reference the noun they modify in terms of number, case, and gender, so that the Latin adjective *bonus*, or "good", is inflected to agree with a noun that is masculine gender, singular number, and nominative case. In many polysynthetic

languages, verbs cross-reference their subjects and objects. In these types of languages, a single verb may include information that would require an entire sentence in English. For example, in the Basque phrase *ikusi nauzu*, or "you saw me", the past tense auxiliary verb *n-au-zu* (similar to English "do") agrees with both the subject (you) expressed by the *n-* prefix, and with the object (me) expressed by the *– zu* suffix. The sentence could be directly transliterated as "see you-did-me"[91]

Syntax

Another way in which languages convey meaning is through the order of words within a sentence. The grammatical rules for how to produce new sentences from words that are already known is called syntax. The syntactical rules of a language determine why a sentence in English such as "I love you" is meaningful, but "*love you I" is not.[note 4] Syntactical rules determine how word order and sentence structure is constrained, and how those constraints contribute to meaning.[93] For example, in English, the two sentences "the slaves were cursing the master" and "the master was cursing the slaves" mean different things, because the role of the grammatical subject is encoded by the noun being in front of the verb, and the role of object is encoded by the noun appearing after the verb. Conversely, in Latin, both *Dominus servos*

Basic constituent structure analysis of a sentence:

In addition to word classes, a sentence can be analyzed in terms of grammatical functions: "The cat" is the subject of the phrase, "on the mat" is a locative phrase, and "sat" is the core of the predicate.

vituperabat and *Servos vituperabat dominus* mean "the master was reprimanding the slaves", because *servos*, or "slaves", is in the accusative case, showing that they are the grammatical object of the sentence, and *dominus*, or "master", is in the nominative case, showing that he is the subject.[94]

Latin uses morphology to express the distinction between subject and object, whereas English uses word order. Another example of how syntactic rules contribute to meaning is the rule of inverse word order in questions, which exists in many languages. This rule explains why when in English, the phrase "John is talking to Lucy" is turned into a question, it becomes "Who is John talking to?", and not "John is talking to who?". The latter example may be used as a way of placing special emphasis on "who", thereby slightly altering the meaning of the question. Syntax also includes the rules for how complex sentences are structured by grouping words together in units, called phrases, that can occupy different places in a larger syntactic structure. Sentences can be described as consisting of phrases connected in a tree structure, connecting the phrases to each other at different levels.[95] To the right is a graphic representation of the syntactic analysis of the English sentence "the cat sat on the mat". The sentence is analyzed as being constituted by a noun phrase, a verb, and a prepositional phrase; the prepositional phrase is further divided into a preposition and a noun phrase, and the noun phrases consist of an article and a noun.[96]

The reason sentences can be seen as being composed of phrases is because each phrase would be moved around as a single element if syntactic operations were carried out. For example, "the cat" is one phrase, and "on the mat" is another, because they would be treated as single units if a decision was made to emphasize the location by moving forward the prepositional phrase: "[And] on the mat, the cat sat".[96]
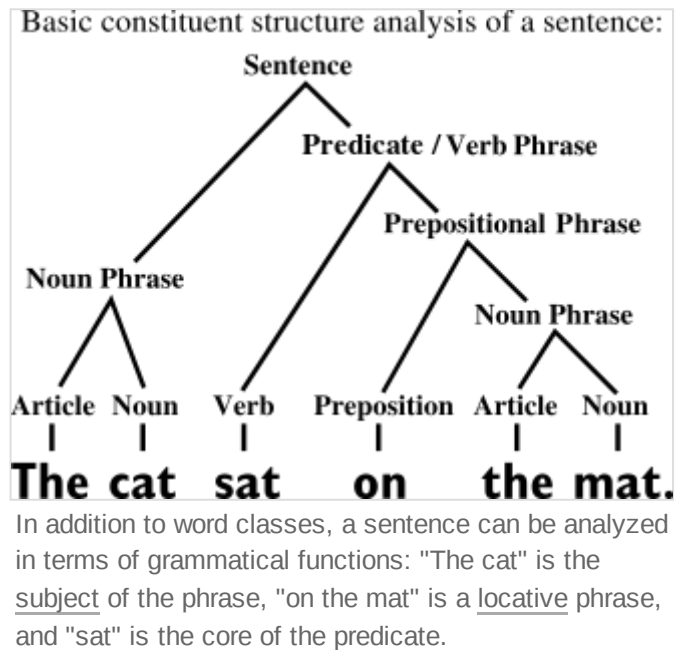
There are many different formalist and functionalist frameworks that propose theories for describing syntactic structures, based on different assumptions about what language is and how it should be described. Each of them would analyze a sentence such as this in a different manner.[22]

Typology and universals

Languages can be classified in relation to their grammatical types. Languages that belong to different families nonetheless often have features in common, and these shared features tend to correlate.[97] For example, languages can be classified on the basis of their basic word order, the relative order of the verb, and its constituents in a normal indicative sentence. In English, the basic order is SVO (subject–verb–object): "The snake(S) bit(V) the man(O)", whereas for example, the corresponding sentence in the Australian language Gamilaraay would be ḏuyugu ṉama ḏayn yiːy (snake man bit), SOV.[98] Word order type is relevant as a typological parameter, because basic word order type corresponds with other syntactic parameters, such as the relative order of nouns and adjectives, or of the use of prepositions or postpositions. Such correlations are called implicational universals.[99] For example, most (but not all) languages that are of the SOV type have postpositions rather than prepositions, and have adjectives before nouns.[100]

All languages structure sentences into Subject, Verb, and Object, but languages differ in the way they classify the relations between actors and actions. English uses the nominative-accusative word typology: in English transitive clauses, the subjects of both intransitive sentences ("I run") and transitive sentences ("I love you") are treated in the same way, shown here by the nominative pronoun *I*. Some languages, called ergative, Gamilaraay among them, distinguish instead between Agents and Patients. In ergative languages, the single participant in an intransitive sentence, such as "I run", is treated the same as the patient in a transitive sentence, giving the equivalent of "me run". Only in transitive sentences would the equivalent of the pronoun "I" be used.[98] In this way the semantic roles can map onto the grammatical relations in different ways, grouping an intransitive subject either with Agents (accusative type) or Patients (ergative type) or even making each of the three roles differently, which is called the tripartite type.[101]

The shared features of languages which belong to the same typological class type may have arisen completely independently. Their co-occurrence might be due to universal laws governing the structure of natural languages, "language universals", or they might be the result of languages evolving convergent solutions to the recurring communicative problems that humans use language to solve.[23]

Social contexts of use and transmission

While humans have the ability to learn any language, they only do so if they grow up in an environment in which language exists and is used by others. Language is therefore dependent on communities of speakers in which children learn language from their elders and peers and themselves transmit language to their own children. Languages are used by those who speak them to communicate and to solve a plethora of social tasks. Many aspects of language use can be seen to be adapted specifically to these purposes.[23] Owing to the way in which language is transmitted between generations and within communities, language perpetually changes, diversifying into new languages or converging due to language contact. The process is similar to the process of evolution, where the process of descent with modification leads to the formation of a phylogenetic tree.[102]

Wall of Love on Montmartre in Paris: "I love you" in 250 languages, by calligraphist Fédéric Baron and artist Claire Kito (2000)

However, languages differ from biological organisms in that they readily incorporate elements from other languages through the process of diffusion, as speakers of different languages come into contact. Humans also frequently speak more than one language, acquiring their first language or languages as children, or learning new languages as they grow up. Because of the increased language contact in the globalizing world, many small languages are becoming endangered as their speakers shift to other languages that afford the possibility to participate in larger and more influential speech communities.[6]

Usage and meaning

When studying the way in which words and signs are used, it is often the case that words have different meanings, depending on the social context of use. An important example of this is the process called deixis, which describes the way in which certain words refer to entities through their relation between a specific point in time and space when the word is uttered. Such words are, for example, the word, "I" (which designates the person speaking), "now" (which designates the moment of speaking), and "here" (which designates the position of speaking). Signs also change their meanings over time, as the conventions governing their usage gradually change. The study of how the meaning of linguistic expressions changes depending on context is called pragmatics. Deixis is an important part of the way that we use language to point out entities in the world.[103] Pragmatics is concerned with the ways in which language use is patterned and how these patterns contribute to meaning. For example, in all languages, linguistic expressions can be used not just to transmit information, but to perform actions. Certain actions are made only through language, but nonetheless have tangible effects, e.g. the act of "naming", which creates a new name for some entity, or the act of "pronouncing someone man and wife", which creates a social contract of marriage. These types of acts are called speech acts, although they can also be carried out through writing or hand signing.[104]

The form of linguistic expression often does not correspond to the meaning that it actually has in a social context. For example, if at a dinner table a person asks, "Can you reach the salt?", that is, in fact, not a question about the length of the arms of the one being addressed, but a request to pass the salt across the table. This meaning is implied by the context in which it is spoken; these kinds of effects of meaning are called conversational implicatures. These social rules for which ways of using language are considered appropriate in certain situations and how utterances are to be understood in relation to their context vary between communities, and learning them is a large part of acquiring communicative competence in a language.[105]

Acquisition

All healthy, normally developing human beings learn to use language. Children acquire the language or languages used around them: whichever languages they receive sufficient exposure to during childhood. The development is essentially the same for children acquiring sign or oral languages.[106] This learning process is referred to as first-language acquisition, since unlike many other kinds of learning, it requires no direct teaching or specialized study. In *The Descent of Man*, naturalist Charles Darwin called this process "an instinctive tendency to acquire an art".[15]

First language acquisition proceeds in a fairly regular sequence, though there is a wide degree of variation in the timing of particular stages among normally developing infants. Studies published in 2013 have indicated that unborn fetuses are capable of language acquisition to some degree.[107][108] From birth, newborns respond more readily to human speech than to other sounds. Around one month of age, babies appear to be able to distinguish between different speech sounds. Around six months of age, a child will begin babbling, producing the speech sounds or handshapes of the languages used around them. Words appear around the age of 12 to 18 months; the average vocabulary of an eighteen-month-old child is around

50 words. A child's first utterances are holophrases (literally "whole-sentences"), utterances that use just one word to communicate some idea. Several months after a child begins producing words, the child will produce two-word utterances, and within a few more months will begin to produce telegraphic speech, or short sentences that are less grammatically complex than adult speech, but that do show regular syntactic structure. From roughly the age of three to five years, a child's ability to speak or sign is refined to the point that it resembles adult language.[109][110]

A lesson at Kituwah Academy, a school where English and the Cherokee language are mediums of instruction

Acquisition of second and additional languages can come at any age, through exposure in daily life or courses. Children learning a second language are more likely to achieve native-like fluency than adults, but in general, it is very rare for someone speaking a second language to pass completely for a native speaker. An important difference between first language acquisition and additional language acquisition is that the process of additional language acquisition is influenced by languages that the learner already knows.[111]

Culture

Languages, understood as the particular set of speech norms of a particular community, are also a part of the larger culture of the community that speaks them. Languages differ not only in pronunciation, vocabulary, and grammar, but also through having different "cultures of speaking." Humans use language as a way of signalling identity with one cultural group as well as difference from others. Even among speakers of one language, several different ways of using the language exist, and each is used to signal affiliation with particular subgroups within a larger culture. Linguists and anthropologists, particularly sociolinguists, ethnolinguists, and linguistic anthropologists have specialized in studying how ways of speaking vary between speech communities.[112]

Arnold Lakhovsky, *The Conversation* (c. 1935)

Linguists use the term "varieties" to refer to the different ways of speaking a language. This term includes geographically or socioculturally defined dialects as well as the jargons or styles of subcultures. Linguistic anthropologists and sociologists of language define communicative style as the ways that language is used and understood within a particular culture.[113]

Because norms for language use are shared by members of a specific group, communicative style also becomes a way of displaying and constructing group identity. Linguistic differences may become salient markers of divisions between social groups, for example, speaking a language with a particular accent may imply membership of an ethnic minority or social class, one's area of origin, or status as a second language speaker. These kinds of differences are not part of the linguistic system, but are an important part of how people use language as a social tool for constructing groups.[114]
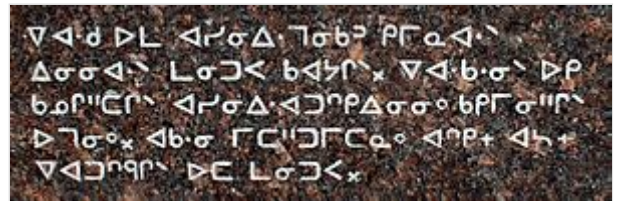
However, many languages also have grammatical conventions that signal the social position of the speaker in relation to others through the use of registers that are related to social hierarchies or divisions. In many languages, there are stylistic or even grammatical differences between the ways men and women speak, between age groups, or between social classes, just as some languages employ different words depending on who is listening. For example, in the Australian language Dyirbal, a married man must use a special set

of words to refer to everyday items when speaking in the presence of his mother-in-law.[115] Some cultures, for example, have elaborate systems of "social deixis", or systems of signalling social distance through linguistic means.[116] In English, social deixis is shown mostly through distinguishing between addressing some people by first name and others by surname, and in titles such as "Mrs.", "boy", "Doctor", or "Your Honor", but in other languages, such systems may be highly complex and codified in the entire grammar and vocabulary of the language. For instance, in languages of east Asia such as Thai, Burmese, and Javanese, different words are used according to whether a speaker is addressing someone of higher or lower rank than oneself in a ranking system with animals and children ranking the lowest and gods and members of royalty as the highest.[116]

Writing, literacy and technology

Throughout history a number of different ways of representing language in graphic media have been invented. These are called writing systems.

An inscription of Swampy Cree using Canadian Aboriginal syllabics, an abugida developed by Christian missionaries for Indigenous Canadian languages

The use of writing has made language even more useful to humans. It makes it possible to store large amounts of information outside of the human body and retrieve it again, and it allows communication across physical distances and timespans that would otherwise be impossible. Many languages conventionally employ different genres, styles, and registers in written and spoken language, and in some communities, writing traditionally takes place in an entirely different language than the one spoken. There is some evidence that the use of writing also has effects on the cognitive development of humans, perhaps because acquiring literacy generally requires explicit and formal education.[117]

The invention of the first writing systems is roughly contemporary with the beginning of the Bronze Age in the late 4th millennium BC. The Sumerian archaic cuneiform script and the Egyptian hieroglyphs are generally considered to be the earliest writing systems, both emerging out of their ancestral proto-literate symbol systems from 3400 to 3200 BC with the earliest coherent texts from about 2600 BC. It is generally agreed that Sumerian writing was an independent invention; however, it is debated whether Egyptian writing was developed completely independently of Sumerian, or was a case of cultural diffusion. A similar debate exists for the Chinese script, which developed around 1200 BC. The pre-Columbian Mesoamerican writing systems (including among others Olmec and Maya scripts) are generally believed to have had independent origins.[80]

Change

All languages change as speakers adopt or invent new ways of speaking and pass them on to other members of their speech community. Language change happens at all levels from the phonological level to the levels of vocabulary, morphology, syntax, and discourse. Even though language change is often initially evaluated negatively by speakers of the language who often consider changes to be "decay" or a sign of slipping norms of language usage, it is natural and inevitable.[118]

Changes may affect specific sounds or the entire phonological system. Sound change can consist of the replacement of one speech sound or phonetic feature by another, the complete loss of the affected sound, or even the introduction of a new sound in a place where there had been none. Sound changes can be *conditioned* in which case a sound is changed only if it occurs in the vicinity of certain other sounds. Sound

change is usually assumed to be *regular*, which means that it is expected to apply mechanically whenever its structural conditions are met, irrespective of any non-phonological factors. On the other hand, sound changes can sometimes be *sporadic*, affecting only one particular word or a few words, without any seeming regularity. Sometimes a simple change triggers a chain shift in which the entire phonological system is affected. This happened in the Germanic languages when the sound change known as Grimm's law affected all the stop consonants in the system. The original consonant *b^h became /b/ in the Germanic languages, the previous *b in turn became /p/, and the previous *p became /f/. The same process applied to all stop consonants and explains why Italic languages such as Latin have *p* in words like *pater* and *pisces*, whereas Germanic languages, like English, have *father* and *fish*.[119]

The first page of the poem *Beowulf*, written in Old English in the early medieval period (800–1100 AD). Although Old English is the direct ancestor of modern English, it is unintelligible to contemporary English speakers.

Another example is the Great Vowel Shift in English, which is the reason that the spelling of English vowels do not correspond well to their current pronunciation. This is because the vowel shift brought the already established orthography out of synchronization with pronunciation. Another source of sound change is the erosion of words as pronunciation gradually becomes increasingly indistinct and shortens words, leaving out syllables or sounds. This kind of change caused Latin *mea domina* to eventually become the French *madame* and American English *ma'am*.[120]

Change also happens in the grammar of languages as discourse patterns such as idioms or particular constructions become grammaticalized. This frequently happens when words or morphemes erode and the grammatical system is unconsciously rearranged to compensate for the lost element. For example, in some varieties of Caribbean Spanish the final /s/ has eroded away. Since Standard Spanish uses final /s/ in the morpheme marking the second person subject "you" in verbs, the Caribbean varieties now have to express the second person using the pronoun *tú*. This means that the sentence "what's your name" is *¿como te llamas?* ['komo te 'jamas] in Standard Spanish, but ['komo 'tu te 'jama] in Caribbean Spanish. The simple sound change has affected both morphology and syntax.[121] Another common cause of grammatical change is the gradual petrification of idioms into new grammatical forms, for example, the way the English "going to" construction lost its aspect of movement and in some varieties of English has almost become a full-fledged future tense (e.g. *I'm gonna*).

Language change may be motivated by "language internal" factors, such as changes in pronunciation motivated by certain sounds being difficult to distinguish aurally or to produce, or through patterns of change that cause some rare types of constructions to drift towards more common types.[122] Other causes of language change are social, such as when certain pronunciations become emblematic of membership in certain groups, such as social classes, or with ideologies, and therefore are adopted by those who wish to identify with those groups or ideas. In this way, issues of identity and politics can have profound effects on language structure.[123]

Contact

One important source of language change is contact and resulting diffusion of linguistic traits between languages. Language contact occurs when speakers of two or more languages or varieties interact on a regular basis.[124] Multilingualism is likely to have been the norm throughout human history and most people in the modern world are multilingual. Before the rise of the concept of the ethno-national state, monolingualism was characteristic mainly of populations inhabiting small islands. But with the ideology
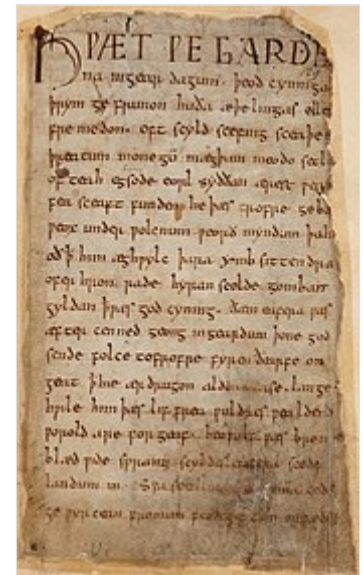
Multi-lingual sign outside the mayor's office in Novi Sad, written in the four official languages of the city: Serbian, Hungarian, Slovak, and Pannonian Rusyn

that made one people, one state, and one language the most desirable political arrangement, monolingualism started to spread throughout the world. Nonetheless, there are only 250 countries in the world corresponding to some 6000 languages, which means that most countries are multilingual and most languages therefore exist in close contact with other languages.[125]

When speakers of different languages interact closely, it is typical for their languages to influence each other. Through sustained language contact over long periods, linguistic traits diffuse between languages, and languages belonging to different families may converge to become more similar. In areas where many languages are in close contact, this may lead to the formation of language areas in which unrelated languages share a number of linguistic features. A number of such language areas have been documented, among them, the Balkan language area, the Mesoamerican language area, and the Ethiopian language area. Also, larger areas such as South Asia, Europe, and Southeast Asia have sometimes been considered language areas, because of widespread diffusion of specific areal features.[126][127]

Language contact may also lead to a variety of other linguistic phenomena, including language convergence, borrowing, and relexification (replacement of much of the native vocabulary with that of another language). In situations of extreme and sustained language contact, it may lead to the formation of new mixed languages that cannot be considered to belong to a single language family. One type of mixed language called pidgins occurs when adult speakers of two different languages interact on a regular basis, but in a situation where neither group learns to speak the language of the other group fluently. In such a case, they will often construct a communication form that has traits of both languages, but which has a simplified grammatical and phonological structure. The language comes to contain mostly the grammatical and phonological categories that exist in both languages. Pidgin languages are defined by not having any native speakers, but only being spoken by people who have another language as their first

Multilingualism is also common in the Indian Republic. The signboard is displayed in the Imphal International Airport in Meitei, Hindi and English, some of the official languages of the Indian Republic.

language. But if a Pidgin language becomes the main language of a speech community, then eventually children will grow up learning the pidgin as their first language. As the generation of child learners grow up, the pidgin will often be seen to change its structure and acquire a greater degree of complexity. This type of language is generally called a creole language. An example of such mixed languages is Tok Pisin, the official language of Papua New-Guinea, which originally arose as a Pidgin based on English and Austronesian languages; others are Kreyòl ayisyen, the French-based creole language spoken in Haiti, and Michif, a mixed language of Canada, based on the Native American language Cree and French.[128]

Linguistic diversity

SIL Ethnologue defines a "living language" as "one that has at least one speaker for whom it is their first language". The exact number of known living languages varies from 6,000 to 7,000, depending on the precision of one's definition of "language", and in particular, on how one defines the distinction between a "language" and a "dialect". As of 2016, *Ethnologue* cataloged 7,097 living human languages.[131] The

Ethnologue establishes linguistic groups based on studies of mutual intelligibility, and therefore often includes more categories than more conservative classifications. For example, the Danish language that most scholars consider a single language with several dialects is classified as two distinct languages (Danish and Jutish) by the *Ethnologue*.[129]

According to the *Ethnologue*, 389 languages (nearly 6%) have more than a million speakers. These languages together account for 94% of the world's population, whereas 94% of the world's languages account for the remaining 6% of the global population.

Language	Native speakers (millions)[129]
Mandarin	848
Spanish	329 [note 5]
English	328
Portuguese	250
Arabic	221
Hindi	182
Bengali	181
Russian	144
Japanese	122
Javanese	84.3

Languages and dialects

There is no clear distinction between a language and a dialect, notwithstanding a famous aphorism attributed to linguist Max Weinreich that "a language is a dialect with an army and navy".[132] For example, national boundaries frequently override linguistic difference in determining whether two linguistic varieties are languages or dialects. Hakka, Cantonese and Mandarin are, for example, often classified as "dialects" of Chinese, even though they are more different from each other than Swedish is from Norwegian. Before the Yugoslav Wars, Serbo-Croatian was generally considered a single language with two normative variants, but due to sociopolitical reasons, Croatian and Serbian are now often treated as separate languages and employ different writing systems. In other words, the distinction may hinge on political considerations as much as on cultural differences as on distinctive writing systems or the degree of mutual intelligibility.[133] The latter is, in fact, a rather unreliable criterion to discriminate languages and dialects. Pluricentric languages, which are languages with more than one standard variety, are a case in point. Standard American English and Standard RP (English) English, for instance, may in some areas be more different than languages with names, e.g. Swedish and Norwegian. A complex social process of "language making"[134] underlies these assignments of status and in some cases even linguistic experts may not agree (e.g. the One Standard German Axiom). The language making process is dynamic and subject to change over time.

Language families of the world
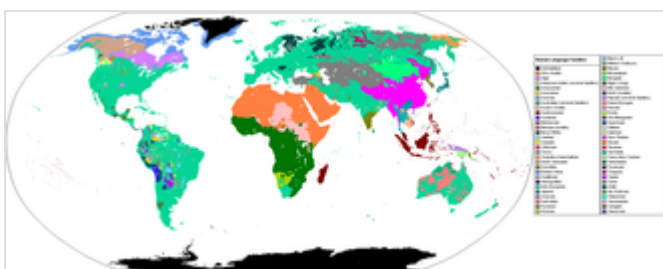
Principal language families of the world (and in some cases geographic groups of families). For greater detail, see *Distribution of languages in the world*.

The world's languages can be grouped into language families consisting of languages that can be shown to have common ancestry. Linguists recognize many hundreds of language families, although some of them can possibly be grouped into larger units as more evidence becomes available and in-depth studies are carried out. At present, there are also dozens of language isolates: languages that cannot be shown to be related to any other languages in the world. Among them are Basque, spoken in Europe, Zuni of New Mexico, Purépecha of Mexico, Ainu of Japan, Burushaski of Pakistan, and many others.[135]

The language family of the world that has the most speakers is the Indo-European languages, spoken by 46% of the world's population.[136] This family includes major world languages like English, Spanish, French, German, Russian, and Hindustani (Hindi/Urdu). The Indo-European family spread first through

hypothesized Indo-European migrations that would have taken place some time in the period c. 8000–1500 BCE,[137] and subsequently through much later European colonial expansion, which brought the Indo-European languages to a politically and often numerically dominant position in the Americas and much of Africa. The Sino-Tibetan languages are spoken by 20%[136] of the world's population and include many of the languages of East Asia, including Hakka, Mandarin Chinese, Cantonese, and hundreds of smaller languages.[138]

Africa is home to a large number of language families, the largest of which is the Niger-Congo language family, which includes such languages as Swahili, Shona, and Yoruba. Speakers of the Niger-Congo languages account for 6.9% of the world's population.[136] A similar number of people speak the Afroasiatic languages, which include the populous Semitic languages such as Arabic, Hebrew language, and the languages of the Sahara region, such as the Berber languages and Hausa.[138]

The Austronesian languages are spoken by 5.5% of the world's population and stretch from Madagascar to maritime Southeast Asia all the way to Oceania.[136] It includes such languages as Malagasy, Māori, Samoan, and many of the indigenous languages of Indonesia and Taiwan. The Austronesian languages are considered to have originated in Taiwan around 3000 BC and spread through the Oceanic region through island-hopping, based on an advanced nautical technology. Other populous language families are the Dravidian languages of South Asia (among them Kannada, Tamil, and Telugu), the Turkic languages of Central Asia (such as Turkish), the Austroasiatic (among them Khmer), and Tai–Kadai languages of Southeast Asia (including Thai).[138]

The areas of the world in which there is the greatest linguistic diversity, such as the Americas, Papua New Guinea, West Africa, and South-Asia, contain hundreds of small language families. These areas together account for the majority of the world's languages, though not the majority of speakers. In the Americas, some of the largest language families include the Quechumaran, Arawak, and Tupi-Guarani families of South America, the Uto-Aztecan, Oto-Manguean, and Mayan of Mesoamerica, and the Na-Dene, Iroquoian, and Algonquian language families of North America. In Australia, most indigenous languages belong to the Pama-Nyungan family, whereas New Guinea is home to a large number of small families and isolates, as well as a number of Austronesian languages.[135] Due to its remoteness and geographical fragmentation, Papua New Guinea emerges in fact as the leading location worldwide for both species (8% of world total) and linguistic richness – with 830 living tongues (12% of world total).[139]

Language endangerment

Language endangerment occurs when a language is at risk of falling out of use as its speakers die out or shift to speaking another language. Language loss occurs when the language has no more native speakers, and becomes a *dead language*. If eventually no one speaks the language at all, it becomes an *extinct language*. While languages have always gone extinct throughout human history, they have been disappearing at an accelerated rate in the 20th and 21st centuries due to the processes of globalization and neo-colonialism, where the economically powerful languages dominate other languages.[6]
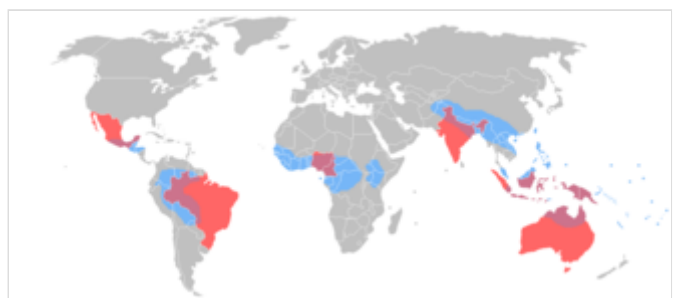
🟥 Together, these eight countries contain more than 50% of the world's languages.

🟦 These areas are the most linguistically diverse in the world, and the locations of most of the world's endangered languages.

The more commonly spoken languages dominate the less commonly spoken languages, so the less commonly spoken languages eventually disappear from populations. Of the between 6,000[5] and 7,000 languages spoken as of 2010, between 50 and 90% of those are expected to have become extinct by the year 2100.[6] The top 20 languages, those spoken by more than 50 million speakers each, are spoken by 50% of the world's population, whereas many of the other languages are spoken by small communities, most of them with less than 10,000 speakers.[6]

The United Nations Educational, Scientific and Cultural Organization (UNESCO) operates with five levels of language endangerment: "safe", "vulnerable" (not spoken by children outside the home), "definitely endangered" (not spoken by children), "severely endangered" (only spoken by the oldest generations), and "critically endangered" (spoken by few members of the oldest generation, often semi-speakers). Notwithstanding claims that the world would be better off if most adopted a single common *lingua franca*, such as English or Esperanto, there is a consensus that the loss of languages harms the cultural diversity of the world. It is a

UNESCO's five levels of language endangerment

common belief, going back to the biblical narrative of the tower of Babel in the Old Testament, that linguistic diversity causes political conflict,[34] but this is contradicted by the fact that many of the world's major episodes of violence have taken place in situations with low linguistic diversity, such as the Yugoslav and American Civil War, or the genocide of Rwanda, whereas many of the most stable political units have been highly multilingual.[140]

Many projects aim to prevent or slow this loss by revitalizing endangered languages and promoting education and literacy in minority languages. Across the world, many countries have enacted specific legislation to protect and stabilize the language of indigenous speech communities. A minority of linguists have argued that language loss is a natural process that should not be counteracted, and that documenting endangered languages for posterity is sufficient.[141]

The University of Waikato are using the Welsh language as a model for their Māori language revitalisation programme as they deem Welsh to be the world's leading example for the survival of languages.[142][143] In 2019 a Hawaiian TV company Oiwi visited a Welsh language centre in Nant Gwrtheyrn, North Wales to help find ways of preserving their Ōlelo Hawaiʻi language.[144]

See also

Language portal

Linguistics portal

Society portal

Religion portal
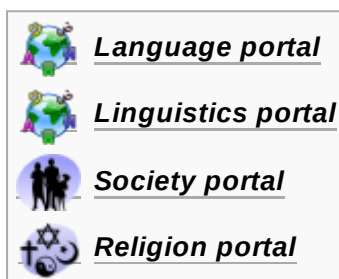
- Father Tongue hypothesis

- Human communication
  - Attitude (psychology)
  - Body language (approachable)
  - Humor
  - Listening
  - Reading
  - Speaking
  - Social skills
- International auxiliary language
- Linguistic rights
- List of language regulators
- Lists of languages
- List of official languages
- Outline of linguistics
- Problem of religious language
- Psycholinguistics
- Speech–language pathology

Notes

1. The gorilla Koko reportedly used as many as 1000 words in American Sign Language, and understands 2000 words of spoken English. There are some doubts about whether her use of signs is based on complex understanding or simple conditioning.[31]
2. "Functional grammar analyzes grammatical structure, as do formal and structural grammar; but it also analyzes the entire communicative situation: the purpose of the speech event, its participants, its discourse context. Functionalists maintain that the communicative situation motivates, constrains, explains, or otherwise determines grammatical structure, and that a structural or formal approach is not merely limited to an artificially restricted data base, but is inadequate even as a structural account. Functional grammar, then, differs from formal and structural grammar in that it purports not to model but to explain; and the explanation is grounded in the communicative situation".[53]
3. While sign is usually a visual medium, there is also tactile signing; and while oral speech is usually an aural medium, there is also lipreading and tadoma.
4. The prefixed asterisk * conventionally indicates that the sentence is ungrammatical, i.e. syntactically incorrect.[92]
5. *Ethnologue's* figure is based on numbers from before 1995. A more recent figure is 420 million.[130]

References

1. Nicholas Evans & Stephen Levinson (2009) 'The Myth of Language Universals: Language Diversity and Its Importance for Cognitive Science'. *Behavioral and Brain Sciences* 32, 429–492.

2. Kamusella, Tomasz (2016). "The History of the Normative Opposition of 'Language versus Dialect': From Its Graeco-Latin Origin to Central Europe's Ethnolinguistic Nation-States" (http://ispan.waw.pl/journals/index.php/ch/article/download/ch.2016.011/2342). *Colloquia Humanistica*. **5** (5): 189–198. doi:10.11649/ch.2016.011 (https://doi.org/10.11649%2Fch.2016.011). Archived (https://web.archive.org/web/20200226064031/https://ispan.waw.pl/journals/index.php/ch/article/download/ch.2016.011/2342) from the original on 26 February 2020. Retrieved 9 February 2020.

3. Tomasello (1996)

4. Hauser, Chomsky & Fitch (2002)

5. Moseley (2010): "Statistics (http://www.unesco.org/culture/languages-atlas/index.php) Archived (https://web.archive.org/web/20140812000456/http://www.unesco.org/culture/languages-atlas/index.php) 12 August 2014 at the Wayback Machine"

6. Austin & Sallabank (2011)

7. Graddol, David (27 February 2004). "The Future of Language". *Science*. **303** (5662): 1329–1331. Bibcode:2004Sci...303.1329G (https://ui.adsabs.harvard.edu/abs/2004Sci...303.1329G). doi:10.1126/science.1096546 (https://doi.org/10.1126%2Fscience.1096546). ISSN 0036-8075 (https://www.worldcat.org/issn/0036-8075). PMID 14988552 (https://pubmed.ncbi.nlm.nih.gov/14988552). S2CID 35904484 (https://api.semanticscholar.org/CorpusID:35904484).

8. "language". *The American Heritage Dictionary of the English Language* (3rd ed.). Boston: Houghton Mifflin Company. 1992.

9. Lyons (1981:2)

10. Lyons (1981:1–8)

11. Trask (2007:129–131)

12. Bett 2010.

13. Devitt & Sterelny 1999.

14. Hauser & Fitch (2003)

15. Pinker (1994)

16. Trask 2007, p. 93.

17. Saussure (1983)

18. Campbell (2001:96)

19. Trask 2007, p. 130.

20. Chomsky (1957)

21. Trask (2007:93, 130)

22. Newmeyer (1998:3–6)

23. Evans & Levinson (2009)

24. Van Valin (2001)

25. Nerlich 2010, p. 192.

26. Hockett, Charles F. (1966). "The Problem of Universals in Language" (https://web.archive.org/web/20121110023354/http://www.ling.ohio-state.edu/~swinters/371/designfeatures.html). Archived from the original (http://www.ling.ohio-state.edu/~swinters/371/designfeatures.html) on 10 November 2012. Retrieved 11 May 2013.

27. Hockett (1960); Deacon (1997)

28. Trask (1999:1–5)

29. Engesser, Sabrina; Crane, Jodie M. S.; Savage, James L.; Russel, Andrew F.; Townsend, Simon W. (29 June 2015). "Experimental Evidence for Phonemic Contrasts in a Nonhuman Vocal System" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4488142). *PLOS Biology*. **13** (6): e1002171. doi:10.1371/journal.pbio.1002171 (https://doi.org/10.1371%2Fjournal.pbio.1002171). PMC 4488142 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4488142). PMID 26121619 (https://pubmed.ncbi.nlm.nih.gov/26121619).

30. Engesser, Sabrina; Ridley, Amanda R.; Townsend, Simon W. (20 July 2017). "Element repetition rates encode functionally distinct information in pied babbler 'clucks' and 'purrs' " (http://wrap.warwick.ac.uk/89798/7/WRAP-element-repetition-rates-encode-functionally-Townsend-2017.pdf) (PDF). *Animal Cognition*. **20** (5): 953–960. doi:10.1007/s10071-017-1114-6 (https://doi.org/10.1007%2Fs10071-017-1114-6). PMID 28730513 (https://pubmed.ncbi.nlm.nih.gov/28730513). S2CID 21470061 (https://api.semanticscholar.org/CorpusID:21470061). Archived (https://web.archive.org/web/20190428232241/http://wrap.warwick.ac.uk/89798/7/WRAP-element-repetition-rates-encode-functionally-Townsend-2017.pdf) (PDF) from the original on 28 April 2019. Retrieved 9 November 2018.

31. Candland (1993)
32. Deacon (1997)
33. Trask (2007:165–166)
34. Haugen (1973)
35. Shanahan, D. (2011). Language, feeling, and the brain: The evocative vector. Transaction Publishers.
36. Ulbaek (1998)
37. Tomasello (2008)
38. Fitch 2010, pp. 466–507.
39. Anderson (2012:107)
40. Anderson (2012:104)
41. Fitch 2010, pp. 250–292.
42. Clark, Gary; Henneberg, Maciej (2017). "Ardipithecus ramidus and the evolution of language and singing: An early origin for hominin vocal capability". *HOMO*. **68** (2): 101–121. doi:10.1016/j.jchb.2017.03.001 (https://doi.org/10.1016%2Fj.jchb.2017.03.001). PMID 28363458 (https://pubmed.ncbi.nlm.nih.gov/28363458).
43. Foley 1997, pp. 70–74.
44. Fitch 2010, pp. 292–293.
45. Chomsky 1972, p. 86.
46. Chomsky 2000, p. 4.
47. Newmeyer (2005)
48. Trask (2007)
49. Campbell (2001:82–83)
50. Bloomfield 1914, p. 310
51. Clarke (1990:143–144)
52. Foley (1997:82–83)
53. Nichols (1984)
54. Croft & Cruse (2004:1–4)
55. Trask (1999:11–14, 105–113)
56. Fisher, Lai & Monaco (2003)
57. Lesser (1989:205–206)
58. Trask (1999:105–107)
59. Trask (1999:108)
60. Sandler & Lillo-Martin (2001:554)
61. MacMahon (1989:2)
62. MacMahon (1989:3)
63. International Phonetic Association (1999:3–8)
64. MacMahon (1989:11–15)
65. MacMahon (1989:6–11)
66. Ladefoged & Maddieson (1996)
67. Nicholas Evans (2017) Listening Here: Ngûrrahmalkwonawoniyan. *Humanities Australia*, Journal of the Australian Academy of the Humanities, vol 8, p 39.
68. Lyons (1981:17–24)
69. Trask (1999:35)
70. Lyons (1981:218–224)
71. Levinson (1983)
72. Goldsmith (1995)
73. International Phonetic Association (1999:27)
74. Trask (2007:214)
75. International Phonetic Association (1999:4)
76. Stokoe, William C. (1960). *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf,* Studies in linguistics: Occasional papers (No. 8). Buffalo: Dept. of Anthropology and Linguistics, University of Buffalo.
77. Stokoe, William C.; Dorothy C. Casterline; Carl G. Croneberg (1965). *A dictionary of American sign languages on linguistic principles.* Washington, D.C.: Gallaudet College Press
78. Sandler & Lillo-Martin (2001:539–540)
79. Trask (2007:326)
80. Coulmas (2002)
81. Trask (2007:123)
82. Lyons (1981:103)
83. Allerton (1989)
84. Payne (1997)
85. Trask (2007:208)
86. Trask (2007:305)
87. Senft (2008)
88. Aronoff & Fudeman (2011:1–2)
89. Bauer (2003); Haspelmath (2002)
90. Payne (1997:28–29)
91. Trask (2007:11)

92. Denham, Kristin; Lobeck, Anne (2009). *Linguistics for Everyone: An Introduction* (https://books.google.com/books?id=n0KVSvqvZKYC&pg=PA9). Cengage Learning. p. 9. ISBN 978-1-4130-1589-8. Archived (https://web.archive.org/web/20200522101450/https://books.google.com/books?id=n0KVSvqvZKYC&pg=PA9) from the original on 22 May 2020. Retrieved 16 April 2020. "An *ungrammatical* sentence is one that is impossible in a given language, one that a native speaker of that variety would never utter naturally. (Remember that ungrammatical sentences are marked with an asterisk, *."

93. Baker (2001:265)

94. Trask (2007:179)

95. Baker 2001, pp. 269–270.

96. Trask (2007:218–219)

97. Nichols (1992);Comrie (1989)

98. Croft (2001:340)

99. Greenberg (1966)

100. Comrie (2009:45); McMahon (1994:156)

101. Croft (2001:355)

102. Campbell (2004)

103. Levinson (1983:54–96)

104. Levinson (1983:226–278)

105. Levinson (1983:100–169)

106. Bonvillian, John D.; Michael D. Orlansky; Leslie Lazin Novack (December 1983). "Developmental milestones: Sign language acquisition and motor development". *Child Development*. **54** (6): 1435–1445. doi:10.2307/1129806 (https://doi.org/10.2307%2F1129806). JSTOR 1129806 (https://www.jstor.org/stable/1129806). PMID 6661942 (https://pubmed.ncbi.nlm.nih.gov/6661942).

107. *Scientific American* (2015:24)

108. Skwarecki, Beth (26 August 2013). "Babies Learn to Recognize Words in the Womb" (https://web.archive.org/web/20151104124655/http://news.sciencemag.org/brain-behavior/2013/08/babies-learn-recognize-words-womb). Science. Archived from the original (https://www.science.org/content/article/babies-learn-recognize-words-womb) on 4 November 2015.

109. O'Grady, William; Cho, Sook Whan (2001). "First language acquisition". *Contemporary Linguistics: An Introduction* (4th ed.). Boston: Bedford St. Martin's.

110. Kennison (2013)

111. Macaro (2010:137–157)

112. Duranti (2003)

113. Foley (1997)

114. Agha (2006)

115. Dixon (1972:32–34)

116. Foley (1997:311–328)

117. Olson (1996)

118. Aitchison (2001); Trask (1999:70)

119. Clackson (2007:27–33)

120. Aitchison (2001:112)

121. Zentella (2002:178)

122. Labov (1994)

123. Labov (2001)

124. Thomason (2001:1)

125. Romaine (2001:513)

126. Campbell (2002)

127. Aikhenvald (2001)

128. Thomason & Kaufman (1988); Thomason (2001); Matras & Bakker (2003)

129. Lewis (2009)

130. "Primer estudio conjunto del Instituto Cervantes y el British Council sobre el peso internacional del español y del inglés" (http://www.cervantes.es/sobre_instituto_cervantes/prensa/2012/noticias/nota-londres-palabra-por-palabra.htm). Instituto Cervantes (www.cervantes.es). Archived (https://web.archive.org/web/20170916204118/http://www.cervantes.es/sobre_instituto_cervantes/prensa/2012/noticias/nota-londres-palabra-por-palabra.htm) from the original on 16 September 2017. Retrieved 17 September 2012.

131. "Ethnologue statistics" (http://www.ethnologue.com/statistics). *Summary by world area | Ethnologue*. SIL. Archived (https://web.archive.org/web/20160108221353/http://www.ethnologue.com/statistics) from the original on 8 January 2016. Retrieved 11 September 2014.

132. Rickerson, E.M. "What's the difference between dialect and language?" (https://web.archive.org/web/20101219070627/http://spinner.cofc.edu/linguist/archives/2005/08/whats_the_diffe.html?referrer=webcluster&). *The Five Minute Linguist*. College of Charleston. Archived from the original (http://spinner.cofc.edu/linguist/archives/2005/08/whats_the_diffe.html?referrer=webcluster&) on 19 December 2010. Retrieved 17 July 2011.

133. Lyons (1981:26)

134. Krämer, Philipp; Vogl, Ulrike; Kolehmainen, Leena; et al. (2022). "What is language making?" (https://doi.org/10.1515%2Fijsl-2021-0016). *International Journal of the Sociology of Language*. **2022** (274): 1–27. doi:10.1515/ijsl-2021-0016 (https://doi.org/10.1515%2Fijsl-2021-0016).

135. Katzner (1999)

136. Lewis (2009), "Summary by language family (http://www.ethnologue.com/statistics/family) Archived (https://web.archive.org/web/20160101054116/http://www.ethnologue.com/statistics/family) 1 January 2016 at the Wayback Machine"

137. Heggarty, P.; Anderson, C.; Scarborough, M.; King, B.; Bouckaert, R.; Jocz, L.; Kümmel, M.J.; Jügel, T.; Irslinger, B.; Pooth, R.; Liljegren, H (28 July 2023). "Language trees with sampled ancestors support a hybrid model for the origin of Indo-European languages". *Science*. **381** (6656): eabg0818. doi:10.1126/science.abg0818 (https://doi.org/10.1126%2Fscience.abg0818). hdl:10234/204329 (https://hdl.handle.net/10234%2F204329).

138. Comrie (2009); Brown & Ogilvie (2008)

139. Briand, Frederic (February 2013). "Silent Plains … the Fading Sounds of Native Languages" (https://www.researchgate.net/publication/235742552). *National Geographic*.

140. Austin & Sallabank (2011:10–11)

141. Ladefoged (1992)

142. "University of Waikato Launches a Strategic Partnership with Cardiff University in Wales" (https://www.scoop.co.nz/stories/ED2111/S00030/university-of-waikato-launches-a-strategic-partnership.htm) (Press release). University of Waikato. 10 November 2021. Archived (https://web.archive.org/web/20211124055642/https://www.scoop.co.nz/stories/ED2111/S00030/university-of-waikato-launches-a-strategic-partnership.htm) from the original on 24 November 2021. Retrieved 21 December 2021 – via Scoop News.

143. Rhiannon James (10 November 2021). "Council investing £6.4m in the future of the Welsh language" (https://nation.cymru/news/council-investing-6-4m-in-the-future-of-the-welsh-language/). *Nation Cymru*. Archived (https://web.archive.org/web/20211111205030/https://nation.cymru/news/council-investing-6-4m-in-the-future-of-the-welsh-language/) from the original on 11 November 2021. Retrieved 21 December 2021.

144. "Hawaiian TV company seeks help to promote language" (https://www.cambrian-news.co.uk/article.cfm?id=130057&headline=Hawaiian%20TV%20company%20seeks%20help%20to%20promote%20language&sectionIs=news&searchyear=2019). *Cambrian News*. 20 August 2019. Archived (https://web.archive.org/web/20211205155429/https://www.cambrian-news.co.uk/article.cfm?id=130057&headline=Hawaiian%20TV%20company%20seeks%20help%20to%20promote%20language&sectionIs=news&searchyear=2019) from the original on 5 December 2021. Retrieved 21 August 2021.

Works cited

Agha, Asif (2006). *Language and Social Relations*. Cambridge University Press.

Aikhenvald, Alexandra (2001). "Introduction". In Alexandra Y. Aikhenvald; R.M.W. Dixon (eds.). *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: Oxford University Press. pp. 1–26.

Aitchison, Jean (2001) [1981]. *Language Change: Progress or Decay?* (3rd ed.). Cambridge, New York, Melbourne: Cambridge University Press.

Allerton, D.J. (1989). "Language as Form and Pattern: Grammar and its Categories". In Collinge, N.E. (ed.). *An Encyclopedia of Language*. London:NewYork: Routledge.

Anderson, Stephen (2012). *Languages: A Very Short Introduction*. Oxford: Oxford University Press. ISBN 978-0-19-959059-9.

Aronoff, Mark; Fudeman, Kirsten (2011). *What is Morphology*. John Wiley & Sons.

Austin, Peter K; Sallabank, Julia (2011). "Introduction". In Austin, Peter K; Sallabank, Julia (eds.). *Cambridge Handbook of Endangered Languages*. Cambridge University Press. ISBN 978-0-521-88215-6.

Baker, Mark C. (2001). "Syntax". In Mark Aronoff; Janie Rees-Miller (eds.). *The Handbook of Linguistics*. Blackwell. pp. 265–295.

Bauer, Laurie (2003). *Introducing linguistic morphology* (2nd ed.). Washington, D.C.: Georgetown University Press. ISBN 978-0-87840-343-1.

Bett, R. (2010). "Plato and his Predecessors". In Alex Barber; Robert J Stainton (eds.). *Concise Encyclopedia of Philosophy of Language and Linguistics*. Elsevier. pp. 569–570.

Bloomfield, Leonard (1914). *An introduction to the study of language* (https://archive.org/details/introductiontos00bloo). New York: Henry Holt and Company.

Brown, Keith; Ogilvie, Sarah, eds. (2008). *Concise Encyclopedia of Languages of the World*. Elsevier Science. ISBN 978-0-08-087774-7.

Clackson, James (2007). *Indo-European Linguistics: An Introduction*. Cambridge University press.

Campbell, Lyle (2002). "Areal linguistics". In Bernard Comrie, Neil J. Smelser and Paul B. Balte (ed.). *International Encyclopedia of Social and Behavioral Sciences*. Oxford: Pergamon. pp. 729–733.

Campbell, Lyle (2004). *Historical Linguistics: an Introduction* (2nd ed.). Edinburgh and Cambridge, MA: Edinburgh University Press and MIT Press.

Campbell, Lyle (2001). "The History of Linguistics". In Mark Aronoff; Janie Rees-Miller (eds.). *The Handbook of Linguistics*. Blackwell. pp. 81–105.

Candland, Douglas Keith (1993). *Feral Children and Clever Animals: Reflections on Human Nature* (https://archive.org/details/feralchildrencle00cand). Oxford University Press US. pp. 293 (https://archive.org/details/feralchildrencle00cand/page/293)–301. ISBN 978-0-19-510284-0. "koko gorilla operant conditioning."

Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam (1972). *Language and Mind* (https://archive.org/details/languagemind00chom). Harcourt Brace Jovanovich. ISBN 978-0-15-549257-8.

Chomsky, Noam (2000). *The Architecture of Language*. Oxford: Oxford University Press.

Clarke, David S. (1990). *Sources of semiotic: readings with commentary from antiquity to the present*. Carbondale: Southern Illinois University Press.

Comrie, Bernard (1989). *Language universals and linguistic typology: Syntax and morphology* (2nd ed.). Oxford: Blackwell. ISBN 978-0-226-11433-0.

Comrie, Bernard, ed. (2009). *The World's Major Languages*. New York: Routledge. ISBN 978-0-415-35339-7.

Coulmas, Florian (2002). *Writing Systems: An Introduction to Their Linguistic Analysis*. Cambridge University Press.

Croft, William; Cruse, D. Alan (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Croft, William (2001). "Typology". In Mark Aronoff; Janie Rees-Miller (eds.). *The Handbook of Linguistics*. Blackwell. pp. 81–105.

Deacon, Terrence (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton & Company. ISBN 978-0-393-31754-1.

Devitt, Michael; Sterelny, Kim (1999). *Language and Reality: An Introduction to the Philosophy of Language*. Boston: MIT Press.

Dixon, Robert M.W. (1972). *The Dyirbal Language of North Queensland* (https://archive.org/details/dyirballanguageo00dixo). Cambridge: Cambridge University Press. ISBN 978-0-521-08510-6.

Duranti, Alessandro (2003). "Language as Culture in U.S. Anthropology: Three Paradigms". *Current Anthropology*. **44** (3): 323–348. doi:10.1086/368118 (https://doi.org/10.1086%2F368118). S2CID 148075449 (https://api.semanticscholar.org/CorpusID:148075449).

Evans, Nicholas; Levinson, Stephen C. (2009). "The myth of language universals: Language diversity and its importance for cognitive science" (https://doi.org/10.1017%2Fs0140525x0999094x). *Behavioral and Brain Sciences*. **32** (5): 429–492. doi:10.1017/s0140525x0999094x (https://doi.org/10.1017%2Fs0140525x0999094x). PMID 19857320 (https://pubmed.ncbi.nlm.nih.gov/19857320).

"First Impressions: We start to pick up words, food preferences and hand-eye coordination long before being born". *Scientific American*. Vol. 313, no. 1. July 2015. p. 24.

Fisher, Simon E.; Lai, Cecilia S.L.; Monaco, Anthony P. (2003). "Deciphering the Genetic Basis of Speech and Language Disorders". *Annual Review of Neuroscience*. **26**: 57–80. doi:10.1146/annurev.neuro.26.041002.131144 (https://doi.org/10.1146%2Fannurev.neuro.26.041002.131144). hdl:11858/00-001M-0000-0012-CB7D-6 (https://hdl.handle.net/11858%2F00-001M-0000-0012-CB7D-6). PMID 12524432 (https://pubmed.ncbi.nlm.nih.gov/12524432).

Fitch, W. Tecumseh (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.

Foley, William A. (1997). *Anthropological Linguistics: An Introduction*. Blackwell.

Goldsmith, John A (1995). "Phonological Theory". In John A. Goldsmith (ed.). *The Handbook of Phonological Theory*. Blackwell Handbooks in Linguistics. Blackwell Publishers. ISBN 978-1-4051-5768-1.

Greenberg, Joseph (1966). *Language Universals: With Special Reference to Feature Hierarchies*. The Hague: Mouton & Co.

Haspelmath, Martin (2002). *Understanding morphology*. London: Arnold, Oxford University Press. (pbk)

Haugen, Einar (1973). "The Curse of Babel". *Daedalus*. **102** (3, Language as a Human Problem): 47–57.

Hauser, Marc D.; Chomsky, Noam; Fitch, W. Tecumseh (2002). "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?". *Science*. **298** (5598): 1569–1579. doi:10.1126/science.298.5598.1569 (https://doi.org/10.1126%2Fscience.298.5598.1569). PMID 12446899 (https://pubmed.ncbi.nlm.nih.gov/12446899).

Hauser, Marc D.; Fitch, W. Tecumseh (2003). "What are the uniquely human components of the language faculty?" (https://web.archive.org/web/20140808045239/http://mind.cog.jhu.edu/faculty/smolensky/050.341-641/Readings/Hauser%20%26%20Fitch%2003%20Evolution%20of%20Lg.pdf) (PDF). In M.H. Christiansen; S. Kirby (eds.). *Language Evolution: The States of the Art*. Oxford University Press. Archived from the original (http://mind.cog.jhu.edu/faculty/smolensky/050.341-641/Readings/Hauser%20&%20Fitch%2003%20Evolution%20of%20Lg.pdf) (PDF) on 8 August 2014.

Hockett, Charles F. (1960). "Logical considerations in the study of animal communication". In W. E. Lanyon; W. N. Tavolga (eds.). *Animals sounds and animal communication*. pp. 392–430.

International Phonetic Association (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press. ISBN 978-0-521-65236-0.

Katzner, Kenneth (1999). *The Languages of the World*. New York: Routledge.

Kennison, Shelia (2013). *Introduction to Language Development*. SAGE.

Labov, William (1994). Principles of Linguistic Change *vol.I* Internal Factors. Blackwell.

Labov, William (2001). Principles of Linguistic Change *vol.II* Social Factors. Blackwell.

Ladefoged, Peter (1992). "Another view of endangered languages". *Language*. **68** (4): 809–811. doi:10.1353/lan.1992.0013 (https://doi.org/10.1353%2Flan.1992.0013). S2CID 144984900 (https://api.semanticscholar.org/CorpusID:144984900).

Ladefoged, Peter; Maddieson, Ian (1996). *The sounds of the world's languages*. Oxford: Blackwell. pp. 329–330. ISBN 978-0-631-19815-4.

Lesser, Ruth (1989). "Language in the Brain: Neurolinguistics". In Collinge, N.E. (ed.). *An Encyclopedia of Language*. London:NewYork: Routledge.

Levinson, Stephen C. (1983). *Pragmatics* (https://archive.org/details/pragmatics00levi). Cambridge: Cambridge University Press.

Lewis, M. Paul, ed. (2009). "Ethnologue: Languages of the World" (http://www.ethnologue.com/ethno_docs/distribution.asp?by=size) (16th ed.). Dallas, Tex.: SIL International. Archived (https://web.archive.org/web/20120404174730/http://www.ethnologue.com/ethno_docs/distribution.asp?by=size) from the original on 4 April 2012. Retrieved 25 August 2012.

Lyons, John (1981). *Language and Linguistics* (https://archive.org/details/languagelinguist0000lyon). Cambridge University Press. ISBN 978-0-521-29775-2.

Macaro, Ernesto, ed. (2010). *Continuum companion to second language acquisition*. London: Continuum. pp. 137–157. ISBN 978-1-4411-9922-5.

McMahon, April M.S. (1994). *Understanding Language Change*. Cambridge University Press. ISBN 978-0-521-44119-3.

MacMahon, M.K.C. (1989). "Language as available sound:Phonetics". In Collinge, N.E. (ed.). *An Encyclopedia of Language*. London:NewYork: Routledge.

Matras, Yaron; Bakker, Peter, eds. (2003). *The Mixed Language Debate: Theoretical and Empirical Advances*. Berlin: Walter de Gruyter. ISBN 978-3-11-017776-3.

Moseley, Christopher, ed. (2010). *Atlas of the World's Languages in Danger* (http://www.unesco.org/culture/en/endangeredlanguages/atlas) (3rd ed.). Paris: UNESCO Publishing. Archived (https://web.archive.org/web/20140820135739/http://www.unesco.org/culture/en/endangeredlanguages/atlas/) from the original on 20 August 2014. Retrieved 7 August 2012.

Nerlich, Brigitte (2010). "History of pragmatics". In Cummings, Louise (ed.). *The Pragmatics Encyclopedia*. London/New York: Routledge. pp. 192–193.

Newmeyer, Frederick J. (2005). *The History of Linguistics* (http://www.linguisticsociety.org/resource/history-modern-linguistics). Linguistic Society of America. ISBN 978-0-415-11553-7. Archived (https://web.archive.org/web/20141231094048/http://www.linguisticsociety.org/resource/history-modern-linguistics) from the original on 31 December 2014. Retrieved 18 September 2014.

Newmeyer, Frederick J. (1998). *Language Form and Language Function* (https://web.archive.org/web/20111216100353/http://www.isc.cnrs.fr/FN_chapter1.pdf) (PDF). Cambridge, MA: MIT Press. Archived from the original (http://www.isc.cnrs.fr/FN_chapter1.pdf) (PDF) on 16 December 2011. Retrieved 29 August 2012.

Nichols, Johanna (1992). *Linguistic diversity in space and time*. Chicago: University of Chicago Press. ISBN 978-0-226-58057-9.

Nichols, Johanna (1984). "Functional Theories of Grammar". *Annual Review of Anthropology*. **13**: 97–117. doi:10.1146/annurev.an.13.100184.000525 (https://doi.org/10.1146%2Fannurev.an.13.100184.000525).

Olson, David R. (1996). "Language and Literacy: what writing does to Language and Mind". *Annual Review of Applied Linguistics*. **16**: 3–13. doi:10.1017/S0267190500001392 (https://doi.org/10.1017%2FS0267190500001392). S2CID 145677801 (https://api.semanticscholar.org/CorpusID:145677801).

Payne, Thomas Edward (1997). *Describing morphosyntax: a guide for field linguists* (https://books.google.com/books?id=LC3DfjWfCiwC&q=%22perfect+aspect%22+%22perfective+aspect%22&pg=PA239). Cambridge University Press. pp. 238–241. ISBN 978-0-521-58805-8. Archived (https://web.archive.org/web/20210331142644/https://books.google.com/books?id=LC3DfjWfCiwC&q=%22perfect+aspect%22+%22perfective+aspect%22&pg=PA239) from the original on 31 March 2021. Retrieved 14 November 2020.

Pinker, Steven (1994). The Language Instinct: How the Mind Creates Language. Perennial.

Romaine, Suzanne (2001). "Multilingualism". In Mark Aronoff; Janie Rees-Miller (eds.). *The Handbook of Linguistics*. Blackwell. pp. 512–533.

Sandler, Wendy; Lillo-Martin, Diane (2001). "Natural Sign Languages". In Mark Aronoff; Janie Rees-Miller (eds.). *The Handbook of Linguistics*. Blackwell. pp. 533–563.

Saussure, Ferdinand de (1983) [1913]. Bally, Charles; Sechehaye, Albert (eds.). *Course in General Linguistics*. Translated by Roy Harris. La Salle, Illinois: Open Court. ISBN 978-0-8126-9023-1.

Senft, Gunter, ed. (2008). *Systems of Nominal Classification*. Cambridge University Press. ISBN 978-0-521-06523-8.

Tomasello, Michael (1996). "The Cultural Roots of Language". In B. Velichkovsky and D. Rumbaugh (ed.). *Communicating Meaning: The Evolution and Development of Language*. Psychology Press. pp. 275–308. ISBN 978-0-8058-2118-5.

Tomasello, Michael (2008). *Origin of Human Communication*. MIT Press.

Thomason, Sarah G.; Kaufman, Terrence (1988). *Language Contact, Creolization and Genetic Linguistics*. University of California Press.

Thomason, Sarah G. (2001). *Language Contact – An Introduction*. Edinburgh University Press.

Trask, Robert Lawrence (1999). *Language: The Basics* (2nd ed.). Psychology Press.

Trask, Robert Lawrence (2007). Stockwell, Peter (ed.). *Language and Linguistics: The Key Concepts* (2nd ed.). Routledge.

Ulbaek, Ib (1998). "The Origin of Language and Cognition". In J. R. Hurford & C. Knight (ed.). *Approaches to the evolution of language*. Cambridge University Press. pp. 30–43.

Van Valin, Robert D. Jr. (2001). "Functional Linguistics". In Mark Aronoff; Janie Rees-Miller (eds.). *The Handbook of Linguistics*. Blackwell. pp. 319–337.

Zentella, Ana Celia (2002). "Spanish in New York". In García, Ofelia; Fishman, Joshua (eds.). *The Multilingual Apple: Languages in New York City*. Walter de Gruyter.

Further reading

- Crystal, David (1997). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.
- Cysouw, Michael; Good, Jeff (2013). "Languoid, doculect and glossonym: Formalizing the notion 'language' " (http://scholarspace.manoa.hawaii.edu/handle/10125/4606). *Language Documentation and Conservation*. **7**: 331–359. hdl:10125/4606 (https://hdl.handle.net/10125%2F4606).

- Swadesh, Morris (1934). "The phonemic principle". *Language*. **10** (2): 117–129. doi:10.2307/409603 (https://doi.org/10.2307%2F409603). JSTOR 409603 (https://www.jstor.org/stable/409603).

External links

- World Atlas of Language Structures: a large database of structural (phonological, grammatical, lexical) properties of languages (http://wals.info/)
- Ethnologue: Languages of the World (http://www.ethnologue.com/) is a comprehensive catalog of all of the world's known living languages

-