MARIA
STAMBOLIEVA

TEN LECTURES IN LINGUISTICS

MARIA STAMBOLIEVA

TEN LECTURES IN LINGUISTICS

Maria Stambolieva

TEN LECTURES IN LINGUISTICS

NEW BULGARIAN UNIVERSITY

Ten Lectures in Linguistics
By Maria Stambolieva

Reviewed by Prof. D. Vesselinov and Dr. M. Hajikoteva

Cover designed and text typeset by MT Studio, Sofia

21 Montevideo St.
1618 Sofia, Bulgaria

ISBN 978-619-233-139-9 (online)

Contents

Lecture 1

Communication and Language

The birth of the science of language

Through the ages, almost all human civilisations have been interested in the origins, nature and functions of human language.

The earliest linguistic tradition that we know of, that of **Babylonia**, dates back **4000 years**. It emerged in an effort to **preserve the knowledge of Sumerian**, the prestigious language of legal texts and religion, following its replacement with Akkadian as the language of everyday communication. This linguistic tradition existed for over 2500 years!

During the **first millennium BC**, emerging **changes in Sanskrit** – the sacred language of religious texts – necessitated the description of **the Vedas**. **Pāṇini**, an Indian grammarian from the 4th century BC, created a remarkable grammar that remained unsurpassed for a long time to come.

Early linguistics emerges as a need to preserve the memory of a changing or disappearing language. All natural languages change over time. A number of linguists have noticed the following, among other, ongoing changes in the English language:

1. *The rise of the present participle (*-ing *form) at the expense of the* to-*infinitive after some verbs (e.g.* start, begin, like, love, hate, fear*).*
2. *An increased use of the progressive, especially with the Passive Voice* (The house is being built) *or with the link verb* to be, *followed by an adjective* (I am being serious)
3. *A decreased use of* shall *and* ought.
4. *The rise of the* get-*passive.*

Can you think of similar changes observable in English, in another foreign language you have studied, or in your mother tongue?

The grammatical tradition of **Greece** also emerged in response to changes in the language and the necessity to explain **the language of Homer**. It then took a different turn, however – towards more general questions (such as the classification of words) and **philosophical issues** – the origin of language, **the relation between language and thought or form and meaning**.

The following is an abridged text from the Stanford Encyclopedia of Philosophy. Which of the two interlocutors would you side with? Can you provide arguments for your position?

One of the most famous dialogues of Plato, *Cratylus,* revolves around the topic of 'correctness of names' – clearly, an important issue in the philosophical disputes of the late fifth century BC. In this dialogue, Socrates has two interlocutors, Hermogenes and Cratylus, who represent opposed views on the issue. These views have come to be known as 'conventionalism' and 'naturalism,' respectively. Hermogenes, a conventionalist, holds that nothing but convention determines which word is used to designate an object. Cratylus, a linguistic naturalist, holds that names cannot be arbitrarily chosen because they belong *naturally* to their specific objects. If you try to speak of something with any name other than its natural name, you are simply failing to refer to it at all.

Dionysius Thrax (around 100 BC) wrote a grammar of Greek with a focus on phonetics and morphology. This is the **oldest surviving grammar of a European language**. **Apollonius Dyscolus** (around 110–175 AD) was the author of the **first known work on syntax**.

The Romans, though following the Greek tradition, did not seem to be much interested in syntax. The grammars of Marcus Terentius **Varro** (116 BC – 27 BC), Aelius Donatus (4th C. AD) and Priscianus Caesariensis (a.k.a. **Priscian**, around 500 AD – whose work *Institutes of Grammar* was extremely influential during the Middle Ages), discussed **issues of morphology, with a special focus on parts of speech**.

The 12th century saw the rise of interest in **universalist aspects of grammar** – a line of thought to be taken up in the 13th century by **Roger Bacon** and, much later, by the American generativists.

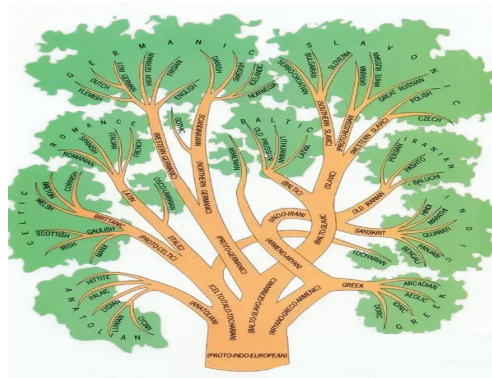
Which languages, according to you, were the observations of the universalists based on? From your experience of languages, can you think of arguments in favour of the postulation of universal rules of grammar?

With the rise of colonialism, from the 15th century on Europeans entered into contact with other continents, foreign cultures and languages. By and by, linguists came to realise, first, that **not all languages are related**, and that for some of them the categories of Latin cannot be applied; second, that **most European languages** are similar not only because of universalist principles of grammar but, simply, because they **are related**. The Danish linguist **Rasmus Rask** (1787–1832) and the German **August Schleicher** (1821–1868) were among the first to establish language relatedness based on methods of comparative analysis. **Jacob Grimm** (1785–1863) described in detail a **consonant shift**, known as **Grimm's Law**.

Grimm's Law

states that a chain of consonant changes took place from Proto-Indo-European to Proto-Germanic in the first millennium BC, which accounts for regular phonetic contrasts between otherwise related words in Germanic and other Indo-European languages – Cf: *frère* à *brother*, *dent* à *tooth*, etc.

A. Schleicher introduced the model of a common language family tree. Most European languages belong to language families which evolved from a common proto-language – **Proto-Indo-European (PIE).** PIE is a theoretical construct, a reconstruction based on the comparison of related languages and language subgroups. It was spoken during the Neolithic age, from around 4500 BC to around 2500 BC. The **comparative method of analysis** demonstrating genealogical relatedness is considered to be **the most outstanding achievement of 19th century linguistics**.



The languages of Europe

Although the European Union has 28 member states, its official languages are 24 – some languages are shared by two or more states. Three languages have semi-official status.

Most of the EU languages belong to three **Indo-European language groups**:

- Balto-Slavonic or Baltic (Latvian and Lithuanian) and Slavonic (Bulgarian, Croatian, Czech, Polish and Slovak). Lithuanian has the best preserved archaic structure of the mother Indo-European language; Bulgarian has the oldest literary tradition among the languages of the Slavonic group.
- Germanic (German, English, Dutch, Danish and Swedish). English is a language which has international status and is one of the three "procedural languages" of the European Commission.
- Romance languages (French, Italian, Spanish, Portuguese and Romanian and Catalan – the latter with semi-official status in the EU) are the heirs of Latin.
- Irish and Scottish Gaelic (another semi-official language) belong to the Gaelic group of Celtic languages.
- Greek, with the oldest literary tradition in the EU, is the only living language in its group.

Of the **non-Indo-European languages**, the largest group is Finno-Ugric, with three members – Finnish, Estonian and Hungarian. Maltese is a Semitic language. Finally, Basque (the third semi-official EU language) is unrelated to any other known language (it is *a language isolate*); it is the language of a pre-Indo-European population of the European Continent.

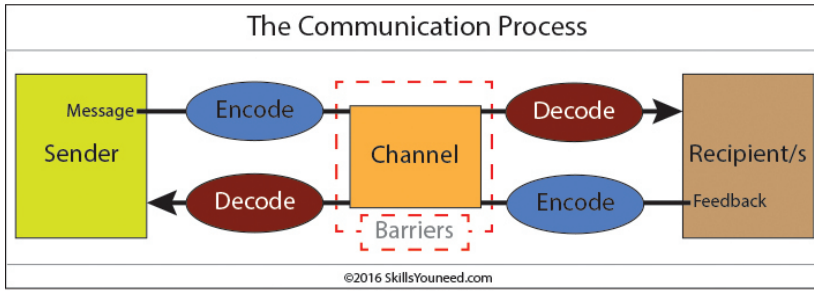
Find information about:

- *which subgroup of Germanic languages English belongs to;*
- *which language group your mother tongue belongs to;*
- *what other large families of languages exist;*
- *how many languages are spoken around the world.*

Most official EU languages are written in the Latin script. The two exceptions are Greek (written in the Greek script) and Bulgarian (written in the Cyrillic script).

Communication in the plant and animal world

Communication (from Latin *commūnicāre*, "to share") is the act of conveying intended meanings from one entity or group to another through the use of mutually understood signals (a common "code"). The sender conveys information to the receiver through a "communication channel":



To make sure that the message reaches its recipient successfully, the form of encoding should be appropriately chosen so as to overcome potential barriers.

Communication is not specific to human language: it is part of the life of all biological organisms. One way or another, **all living creatures communicate**.

E. O. Wilson[1] defines **biological communication** as "Actions on the part of one organism that alter the probability pattern of behavior in another organism in an adaptive fashion". **The channel of communication can be visual, auditory, tactile and haptic, olfactory, electromagnetic, or biochemical**. Some species make use of a number of channels. Elephants, for instance, communicate with every sense they have: touch, taste, smell, vision and hearing.

[1] Wilson, E. O., 1971. The Insect Societies. Belknap Press of Harvard University Press.

Did you know that?

Not only animals, plants too can communicate.

Shaunacy Ferro* relates five types of situations in which **plants use chemicals to communicate with each other or with insects**. Here is one:

"The wild tobacco plant can identify a hornworm caterpillar by its saliva. When attacked by this caterpillar, the tobacco plant emits a chemical signal that appeals to the insect's enemies. Within hours, caterpillar predators like the big-eyed bug show up, ideally driving the pest away".

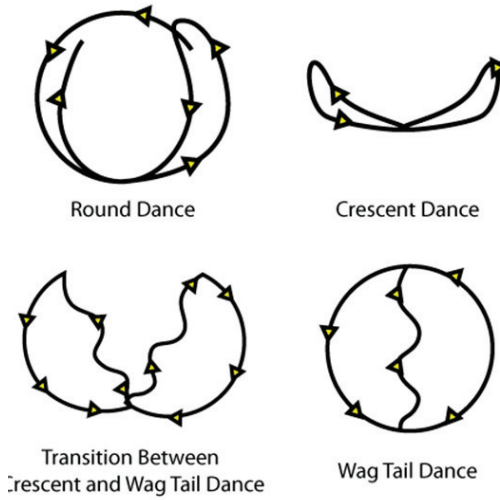
* *http://mentalfloss.com/article/66302/5-ways-plants-communicate*

Spiders, like plants, have mainly chemical communication but can also communicate with visual and acoustic (vibratory) signals. In the photograph below, you can see a jumping spider displaying rich colour and ornamentation during the courtship period.

Insects (arthropods) communicate using sounds – Cf. the chirping of crickets or the buzzing of cicadas.[2]

[2] Cf. e.g. Busnel 1963, Sebeok 1977, Pollack 2016.

The complex dances of honeybees[3] (round dance, crescent (or sickle) dance and tail wagging) have long attracted the attention of researchers.



The round dance indicates locations near the hive, the crescent (or sickle) dance – intermediate distance, the wag tail: a greater distance. The quality of the food source is indicated by the number of repetitions of the basic pattern. The angle made by the direction of the open end of the sickle in the sickle dance informs on direction.

In all the above cases, communication is innate. There are animal species, however, which use for communication both innate systems and systems of sounds which involve experience and learning. An example in point are bird songs.[4]

[3] Seeley, Thomas D. 1995. The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies. Cambridge, Mass.: Harvard University Press.

[4] Aitchison 1996: 7–9: *http://downloads.bbc.co.uk/rmhttp/radio4/transcripts/1996_reith1.pdf*

The song of the robin consists in alternations between high-pitched and low-pitched notes. These notes can be used in different alternations, sometimes quite complex, which allows for some creativity; the message itself, however, remains fixed and restricted.

It has been asserted[5] that dolphins can understand gestural communication, that they can learn an artificial language, use its words to refer to objects and even make sense of artificial grammar.

Much research has been carried out on the communication of great apes, who are genetically closest to humans (we share over 99 per cent of genetic material with chimpanzees). While apes cannot speak, they can communicate with other apes and can be taught to use signs in establishing cause-result relations for the purpose of communication with humans. It is not clear to what extent this use can be creative and whether it involves a symbolic use of signs.

Two American researchers, Allen and Beatrice Gardner, taught American Sign Language to Washoe, a chimpanzee brought up in a human environment. Find out more about their experiment and its results.

Some of the important questions which researchers in language and communication seek to answer are:

- In what respects and to what extent do the communication abilities of non-human animals differ from those of humans?
- Are the abilities that underlie human language general or species-specific?
- What sets human language apart from other forms of communication?
- What are the origins of human language?

[5] *http://acp.eugraph.com/cetaceans/index.html#*

The origins of human language

The earliest registered traces of conscious human-like activity, such as the creation and use of stone tools, date back 3 million years ago. For most researchers, this is a first sign of conscious thought. Our ancestors *Homo sapiens*, however, emerged much later, around 150 000 years ago.

Between 150 000 and 70 000 years ago, as a result of a micro mutation, the volume of the human brain increased considerably. This was probably the decisive factor for the emergence of human language as a species-specific system of communication displaying the design features defined by Charles Hockett (Cf. Lecture 2).

A topic of debate for anthropology and linguistics is the name *Homo sapiens*.

Human language probably developed simultaneously with the appearance of our species, of which it is the most defining feature. Some authors have therefore suggested that *Homo loquens* ("Speaking man") is a more appropriate name than *Homo sapiens*:

> *In any case,* Homo sapiens, *"the wise human," should perhaps really be called* Homo loquens, *"the speaking human", because language and humans are everywhere found together, whereas wisdom among humans is much more selectively distributed.*
>
> Edward Vajda. Retrieved from:
> *http://pandora.cii.wwu.edu/vajda/ling201/test1materials/origin_of_language.htm*

There are two **major hypotheses about the emergence of human language**.

According to **creationists**, a superior species or a deity created men and gave them language. Creationists assert that there is a qualitative difference between the cognitive, mental and language capacities of men and other species – a gap that evolution cannot account for.

Evolutionists, on the other hand, view language as a higher stage of the general cognitive capacity. A number of evolutionist theories have been put forward. To most of them critics have given somewhat sarcastic names.

According to the **ding-dong** and **bow-wow hypotheses**, language developed from the naming by imitation of features of objects and situations, e.g. Chinese *tun-tun*, "heart" (ding-dong hypothesis) or of animals, e.g. English *moo* (bow-wow hypothesis). This iconicity of naming is called **onomatopoeia**.

Can you think of other onomatopoeic words?

The **pooh-pooh hypothesis** holds that words came from exclamations such as *Ouch! Wow! Ha-ha*! and others. For supporters of the **ta-ta hypothesis**, speech developed as a vocal pantomime of gestures.

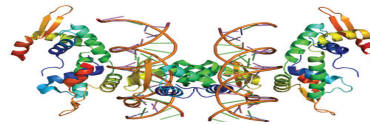
Animals can signal other members of their clan for danger, and early humans might have done the same. This forms the foundation of the **Warning hypothesis**. Early humans might also have developed specific sounds or chants during common activity – hence the **Yo-ho-ho hypothesis**.

As the early language of humans has left no material traces, none of these hypotheses can be supported by solid proof. Not surprisingly, in 1866 the French Academy of Sciences placed a ban on publications about the origin of human language. The topic, however, is too exciting to be dropped, and their ban did not last long.

Which of the above hypotheses would you side with? Find more information on its authors and arguments. Find information on other hypotheses about the origin of human language.

Modern science tells us that the proper development of speech and language is due to a protein, encoded in humans by the gene FOXP2. Mutations in this gene cause a number of severe disorders, including disorders in speech.

An interesting fact: FOXP2 is more active in females than in males – which would come to explain the greater ease with which females acquire and learn languages!



Quiz 1

1. **Linguistic tradition is believed to have begun:**
   a. 2500 years ago
   b. 3000 year ago
   c. 4000 years ago
   d. 5000 years ago.
2. **The first language to be described was:**
   a. Greek
   b. Latin
   c. Sanskrit
   d. Sumerian.
3. **The author of the first book on syntax was:**
   a. Plato
   b. Panini
   c. Apollonius Dyscolus
   d. Dionysius Thrax.
4. **Grimm's Law is about:**
   a. rules of language development
   b. language simplification
   c. consonant changes
   d. vowel changes.
5. **Which of these languages is not Indo-European?**
   a. Hungarian
   b. Welsh
   c. Croatian
   d. Portuguese.
6. **Biochemical communication is typical of:**
   a. plants
   b. birds
   c. caterpillars
   d. plants and animals
7. **The round dance of bees indicates:**
   a. the quality of the food source
   b. the direction of the source
   c. a great distance from the source
   d. a location near the hive.

Further reading

Anderson, Stephen R., 2016. Languages: A Very Short Introduction. Oxford: Oxford University Press.

Busnel, R.G., 1963. In "The Acoustic Behaviour of Animals", Chapter 4, Elsevier, Amsterdam.

Givon, Talmy, Malle, Bertram F., 2002. The Evolution of Language out of Pre-language. Amsterdam, Philadelphia: John Benjamins.

Hockett, Charles F., 1960. The Origin of Speech, *Scientific American* 203.

Hockett, Charles F., 1966. The Problem of Universals in Language. Retrieved from *http://www.ling.ohio-state.edu/~swinters/371/designfeatures.html*

MacNeilage, P., 2008. The Origin of Speech. Oxford: Oxford University Press.

Pollack, G.S. et al (eds.), 2016. Insect Hearing. Springer Handbook of Auditory Research. Springer International Publishing, Switzerland.

Sebeok, Thomas A. (ed.), 1977. How Animals Communicate. Bloomington: Indiana University Press.

Tomasello, M, 2008. Origins of Human Communication. Cambridge, MA, MIT Press.

Wilson, E.O., 1971. The Insect Societies. Belknap Press of Harvard University Press.

Zuberbuhler, Klaus. Primate Communication, the Knowledge Project. Retrieved 12 May 2013 from *http://www.nature.com/scitable/knowledge/library/primate-communication-67560503*

Lecture 2

Human Language

The communicative function of language

Language and human society are inseparable: **language exists wherever human societies exist**. In a society where humans interact and cooperate, one of the major functions of language is to provide a **means of communication** with fellow men. Communication is the major motivation for the use and development of language.

Which of the hypotheses on the origins of language listed in Lecture 1 takes into account human interaction and cooperation?

Which of these hypotheses can account for language development?

Michael Tomasello (born 18 January 1950**)**, an American developmental psychologist and evolutionary anthropologist, views the development of language as part and parcel of its communicative functions.

Humans alone, asserts Tomasello, share the **cognitive capacity of interacting with others socially and reading intentions**. As a result of their social environment, humans, according to him, are endowed with a number of species-specific cognitive capacities:

- the capacity to imitate others;
- the capacity to share attention with others;
- the capacity to understand the intentions of others.

Based on his study of primate cognition and child language acquisition, Tomasello sets out and defines **three distinct stages in the development of language structure**: a **grammar of wanting**, a **grammar of informing** and a **grammar of sharing and telling**.

The **first stage** of language development, which is characteristic of the signaling communication of primates and of infants under the age of two, is restricted to situations taking place "here and now", observable at the time of speech. It **expresses qualification** or, most frequently, **a wish or a demand**. Its linguistic form is a simple structure without syntactic shape: an object word and a quality word or action word, e.g. *Uncle bad* or *Bonny eat*.

The second and third stages are specific to human language: communication becomes more abstract, moves away from the "here and now" only and includes participants, objects and events which are removed in space and time.

At the **second stage**, the need to inform about events taking place elsewhere, at another time and with **additional participants** presupposes the introduction of "third party" actors and **markers of place and time**.

Can you think of examples of such "second stage" communication?
What could the markers of removed place and time be?

At the **third stage**, several situations can be linked to present a **coherent story**. Situations become clearly marked as **states, processes or events**, the role of the participants is better expressed, **anaphora** (the use of pronouns and other pro-forms) emerges. **Syntactic structuring** becomes important.

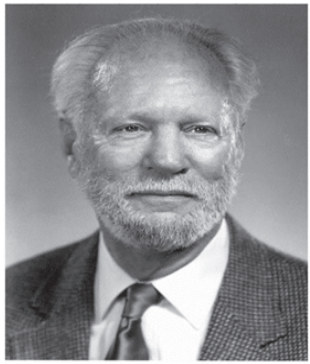
Why should communicating a more coherent story involve a more complex syntactic structure?

In 3 or 4 simple sentences, relate a recent past event. Then, make a note of the number of participants, situation types, anaphoric means and markers of time and space.

Recall the information on animal communication presented in Lecture 1. Does it support Tomasello's theory of language development?

The design features of human language

Charles Hockett (1916–2000), an influential American linguist working in the structuralist tradition, **defined 16 features characterising (human) language**, most of which set it apart from animal communication. According to Hockett, only communication systems manifesting *all* of these 16 features qualify as "language".[1]



The features set out by Hockett are:

1. Use of a **vocal-auditory channel** based on speaking and hearing. Because this feature did not take into account sign language, it was later modified to "tactile-visual" (as opposed to chemical-olfactory).

2. Broadcast transmission and **directional reception** – listeners are aware of the direction from which the message comes.

3. **Transitoriness** (rapid fading): signals, unless they are fixed (as in writing, quipu or other) only last a short time.

4. **Interchangeability**. Not all biological organisms can both send and receive signals. (An often cited example is the silkworm, the males of which can only receive the chemical signals emitted by females.) In human communication, utterances that are understood can also be produced.

5. **Total feedback** – the sender of a message also perceives the message: you can hear what you say.

6. **Specialisation**: the signal is deliberately produced for the purpose of communication; communication is not an additional side effect of some other behavior.

[1] Cf. Hockett, Charles F. (1960). The Origin of Speech, Scientific American 203. 1960.

7. **Semanticity**: the signals in the system of communication have a clear meaning.
8. **Arbitrariness**: the form of the signal is related to its meaning by convention, there is no logical relation between them.
9. **Displacement**: communication need not be about a situation that is sensorily accessible to the participants in the act – it can relate to things removed in space or time.
10. **Discreteness**: A message can be complex, made up of smaller, discrete parts. Changing one of the subparts can change the meaning of the message.
11. **Duality of patterning**: A large numbers of meaningful signals (e.g. words, phrases or sentences) can be produced from a small number of meaningless units (e.g. sounds and strings of sounds).
12. **Productivity**. A system of communication is productive if it is open-ended, if it allows the production of an unlimited number of different messages.
13. **Traditional (cultural) transmission** implies that, even if the communication system relies on an inborn capacity, it is acquired in a social setting.
14. **Learnability** is a feature of human language which allows a speaker of one language to learn to communicate in one or more additional languages.
15. **Prevarication** is the feature of a system of communication which allows the production of deliberately deceptive messages.
16. **Reflexiveness** allows communication about the system of communication itself.

Which of these design features, according to you, define all forms of communication?

Which features are unique to human language?

Which of Hockett's design features of human language form part of Tomasello's stages of language development?

The innate language acquisition device and Universal Grammar

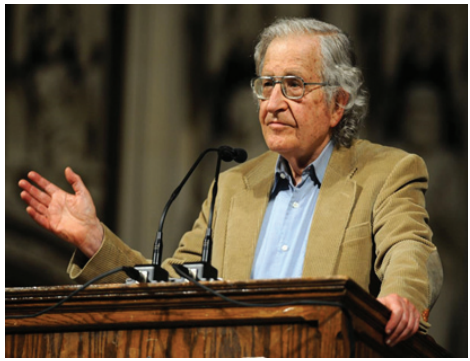
Not all linguists believe in a gradual evolutionary transition from animal to human communication. Probably the most notable opponent of this hypothesis is the American linguist **Noam Chomsky** (born 7 December 1928). Chomsky holds that human language capacity is an inborn characteristic of Homo sapiens as a species, which developed as a result of a micromutation some 60-70 000 years ago. (Cf. Lecture 1 on FOXP2.)

According to Chomsky, humans are genetically equipped with grammar – not in the form of the three grammars proposed by Tomasello but, rather, with a mechanism or genetic predisposition for acquiring language – a **language acquisition device** that he calls **Universal Grammar**.

Possible support for this thesis could be found in data on the phonetic, phonological and grammatical complexity of existing languages – demonstrating that language complexity is not related to the level of complexity of other aspects of human activity, such as technology. While we may speak (with good reason or not) of primitive societies, there are no "primitive human languages" that we know of. Quite the reverse – it would seem, from what we are in a position to observe at first blush, that language development leads to simplification. **Otto Jespersen** (1860–1943), a Danish linguist, studied trends in language development and hypothesised that the language of early man was full of difficult sounds, used a very wide range of pitch and tone and was structurally quite complex – as also noted by other linguists:

> "*Our Australian verb... rivals and excels the Greek and the Sanskrit, for it has four futures, and, for time past, it has three forms, marking the past time as instant, proximate, and remote. Corresponding to these tenses, there are nine participles, each of which may be used as a finite verb. Besides an imperative mood and a subjunctive mood, there are reflexive and reciprocal forms, forms of negation, forms to express continuance, iteration, imminence, and contemporary circumstances...And, in Australian, this copiousness of diction is not confined to verbs: it shows itself also in the building up of other words.*"
>
> *From* J. Fraser's Appendix D to L. E. Threlkeld, 1892. *An Australian language as spoken by the Awabakal.*

Of course, it does not follow from observations on contemporary primitive societies that the language of the first humans was structurally complex. Other linguists, who study pidgins or early language acquisition in children, present evidence of poorer levels of grammaticalisation at earlier stages of language acquisition and development.



For Prof. Noam Chomsky, Universal Grammar is the answer not only to problems of language complexity, but also to the questions of early language acquisition by children.

Children are not taught language by their parents in any systematic way. This is a task most parents would not be equal to, as in most people language use is not based on conscious knowledge of language structure. Children do not *learn*, they *acquire* their mother tongue by listening to conversation, which can often consist of incomplete statements or poorly formed structures. Studies of language acquisition demonstrate that children do not acquire language by simple imitation or stimulus-response: they form hypotheses about its structure on the basis of a poverty of stimuli. This is the essence of Chomsky's **poverty of stimulus argument**.

> *"Evidently, development of language in the individual must involve three factors: (1) genetic endowment, which sets limits on the attainable languages, thereby making language acquisition possible; (2) external data, converted to the experience that selects one or another language within a narrow range; (3) principles not specific to the Faculty of Language."*

Universal Grammar is a blueprint for language consisting of universal principles and sets of parameters, to be fixed upon exposure to a specific language.[2] It implies that **behind the structural diversity of the over 5000 existing languages and dialects, there are general rules or principles valid for all languages, as well as a set of no less general parameters, each with a restricted number of options**. Setting these parameters one way or another results in the structure of a specific language; and it is by exposure to a specific language that children set its parameters. The comparison of languages with different values for the universal parameters would thus provide an interesting basis for contrastive and typological studies.

[2] Cf. Chomsky, Noam 1986. Knowledge of Language: Its Nature, Origin and Use. New York: Praeger; Hyams, Nina M., 1986. Language Acquisition and the Theory of Parameters. Dordrecht: Reidel.

Situations, predications, judgments

Whether human language is the gift of a creator, the result of slow evolution or of a sudden natural change; whether it is the product of social interaction leading to enhanced cognitive ability or of a micro mutation resulting in the enlargement of the human brain and the development of an innate language faculty, the existence of a close relation between language and thought has remained undisputed through the ages: they both refer to real world entities with their properties and interrelations as constituents of **situations**. As Kenneth Jon Barwise and John Perry, two American logicians and philosophers of language wrote in their first work on Situation Semantics:[3]

> "*The world consists not just of objects, or of objects, properties and relations, but of objects having properties and standing in relations to one another. And there are parts of the world, clearly recognized (although not precisely individuated) in common sense and human language. These parts of the world are called situations*."

The entities in a situation are its "**participants**". The number of participants and their roles ("**thematic roles**") in the situation are dependent on its type. Sleeping or running situations typically involve one participant role, fighting situations – two, writing situations – three (writer, written product, instrument), etc.

There is no unanimity on the number and definition of thematic roles, but the following are considered standard. They can be observed in all languages because they refer to entities and relations in a world that humans share and perceive in similar ways – due to common cognitive capacity and thought.

Agent: an active participant who/which brings about the state of affairs.

Ex. ***Tristan*** *opened the box and took out the key*.

Experiencer: a participant who/which undergoes an experience (emotional, sensory, etc.)

Ex. **The children** *are feeling so excited*!

[3] Barwise, J. and Perry, J. The Situation Underground. In: Stanford Working Papers in Semantics, vol. I, eds. J. Barwise and I. Sag. Stanford Cognitive Science Group, 1980, Section D, pp. 1–55.

Recipient (or **Beneficiary**): a passive participant who/which receives something in the situation.

Ex. *Did you send **Granny** a birthday card*?

Instrument: the unconscious, non-sentient means/tool acting on a participant in the situation.

Ex. *My father always slices the meat with **his very special carving knife***.

Cause: a usually unconscious and non-sentient phenomenon or state of affairs which brings about a change in the other participants.

Ex. ***The hurricane** damaged many buildings and wounded three people.*

Path: the distance traversed by one or more participants.

Ex. *We crossed **the Atlantic** in around 5 days*.

Location: the place where the situation takes place or the endpoint of the path.

Ex. *When we finally arrived in **Torino**, it was past midnight*.

Measure: the amount of time, quantity, etc.

Ex. *This pair of shoes cost me **an arm and a leg***.

Theme: an argument undergoing, literally or metaphorically, a change of state.

Ex. *Who painted **the roses** red*?

In logical models of situations, such as predicate logic, situations are formulated as **predications**, the participants in which are **arguments**. **Situations are** thus presented as **relations between a predicate and its arguments**. A situation with one participant is a **one-place predicate**, p (x), e.g. sleep (x). A situation with two participants is a **two-place predicate**, p (x, y) e.g. write (x, y).

When the variables x, y, etc. **are substituted by specific values**, the predication becomes a **proposition**: sleep (John) and write (Mary, letter).

Each predication or proposition states a property, expresses characterisation. Establishing the characteristics of entities is judging; the thought expressed is a judgment. **The judgment is the simplest unit (or atom) of thought**. The judgement is a thought consisting of two parts: a thought about the characterised (the subject of thought) and a thought about the predicated characterisation (the predicate). As pointed out by **W. M. Wundt** (1832–1920), a German psychologist and philosopher, **the simplest form of thought** is the splitting up (analysis) of a total thought (or representation) into **two components**: **a predicate and a subject**.

A proposition can form the basis for the formulation of different judgements, in which the situation is split up in different ways. To take an example, the proposition *loves (John, Mary)* can form the basis for the judgments:

[John $_{subject}$] [loves Mary $_{predicate}$] or
[Mary $_{subject}$] [is loved by John $_{predicate}$] or
[The one who loves Mary $_{subject}$] [is John $_{predicate}$], etc.

How many judgments can you formulate from the proposition read (Peter, story)?

The ability to present situations in the form of propositions and judgments and to analyse judgments into subjects and predicates is common to human beings. Also common to human beings is the linguistic expression of judgments. Like judgments, most (though not all) clauses and sentences contain two major parts – a (grammatical) subject and a (grammatical) predicate. Just as the arguments of the proposition can be subjects or predicates of a judgment, the phrases expressing thematic roles can be assigned different functions – due to the variations that the syntax of human language allows.

How language affects the development of thought

All conscious processes of thought, the results of which can be used in future, take place in language form.[4]

- Language allows us to lend structure to our processes of thought and is essential to the processes of concept-formation.
- Language gives focus to processes of thought, makes them conscious and facilitates the establishing of cause-result relations.
- Processes of cognition build on explicit or implicit comparisons for the purpose of establishing similarities. Our ability to compare is innate and is based on another ability – to create structure and impose structure.
- Comparison is a prerequisite for establishing analogies – for which humans have a much greater capacity than other species. This capacity is, to a considerable degree, boosted by the complex net of relations between the elements of language.
- The syntactic structures of language allow complex variations of thought.

As F. de Saussure, the father of modern linguistics, points out,

> "*Language is no longer regarded as peripheral to our grasp of the world we live in, but as central to it. Words are not mere vocal labels or communicational adjuncts superimposed upon an already given order of things. They are collective products of social interaction, essential instruments through which human beings constitute and articulate their world.*"

[4] R. Jackendoff 1996. "How language helps us think". *Pragmatics and cognition* 4, 1–36.

Quiz 2

1. **M. Tomasello views the development of language in relation to:**
   a. primate cognition
   b. the development of thought
   c. its communicative function
2. **Semanticity is related to:**
   a. the function of signals
   b. the meaning of signals
   c. the structure of signals
3. **Arbitrariness is:**
   a. a logical relation between the form and meaning of the sign
   b. a conventional relation between the form and meaning of the sign
   c. a necessary relation between the form and meaning of the sign
4. **Prevarication is:**
   a. rapid fading
   b. openness
   c. deception
5. **A language acquisition device is:**
   a. a tool for foreign language learning
   b. a predisposition for language
   c. a language app
6. **Chomsky's poverty of stimulus argument is presented as proof of:**
   a. universal grammar
   b. learnability
   c. reflexiveness
7. **Which of the following is NOT a thematic role?**
   a. Path
   b. Topic
   c. Measure

Further reading

Barwise, J. and Perry, J., 1980. The Situation Underground. In: *Stanford Working Papers in Semantics*, vol.I, eds. J. Barwise and I. Sag. Stanford Cognitive Science Group, 1980, Section D, pp. 1–55.

Chomsky, Noam, 1986. Knowledge of Language: Its Nature, Origin and Use. New York: Praeger

Hauser, Marc; Chomsky, Noam, Fitch, William Tecumseh, 2002. "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" Science, 298 (5598): 1569–1579.

Hockett, Charles F.,1960. The Origin of Speech. *Scientific American* 203.

Hyams, Nina M., 1986. Language Acquisition and the Theory of Parametrs. Dordercht: Reidel.

Jackendoff , R. 1996. How language helps us think. *Pragmatics and Cognition* 4, 1-36.

Threlkeld, L. E., 1892. An Australian language as spoken by the Awabakal, the people of Awaba, or lake Macquarie (near Newcastle, New South Wales) being an account of their language, traditions, and customs. Sydney: Charles Potter, Government Printer.

Tomasello, Michael; Carpenter, Malinda; Call, Josep; Behne, Tanya; Moll, Henrike, 2005. "Understanding and sharing intentions: The origins of cultural cognition". Behavioral and Brain Sciences. 28: 675–735.

Lecture 3

A system *sui generis*

Structuralism

While 19th century linguistics devoted most of its efforts to the comparison of the sounds, grammar and lexical stock of different languages in view of establishing genealogical relations and reconstructing common ancestors, **the greatest contribution of 20th century linguistics was the focus on the inner organisation and functioning of each language viewed in isolation and in its relatively stable contemporary state.** This methodology of analysis, "**structuralism**", views specific fields of study as complex **systems** made up of interrelated and interdependent parts (**elements**). The complex network of relationships which defines the place of each element is the system's **structure**.

Structuralism became an influential methodology in all of the humanities, but it originated as a theory of language – in the work of a linguist of Swiss origin, **Ferdinand de Saussure** (1857–1913). Saussure was born in Geneva and developed his influential lecture course in general linguistics at the University of Geneva. His lectures were published posthumously, in 1916, by his students, the (future) French linguists Albert Sechehaye and Charles Bally, as *Cours de linguistique générale*. Saussure's major contribution to linguistic theory lies in:

- the separation of the **synchronic** linguistic **analysis** from the historical, **diachronic perspective**;
- the delimitation of **language** (*langue*) and **speech** (*parole*) and the establishment of language as the prime object of study for Linguistics;
- the definition of **language as a system**, the units of which are elements interrelated by a **complex structure**;
- the definition of the elements of language as **signs**, the elaboration of **the nature and structure of the linguistic sign**;
- the analysis of language structure as a complex **set of oppositions**;
- the definition of two major **axes of analysis: paradigmatic and syntagmatic**;
- the view of **language as an abstract system**.

Ferdinand de Saussure's theory of language was further developed by the linguistic circles of Prague and Copenhagen, by the tradition of descriptivism in the USA, and by the French structuralists.

Cours de linguistique générale

***Langue* and *parole*.**

Ferdinand de Saussure applied to linguistics the Humboldtian distinction between active doing (*energeia*) and the product of that doing (*ergon*), introducing the opposition of ***langue*** (language) **– an abstract systematic rule system used by a community of speakers**, yet **independent of the individual speakers –** , and ***parole*** (speech, utterance) **– the use of language**, its concrete manifestation in oral and written form.

Even though we only have access to *langue* through the observation and study of *parole*, it is in the study of langue that Saussure's interest and contribution lies: in its inner organisation and the nature of its elements.

To illustrate the opposition between langue and parole, Saussure made use of **the chess metaphor**. He compared the rules of chess to *langue* (the norms for playing the game) and the individual choices of the players in making their moves – to *parole*.

Look at the chessboard in the photo. How would you comment on the pieces? Would it be possible to use this board and the pieces in a game of chess?

Along with *langue* and *parole*, Saussure also introduced a third term, ***langage* – the general faculty to express oneself with signs**. This faculty is common to natural and artificial languages and to different systems of communication. It is common to all systems of human communication.

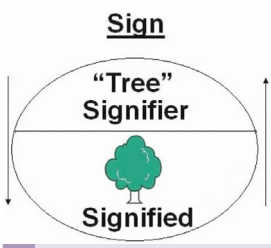
The idea of a general human faculty underlying our use of language was taken up half a century later by Noam Chomsky, and reformulated as "the human **language faculty**" based on an innate "universal grammar" (Cf. Lecture 2). He proposed an opposition of linguistic **competence vs. performance**.

Linguistic competence, as opposed to Saussure's *langue*, is not independent of the separate individual; it is the system that the individual has acquired a/ due to his or her language faculty and b/ by virtue of being exposed to the linguistic performance (Saussure's *parole*) of other members of the language community.

The sign character of language

F. de Saussure was the first to define **language as a system of signs**. Without knowing the signs of a language, *parole* cannot be understood: it would be perceived as a meaningless, unorganized string of sounds or symbols. In Saussure's theory, the sign is an abstract, mental unity of two interrelated parts: a **signified** and a **signifier**.

The signified is a concept, a mental image – in the simplest case, of a real world object, e.g. a tree. Trees can be of different shapes and different types and the mental image of a tree in different individuals may differ slightly, but the basics should be common to most humans: a perennial plant of a considerable height having a trunk, branches and leaves. We share concepts via signifiers. **The signifier** is, again, a mental image: acoustic or other. It is the sign as a unity between signified and signifier that is the basic element of language.

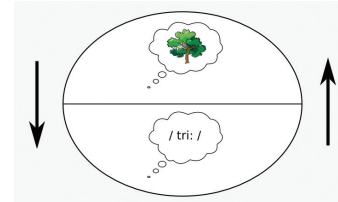


Taken in isolation, the signifier is not, as a rule, predetermined by the nature of the signified. The same signified can have different signifiers in different languages: *arbre* (French), *Baum* (German), *árbol* (Spanish), *дърво* (Bulgarian), *mti* (Swahili), etc. Saussure defines **the relation between signified and signifier** as **arbitrary**.

Do you know the signifiers for this signified in other languages? Can you think of cases where the relation between a signified and its signifier is not (entirely) arbitrary?

Another characteristic feature of the signifier, apart from its arbitrary relation to the signified, is its **linear nature**. The arbitrariness of the sign and linearity of the signifier are two of the major principles underlying the specific functioning of the system of language.

In Saussure's view, **language is form, not substance**, and the linguistic sign has an entirely abstract nature. It is **a relation** not between an object and a name, but **between a concept and an acoustic image**. The picture on the right would thus be a more accurate illustration of Saussure's linguistic sign.



Language as a system of signs System and structure

The linguistic signs do not form an unorganised set; they are connected by nets of relations of different types. The specific organisation of language is its **structure**. The set of structurally organised interrelated elements is called a **system**.

Language elements do not normally appear in speech in isolation; they form sequences, or **strings**. The **linear relations** between adjoining elements in a string are called by Saussure **syntagmatic**. The linear groups themselves he named **syntagms**. The syntagm is a linguistic unit consisting of a set of linguistic units that are in a sequential relationship to one another. It is always composed of two or more units.

In the string *the linguistic signs do not form an unorganised set*, syntagmatic relations exist between *linguistic* and *signs*, between *the* and *linguistic signs*, between *signs* and *do not form*, between *un* and *organised*, etc.

What other syntagmatic relations can you identify in the above example? How many syntagmatic relations can you identify in the following quote: "The limits of my language are the limits of my world." (Ludwig Wittgenstein)?

Paradigmatic relations are relations between sets of linguistic items that form possible choices in particular syntactic positions or for particular syntactic functions. Paradigmatic relations are based on associations of elements having one or more common features, but at the same time differentiated with respect to at least one feature. These relations exist between elements of the system outside the linear strings where they occur.

A **paradigm** is based on a common core and variation with respect to one feature; a minimal paradigm consists of two members. In the quote above, paradigmatic relations can be established between 'the' and 'my'. Paradigmatic relations exist between 'world' and other morphological forms of the word, e.g. *worlds* or *world's*.

Another term for paradigmatic is **associative**. Associative relations can exist between sentence types (e.g. interrogative sentences) or clause types (e.g. non-finite clauses), between synonyms and antonyms, between the forms of a word (word forms), between words of the same type (e.g. nouns, adverbs), etc.

In language, there is nothing but differences

The process of cognition involves a number of interrelated stages, including structuring, mapping of structures, comparisons for the purpose of establishing similarities and differences, establishing of analogies. Human language with its specific structure, with the **net of oppositions** in which the elements of the system enter, strongly supports and enhances cognitive processes.

Similarities and differences can be established, as in any cognitive process, via mapping and comparison. Comparisons in language can be made within two-member or multi-member oppositions. Any opposition however is grounded in similarity: of meaning, category, form or function. The elements of language can be compared and opposed in different ways and along both axes of the system: the syntagmatic and the paradigmatic one.

Syntagmatic relations exist between elements which can be **constituents** (building blocks) of a structure – which forms their common base and similarity – but cannot appear in the same position within that structure – wherein lies their dissimilarity. The two word forms in the string *red rose* are similar in that they are constituents of a phrase – which is due to their grammatical and semantic compatibility: the adjective *red* easily combines with concrete nouns, but not with abstract ones (**red happiness*). The two word forms differ in that they cannot appear in the same position and function in the phrase and are not interchangeable in it.

Associative relations cannot be directly observed in the flow of speech: associatively related elements can be interchangeable in some positions and mutually exclusive in others. These elements are retained in memory in the form of sets (paradigms) of units which are grouped because of a similarity, of a common feature. Within the paradigm, they must differ for one or more other features. Oppositions within a paradigm can be based on sound (*coach/roach, king/ring, code/road*), on a different component of meaning (*boy/girl* or *boy/man* or *boy/calf*), on the expression of a different value of a category (*boys/ boys* or *work/worked*), etc.

The point of Saussure's assertion that *in language, there is nothing but differences* is that **if two elements of the system are identical** in form, meaning, syntactic function or stylistic register, **one of them is redundant**. It will inevitably either undergo structural or functional change, or else disappear from the system.

Synchronic and diachronic linguistics

Synchrony and **diachrony** are two relatively opposed approaches to linguistic analysis. The synchronic approach considers the system of a language in one of its relatively stable states, without taking into account historical change. The diachronic approach views phenomena in development, focusing on linguistic change.

In contrast with most of his predecessors, who focused on the evolution of languages and in the reconstruction of protolanguages, Saussure emphasized the primacy of synchronic analysis to understand the inner functioning of language.

Prior to around 1200, English formed negative structures by placing the particle *ne* before the verb and *not* or *nawt* after it: *Ic **ne** seye **not*** – I don't say. *He **ne** speketh **nawt.*** – He does not speak. Around 1400, ***ne*** was used infrequently and ***not/nawt*** typically occurred by itself after the verb: *I seye **not** the words. We saw **naw**t the knyghtes*. It is several centuries later that the current practice of allowing *not* to occur after certain types of verb only became the rule.

Compare this development to negation in another language that you are familiar with. Are there parallels? Do you think this change could be viewed as evolution?

Does the information presented above help you form grammatical negative sentences in Modern English?

The following suffixes are not used to form new words in Modern English.

-l : especially productive in Middle English, carries the idea of iteration. Ex.: *to twinkle* (from OE *twinclian*), *to wrestle* (from OE *wristlian*), *to rattle, to sparkle*, etc.

-er : to form verbs with a connotation of frequency and iteration. It is no longer productive in Modern English, but has been preserved in: *twitter, chatter, jabber* etc.

-ster: for feminine agent nouns. Only *spinster* has survived with the suffix in its original meaning. In Modern English it is not very productive and gives a degrading nuance: *gamester, gangster* etc.

-dom: appended to noun and adjectival roots, it formed abstract nouns with the connotation of 'condition', 'state', 'dignity'.

-ship: connected with Mod. English *shape*, it was appended to nominal and adjectival roots to form abstract nouns with the connotation 'state or condition of being so-and-so'.

-hood: from OE *had*: state, condition, nature, form, kind.

-the: a very old OE suffix, was appended to adjectival roots to form abstract nouns – *strength, width, length*, etc.

The Prague Linguistic Circle

The linguistic views of Ferdinand de Saussure, as well as the work of his contemporary – the Polish linguist **Jan Baudouin de Courtenay** (1845–1929) and his **functional approach to language**, exerted a strong influence on a group of linguists from Central and Eastern Europe who founded, in the late twenties, a linguistic society known as the **Prague Linguistic Circle**. In 1929 the Circle submitted a seminal paper, *The Prague Theses*, to the First Congress of Slavists. The same year marked the launch of their journal, *Travaux du Cercle Linguistique de Prague*.

The linguistic school of Prague can be characterised as **structural and functional**. The Circle adopted and developed the following main tenets of Saussure's theory of language:

- the **distinction between synchrony and diachrony** (though admitting that this distinction cannot be absolute). Diachrony relates different synchronic stages; and synchronic stages retain archaic elements of structures in development: language evolution can be observed on the synchronic plane;
- the **structural approach to language –** language as a system made up of structurally organised signs which are unities of a signified and a signifier.
- the central **notion of oppositions** – which they developed into a full-fledged theory and applied, **first and foremost, to a new branch of linguistics that they founded – phonology**, but also to morphology and syntax.

The theory of phonological oppositions was first developed by the Russian linguist Nikolai Trubetzkoy (Cf. also Lecture 4). It was **applied to Morphology by the Russian-American linguist Roman Jakobson** (1896–1982), who identified oppositions in grammar and used them to define grammatical categories. The major **contribution of the Prague Circle to Syntax is in the functional approach developed by the Czech linguist Vilém Mathesius** (1882–1945). He is considered to be the father of modern functional syntax.

The Prague scholars **did not favour a strict distinction between *langue* and *parole***. For them, speech was the immediate, observable reality of language, while language was, simply, the result of the scientific analysis of speech.

Descriptivism

In the United States structuralism took a specific turn, having been preceded by **Franz Boas**'s (1858–1942) study of the languages of the native population of the continent and the 1911 publication of his "**Handbook of Amerindian Languages**". Boas was convinced that the development of new, formal methods of analysis was needed – as the categories and norms of Greek and Latin grammars were not applicable to the native languages of the New World.

It is to **Leonard Bloomfield** (1887–1949) that American structuralism is mostly indebted. In his book "**Language**", which appeared in 1933, Bloomfield presented a comprehensive description of the principles of American structural linguistics – descriptivism – and the methods of **Distributional** and **Immediate Constituents (IC) Analysis**.

The **linguistic theory of Bloomfield** is based on:

1/ the hypothesis of **regularity** in language,

2/ the **exclusion of psychology** from the analysis of data;

3/ adherence to **precise**, scientific descriptive **formulations**;

4/ the **inclusion of meaning**.

IC analysis is a recursive procedure of breaking up a sentence or clause into its constituent parts, until the level of the morpheme has been reached.

The methods of **Distributional Analysis** are: A. **Segmentation**: the string is segmented into minimal units. Minimal phonetic sequences that recur with a constant meaning are identified as morphemes. B. **Identification**: sequences that are repeated in different contexts are identified as morphemes; identical contexts are identified, and the elements that go in the slots are identified as morphemes; C. **Substitution**: different elements that appear in identical / similar slots are identified as morphemes.

The sequence in which an element occurs is its **environment**; the set of all environments in which an element occurs is its **distribution**.

For the strings *restructured, renewed, reformulated, revived,* what sequences and what contexts would the method of distributional analysis identify?

Levels of linguistic analysis

A specific manifestation of the systematic organisation of natural language, as demonstrated with the methods of distributional analysis, is its hierarchical structure. In European linguistics, the notion of **levels of linguistic analysis** was first introduced by the French linguist and semiotician **Emile Benveniste** (1902–1976), one of the critics of Saussure's *Course*.

According to Benveniste,

1/ to say that the units of language have structure is to say that they are made up of units of a lower level, constituents, which build up their form;

2/ the units of a level of language can in turn be integrators for the units of the higher level, which determines their meaning/function.

Benveniste proposes a method of consecutive identification of the units of language and the respective levels with the formal procedures of descriptivism: identification, segmentation and substitution. At the bottom of this hierarchy lies the phonological level with, as its main unit, the phoneme. The sentence forms the top of the hierarchy: the syntactic level. The word is seen as a unit of intermediate status in the hierarchy.

Two main **types of relations between the units** of language are defined: **distributional** (between the elements of the same level) and **integrative** (between elements of different levels). Units of the lower levels are integrated within units of the higher ones; for the latter, they are **constituents**.

Only the immediate constituents of a unit are structurally visible at its level. Thus, while lexical units are made up of morphemes and, ultimately, of phonemes, the structural analysis of lexemes stops at the level of morphemes. Similarly, phonemes and morphemes are not immediate constituents of the units of the syntactic level.

Each level is a system in itself, with specific rules which operate on it. The rules of phonology determine the occurrence and combination of phonemes; the rules of morphology – the behaviour of morphemes, and the rules of sentence formation – the combination and positioning of words in a sentence. Each level of analysis corresponds to each level of the structure of language.

Note that this view of language structure is a considerable deviation from the Saussurean understanding of language as an abstract system in which the complexity of the signifier is not related to the status of the sign. Also, it does not find confirmation in the results obtained after application of distributional analysis.

The consecutive top-down application of the Immediate Constituents analysis yields structural positions which can be filled by paradigmatic sets of structurally very different units:

Mary / sent //	[0]	a book.
	[John]	
	[the boy]	
	[her old friend]	
	[the boy I told you about]	
	[whoever asked for it]	

Not words, phrases are the immediate constituents of the sentence and the clause. However, as illustrated in the example above, phrases may in turn be clauses. This feature of language – the possibility to embed units of a given structural type into units of the same structural type, is an example of **recursion**.

Phrases can be complex, but also null: the first phrase in square brackets in the example above is zero-filled. **The structural importance of null or zero units in the system**, both on the level of form and of function, follows naturally from the very nature of language as a system of oppositions, where the presence of a feature is opposed to its absence.

The phenomenon of recursion and the structural importance of zero in linguistic description, together with the asymmetrical relation between the complexity of the signifier and the status of the sign in the system, are among the main features of language which characterise it as a system *sui generis*. These features, as pointed out in Lecture 10, form an integral part of Generative Syntax.

Quiz 3

1. **Ferdinand de Saussure made the analogy with chess to illustrate:**
   a. the sign character of language
   b. the opposition between synchrony and diachrony
   c. the opposition between language and speech
2. **According to Saussure, the elements of the system of language are:**
   a. signifieds
   b. signifiers
   c. signs.
3. **A structure is:**
   a. an organised string
   b. a specific set of elements
   c. a specific organisation of elements
4. **Paradigmatic relations are:**
   a. associative relations
   b. syntagmatic relations
   c. relations of formal similarity
5. **It is considered that the main contribution of the Prague Circle is the theory of:**
   a. phonology
   b. morphology
   c. syntax
6. **The linguistic theory of Bloomfield is based on:**
   a. the inclusion of both psychology and meaning
   b. the inclusion of psychology and the exclusion of meaning
   c. the inclusion of meaning and the exclusion of psychology
7. **The immediate constituents of the word are:**
   a. phonemes
   b. morphemes
   c. distinctive features.

Further reading

Benveniste, Emile, 1981. Problèmes de linguistique générale. Paris, Editions Gallimard.

Bloomfield, Leonard, 1933. Language. New York, Henry Holt.

Fries, Charles C., 1952. The Structure of English. New York: Harcourt, Brace.

Hockett, C. F. (ed.), 1970. A Leonard Bloomfield Anthology. Bloomington.

Holenstein, E., 1975. Roman Jakobson's Approach to Language: Phenomenological Structuralism. Bloomington and London: Indiana University Press.

Lepschy, Giulio C., 1970. A Survey of Structural Linguistics. London, Faber and Faber.

Rivkin, Julie, Michael Ryan (eds.), 2004. F. de Saussure. Course in General Linguistics. Literary Theory: An anthology. Blackwell.

Saussure, Ferdinand de. Course in General Linguistics. McGraw-Hill, 1966.

Seuren, Pieter A. M., 2009. Western Linguistics: An Historical Introduction. Oxford, Blackwell.

Trubetzkoy, N. S.,1969. Principles of Phonology. Berkeley.

Lecture 4

Phonetics and Phonology

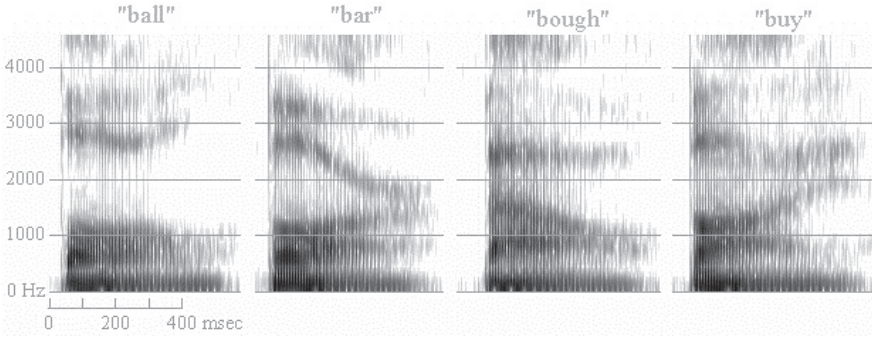
Phonetics

Phonetics is the branch of linguistics which studies the sound structure of language and two interrelated systems: a) the system of segmental units – the speech sounds: how they are formed and what their articulatory and acoustic features are; b) the system of suprasegmental units and features – vowel length, syllable stress, intonation, tone.

Phonetics can be subdivided into three branches:

- the study of the production of sounds: articulatory phonetics;
- the study of the perception of speech sounds: auditory phonetics;
- the study of the physical properties of speech sounds: acoustic phonetics.

Acoustic phonetics involves the analysis of sound vibrations – their amplitude, their duration and the frequency of waveforms. It makes use of instruments such as (in its early years) the phonograph and (nowadays) the spectrograph.



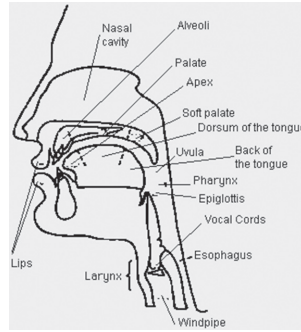
Auditory phonetics is a field related to anatomy and physiology which studies:

- how we perceive messages (strings of sounds) with our hearing (auditory) apparatus;
- how these messages are transformed into neural signals;
- how the information received is processed in the brain;
- how the signals are decoded into messages.

Articulatory phonetics studies sounds: how they are produced by speakers, what their nature is and how to classify them.

The production of speech sounds

Producing speech sounds involves the conversion of the stream of air that comes from the lungs and passes through the oral and/or nasal cavities into soundwaves. The following parts of our physiology participate in this process:



1/ the **larynx**, situated at the top of the **trachea**;

2/ the **subglottis**, situated just below the larynx

3/ the **vocal tract**, above the larynx, composed of the oral and nasal cavities.

The air passing through the larynx and glottis is the source of **sound waves**.

Moving the stream of air out of the lungs and through the larynx and the vocal tract produces speech sounds.

The larynx contains folds of muscle called **vocal folds**. The opening between them is the **glottis**. The vocal folds are attached at the front of the larynx and free at the back. When the vocal folds are relaxed, the air passes freely through them and the sounds produced are voiceless. When the vocal folds come close together, the air passing makes them vibrate and the sounds produced are voiced.

Put a hand lightly on your throat and pronounce the following pairs of words:

fat/vat; sip/zip; dilution/ delusion; rich/ridge; pat/bat; tap/dab; kill/gill

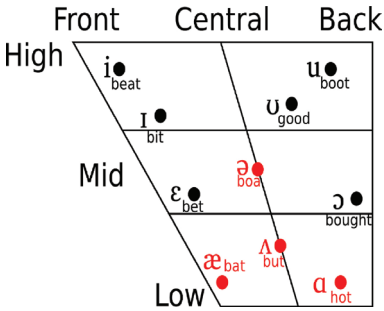
Can you feel vibration? Which of the sounds are voiced?

Vocal fold vibration is the sound source for **vowels**. The vocal tract above the glottis acts as a **resonator** affecting the sound made by the vocal folds. The shape of this resonator determines the quality of the vowel.

There are several ways in which speakers can change the vocal tract and thus change vowel quality. They do this by:

- raising or lowering the body of the tongue
- advancing or retracting the body of the tongue
- rounding or not rounding the lips
- making these movements in a tense or in a lax manner.

When describing a vowel, it is necessary to provide information about these four aspects of its articulation.



- [i] high, front, unrounded, tense (as in *meat*)
- [ɔ] mid, back, rounded and lax (as in *bought*)
- [a] low, back, unrounded and lax (as in *got*)
- [ʌ] mid, central, unrounded, and lax (as in *hut*)
- [e] mid, front, unrounded and tense (as in *make*)

One single vowel sound is called a **monophthong**. **Diphthongs** are gliding vowels – combination of two adjacent vowel sounds within the same syllable. Some languages (mainly of the Romance and Turkic groups) also have **triphthongs** – similar combinations of three vowels.

Describe the vowel sounds in the following words:

TAKE
BEER
JOY
HIT
POT
AIR
GRAB

How many diphthongs and triphthongs are there in English? Find examples.

Consonants are speech sounds produced with a narrowing in the vocal tract. When describing a consonant, it is necessary to provide information about three aspects of its articulation: 1/ whether it is voiced or voiceless; 2/ the place where the airstream is constricted; 3/ the type of airstream constriction. Consonants can be **voiced** or **voiceless** depending on whether they are made with or without vibration of the vocal folds.

The **place in the vocal tract** where the air is constricted defines the consonants produced as:

- **Bilabial** – produced by bringing the lips together: [p], [b], [m], [w].
- **Labiodental** – produced when the lower lip is brought against the upper front teeth: [f], [v].
- **Interdenta**l – produced when the tip of the tongue is placed between the front teeth: [θ] (as in *things*), [ð] (as in *there*).
- **Alveolar** – produced when the tip of the tongue is placed against the ridge behind the upper front teeth called **alveolar ridge**: [t], [d], [s], [z], [n], [l], [r]
- **Palatal** – produced when the tip of the tongue is placed against the hard **palate**: [ʃ] (as in *sure*), [ʒ] (as in *pleasure*), [tʃ] (as in *watch*), [dʒ] (as in *lodge*), [j] (as in *yes*).
- **Velar** – produced when the tip of the tongue is placed against the **velum** (the soft palate): [k], [g] and [ŋ] (as in *bring*).
- **Glottal** – produced in the glottis: [h] (as in *house*) and [ʔ] (as in the cockney pronunciation of *bottle*).

The **manner of articulation** of a consonant sound depends on how close together the articulators are.

- **Stops** are produced by completely obstructing the airstream: [t], [d], [p], [b], [k], [g] or [ʔ].
- **Fricatives** are produced by making a very small opening, so that friction is produced: [f], [v], [s], [z], [ʃ], [ʒ].
- **Affricates** are produced by first briefly stopping the airstream, then releasing the articulators slightly – so that friction is produced: [tʃ], [dʒ].
- **Liquids** are produced with an obstruction which does not stop the airflow and does not cause friction: [l], [r].
- **Glides** or **Semi-vowels** are produced as for high vowels, but with a very slight closure of the articulators: [w], [j].
- **Nasals** are produced by lowering the velum so that the air cannot enter the oral cavity and passes through the nasal cavity: [m], [n] and [ŋ].

Phonology

While phonetics studies the sounds of speech from the point of view of their articulation, acoustic characteristics and auditory perception, **phonology views speech sounds in their functional aspect**. Two languages can have the same **phones**, the same sounds, and still sound different. The reason is that each language organises the sounds that it makes use of in a specific way. Making use of the same set of sounds, these languages will have different phonological systems.

The theory of distinctive features

The theory of **distinctive features** studies aspects of articulation as constituent features of speech sounds: voiced or voiceless, bilabial, labiodental or palatal, nasal, etc. Thus, [m] is voiced, bilabial, nasal; [f] is labiodental and voiceless. The goal of the theory is to identify a set of features which is sufficient for the description of the phonological systems of as many languages as possible.

The theory of oppositions

The contrasts between phonemes can now be described in terms of oppositions of sets of distinctive features, or of oppositions of sets of speech sounds with respect to a distinctive feature, e.g. the set of sounds having the distinctive feature [+nasal] as opposed to those having the feature [-nasal]. Modern phonology was first elaborated by the Russian linguist **Nikolai Trubetzkoy** (1890–1938), who identified three major **types of oppositions**:

- **Privative**: one member possesses a feature (it is marked for it, it is the marked member of the opposition), the other member does not possess the feature (it is the unmarked member of the opposition);
- **Equipollent** (the two members of the opposition have equal standing);
- **Gradual** (the members of the opposition are all marked for the feature, but to a different degree).

Phonemes, though they cannot be further segmented, can be analysed in terms of distinctive features. It is due to these characteristic features that speakers of language can tell one phoneme from another. With respect to its bundle of distinctive features, **a phoneme enters into oppositions with all the other phonemes of the language**. Some phonemes are in opposition for one distinctive feature only (e.g. +voiced/-voiced:[d]/[t]); others are opposed for a number of features: (e.g. [m] (+nasal, +voiced) and [p] (-nasal,-voiced).

We do not always perceive oppositions for one or more distinctive feature as oppositions of sounds. Some pairs of sounds which differ with respect to one or more distinctive features in the bundle can be perceived by the speakers of one language to be two sounds, while speakers of other languages consider them to be one sound. Speakers of English perceive the alveolar sound in *table* and the glottal stop in *bottle* as the same [t] sound. Similarly, the different quality of the initial sound in *боб* and *бял* is not felt to be distinctive by speakers of Bulgarian. The sounds [l] and [r] are not distinctive in Swahili. For speakers of other languages however, such differences are important and the change of sound would lead to changes in the meaning of the message.

Phonemes and allophones

Following the work of the Prague Linguistic Circle, a set of speech sounds (phones) identified by the native speaker as the same sound is called a **phoneme**. The members of these sets are called **allophones**. An allophone is a phone that has been classified as belonging to some set or class of phones, i.e. to a phoneme. Phonemes can also be defined as the forms in which we store sounds in our memory. While phones are studied as elements (segments) of speech, phonemes are elements of language. They are differentiators of meaning, the smallest phonological units which cannot be decomposed into smaller successive units.

Environment and distribution

The phonological system of a language is made up of its phonemes and their interrelations with other phonemes. Phonemes seldom appear isolated in speech. Most often, they appear in the **environment** of other phonemes. This environment can affect their features; it can also **neutralise their oppositions**. The set of the possible positions and environments of the phoneme is its **distribution**.

Some phonological processes can best be observed in the context of the next larger segment of speech – the **syllable**. The syllable consists of one vowel sound, of a diphthong (a sound formed by the combination of two vowels in a single syllable (as in shout), or else of a syllabic consonant (as in English even [ˈiːvn̩] or Slovak vŕba [ˈvr̩ːba]), with or without preceding or following consonants.

Phonological rules and processes

Three important rules of phonology were formulated by N. Trubetzkoy:

First rule: If in a given language two sounds occur in the same position and are interchangeable without changing the meaning of the string, they are **facultative variants of a phoneme,** or allophones. For such sounds, we say that they are in **free variation**.

Second rule: If in a given language two sounds appear in the same position but are not interchangeable without changing the meaning of the string, they are **realisations of two different phonemes**. For such sounds, we say that they are in **constrastive distribution**.

Third rule: If two acoustically similar sounds never appear in the same position, then they are **combinatorial variants of the same phoneme**. For such sounds, we say that they are in **complementary distribution**.

Several processes of change in the distinctive features of phonemes can be observed as a result of interrelations with context. They are called "**phonological processes**".

- Processes of **assimilation** cause a phoneme sound to become more like a neighbouring sound with respect to some feature – Cf. the pronunciation of *handbag* in rapid speech as [ˈhæmbæg].
- Processes of **dissimilation** cause two neighbouring phonemes to become less alike with respect to some feature – Cf. fricative dissimilation in English: *fifth* [fɪft].
- Processes of **insertion**: In English, between a nasal and a voiceless fricative, a voiceless stop with the same place of articulation is inserted: Cf. the insertion of [k] in *strength* – [strɛŋkθ].
- Processes of **deletion** eliminate sounds. Such rules apply more frequently in unstressed syllables and in casual speech. A typical example in English is [h]-deletion in unstressed position – *Did you tell her* [ə] *the news*?

Suprasegmental features

Prosody is concerned with those elements of speech that are not individual phonetic segments (vowels and consonants) but are properties of syllables and larger units of speech. These contribute to linguistic functions such as intonation, tone, stress, and rhythm – called **suprasegmental features**.

In auditory terms, the major variables are:

- the **pitch** of the voice (varying between low and high)
- **length** of sounds (varying between short and long)
- loudness, or **prominence** (varying between soft and loud)
- **timbre** (the quality of the sound).

In acoustic terms, these correspond reasonably closely to:

- fundamental **frequency** (measured in hertz, or cycles per second)
- **duration** (measured in time units such as milliseconds or seconds)
- **intensity**, or sound pressure level (measured in decibels)
- **spectral characteristics** (distribution of energy at different parts of the audible frequency range).

Different combinations of these variables are exploited in the linguistic functions of intonation and stress, as well as other prosodic features such as rhythm, tempo and loudness. Additional prosodic variables have been studied, including voice quality and pausing.

Stress

Stress or accent is relative emphasis or prominence given to a certain syllable in a word, or to a certain word in a phrase or sentence. This emphasis is typically caused by such properties as increased loudness and vowel length, full articulation of the vowel, and changes in pitch (high/low).

Intonation is variation of spoken pitch that is not used to distinguish words; it is used for a range of functions such as indicating the attitudes and emotions of the speaker, signalling the difference between statements and questions, and between different types of questions, focusing attention on important elements of the spoken message and also helping to regulate conversational interaction.

Tone is the use of pitch in language to distinguish lexical or grammatical meaning – that is, to distinguish or to inflect words. Languages that do have this feature are called **tonal** languages. As many as seventy percent of world languages may be tonal**.** Tonal languages are extremely common in Africa, East Asia, and Mexico.

Phonetic transcription

There are languages, like Bulgarian or Italian, where the relationship between the ways in which words are written and pronounced is more or less consistent. For such languages we say that they have **phonemic orthography**. There are however also languages, like English, Irish or French, where the pronunciation of words changes faster than their written form (their **orthography**) – so much so that at some point it becomes necessary to provide a representation of the way the written words of these languages should be pronounced. Phonetic transcription is the representation of the pronunciation of speech sounds with the help of a special phonetic alphabet. In phonetic transcription, there is a one-to-one relationship between the letters of the phonetic alphabet and the sounds represented. The most popular phonetic alphabet is the **International Phonetic Alphabet**, devised to represent the sounds of all existing languages.

International Phonetic Alphabet (IPA)

iː READ	ɪ SIT	ʊ BOOK	uː TOO	ɪə HERE	eɪ DAY	John & Sarah Free Materials 1996	
e MEN	ə AMERICA	ɜː WORD	ɔː SORT	ʊə TOUR	ɔɪ BOY	əʊ GO	
æ CAT	ʌ BUT	ɑː PART	ɒ NOT	eə WEAR	ɑɪ MY	ɑʊ HOW	
p PIG	b BED	t TIME	d DO	ʧ CHURCH	ʤ JUDGE	k KILO	g GO
f FIVE	v VERY	θ THINK	ð THE	s SIX	z ZOO	ʃ SHORT	ʒ CASUAL
m MILK	n NO	ŋ SING	h HELLO	l LIVE	r READ	w WINDOW	j YES

Read the words, type them in and look them up in a dictionary

/ ɒlˈfakt(ə)ri/

/ ˌdʌɪəˈkrɒnɪk/

/ ˌɒnə(ʊ)mətəˈpiːə/

/ rɪˈdʌnd(ə)nsi/

/ ˌsɪntagˈmatɪk/

/ ˈhʌɪərɑːki/

/ əˈnaf(ə)rə/

Sound symbolism

> "The mute consonants represent the earth, the sibilants the sky, the vowels heaven. The mute consonants represent fire, the sibilants air, the vowels the sun… The mute consonants represent the eye, the sibilants the ear, the vowels the mind". Aitareya Aranyaka III.2.6.2.
>
> *The Upanishads. Translation: Max Muller, 1879*

Phonemes are part of the system of language and participate in the construction of the meaningful units of the higher levels but they are not themselves carriers of meaning. Exceptions to this statement – cases of **sound symbolism** – are studied by Phonosemantics.

The simplest type of sound symbolism is **onomatopoeia**, the imitation of sounds made by animals, doors slamming or the wind blowing. Examples of onomatopoeic words in English are *bow-wow, moo, murmur, bang, rumble, pop, smack, splash*, etc.

Words that share a sound sometimes seem to have something in common. It has been observed that one large group of words beginning with [b] is about barriers, bulges and bursting, another – about being banged, beaten, battered, bruised, blistered and bashed. This may also be due to analogy: if a word begins with a particular phoneme, a similar meaning is likely to be expressed by words with the same initial phoneme. A well-known claim is that if the basic word for 'house' in a given language has an initial [h], then there is a probability that a large number of words related to housing will begin with an [h] – Cf. English *hut, home, hovel, habitat*. This phenomenon is called **clustering**.

Another interesting object of study is the meaning conveyed by certain combinations of sounds and the role of the separate sounds in them. For English, placing an [m] in front of [p] seems to lend additional force to the situation of motion: Cf. *stamp* or *tamp*. This phenomenon is called "iconism".

Quiz 4

1. **The analysis of sound vibration is called:**
   a. auditory phonetics
   b. acoustic phonetics
   c. articulatory phonetics.
2. **The vocal folds are situated in:**
   a. the trachea
   b. the larynx
   c. the nasal cavity.
3. **When the vocal folds come close together, they:**
   a. vibrate
   b. stop the air coming from the lungs
   c. let the air pass freely.
4. **The consonants [s] and [z] are:**
   a. alveolar
   b. palatal
   c. velar.
5. **The sounds [w] and [j] are:**
   a. liquids
   b. glides
   c. affricates.
6. **The sounds [m] and [p] are both:**
   a. nasal
   b. voiced
   c. bilabial.
7. **The pronunciation of *handbag* as [ˈhæmbæg] is an example of:**
   a. assimilation
   b. dissimilation
   c. deletion.
8. **Stress is related to:**
   a. the relative emphasis of syllables
   b. the variation in pitch
   c. the duration of sound.
9. **In her book *Gods in the Word* Margaret Magnus suggests that the [r] sound sets r-words in motion. If this is true, it would exemplify:**
   a. onomatopoeia
   b. clustering
   c. iconism.

Further reading

Phonetics

Ashby, P., 2011. Understanding Phonetics. London: Hodder Education.

Ashby, M., Maidment, J., 2005. Introducing phonetic science. Cambridge: Cambridge University Press.

Ball, M. J., Rahilly, J., 1999. Phonetics: The Science of Speech. London: Arnold.

Davenport, M., Hannahs, S. J., 2010. Introducing Phonetics and Phonology (3rd ed.). London: Hodder Education.

Hewlett, N., Beck, J., 2006. An Introduction to the Science of Phonetics. Mahwah, NJ: Lawrence Erlbaum.

Jones, M. J., Knight, R. (eds.), 2013. The Bloomsbury Companion to Phonetics. London: Bloomsbury Academic.

Phonology

Bybee, J. L., 2001. Phonology and Language Use. Cambridge; New York: Cambridge University Press.

Burquest, D. A., 2001. Phonological Analysis: A Functional Approach. Dallas, Texas: SIL International.

Clark, J. and Yallop, C., 1995. An Introduction to Phonetics and Phonology 2nd edition. Oxford: Blackwell.

Crystal, D., 1969. Prosodic System and Intonation in English. Cambridge: Cambridge University Press.

Durand, J. and Laks, B. (eds.), 2002. Phonetics, Phonology, and Cognition. Oxford: Oxford University Press.

Halliday M. A. K., 1970. A Course in Spoken English: Intonation. Oxford: Oxford University Press.

Roach, P., 2000. English Phonetics and Phonology: Practical Course. Cambridge: Cambridge University Press.

Trubetzkoy, N. S.,1969. Principles of Phonology. Berkeley.

Phonetics and Phonology

Cipollone, N., S. H. Keiser, Sh. Vasisht., 1994. Language Files. Ohio State University Press. Columbus.

Crystal, David, 2003. The Cambridge Encyclopedia of the English Language. Cambridge University Press.

Lecture 5

The Lexicon

Words and lexemes

The word is the smallest independent building block of language possessing meaning. It is perceived by speakers as a basic unit of language and is more often than not written as a string of letters with spaces on both sides. The term "word", however, is rather loose and can sometimes mean very different things. Thus, answering a seemingly simple question such as "How many words are there in the excerpt below?" is not an easy task:

> *"When I use a word,' Humpty Dumpty said in rather a scornful tone, 'it means just what I choose it to mean – neither more nor less.'*
>
> *'The question is,' said Alice, 'whether you can make words mean so many different things.'*
>
> *'The question is,' said Humpty Dumpty, 'which is to be master – that's all."*
>
> Lewis Carroll, Through the Looking Glass

If we count all the strings of letters between intervals, we will say that there are 54 words. This is the **length of the text in running words** or word **occurrences.** But, some of the strings occur more than once. If we count the different forms (graphic words), the words will be 41: *I, a, Humpty, Dumpty, said, it, mean, to, the, question, is* occur more than once. Again, some of the strings, even if differing in form, belong to the same "word": *words* is the plural of *word*, *mean/means* and *be/is* are different forms (**word forms**) of the verbs *to mean* and *to be*, respectively; they are forms of the same lexical unit or **lexeme (**hence also the term, though seldom used, **allolexes)**. On the other hand, the two forms of *to be* in '*which is to be master'* can be argued to stand for different units of the lexicon: a modal verb and a link verb; *that's*, although a single string of letters between intervals, is a contracted form of *that is*, i.e. of two word forms. Finally, the proper name *Humpty Dumpty* should be analysed as a unity.

The stock of lexemes of a language form its **Lexicon**. Viewed with a focus on their meaning (semantics), etymology, of their semantic relations with other lexemes, their derivation from other lexemes and their types, lexemes are the object of study of a separate branch of Linguistics – **Lexicology**.

The Lexicon

The term "lexicon" is used in linguistics to refer to a/ the lexical stock of a language; b/ the mental dictionary of the speakers of a language; c/ one of the major components in systems of automatic language processing.

As a central component of theoretical models of language, the lexicon appeared in the mid-sixties, following a seminal publication by two American linguists, J. J. Katz and J. A. Fodor – "The Structure of a Semantic Theory". With this new view of the lexicon, it was no longer seen as a source of gap-fillers for structural positions (lexicalisation of grammar), but as a central component in the generation of speech, as a repository of the lexical elements of language and the structures and restrictions with which they are projected in speech (grammaticalised lexis). The proposed lexicon entries had the following structure:

1/ grammatical: providing information on grammatical categories and syntactic requirements and restrictions;
2/ semantic: based on the atomisation of meaning and the definition of the units, making use of semantic markers and semantic differentiators. Semantic markers (also known as semantic features) have a general universal character and represent, in analogy with the non-linear distinctive features of phonemes, semantic distinctive features. Semantic differentiators provide more specific additional information in the format of standard dictionary entries. Selectional restrictions define the contexts in which a lexical unit can appear.

To take an example, in one of its senses the adjective *colourful*, due to selectional restrictions, can only appear in phrases with nouns having 'physical object' as a semantic marker, e.g. *ball*, *dress*, *pen*, *car*, etc. Nouns with different semantic markers will be blocked from appearing as its arguments; Cf. **colourful knowledge* or ?*Their education was colourful.*

Entry:	COLOURFUL
Grammatical marker:	ADJECTIVE
Semantic marker:	COLOUR
Semantic differentiator:	ABOUNDING IN CONTRASTS AND OF MULTIPLE COLOURS
Selectional restrictions:	PHYSICAL OBJECT

Note that in another of its senses – 'vulgar, rude', the adjective *colourful* imposes even stricter selectional restrictions on context – it is almost exclusively used with the noun *language*.

Types of lexemes

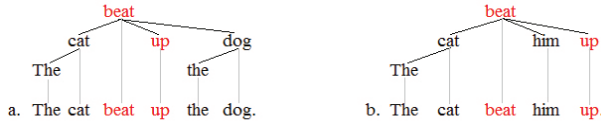
The units of the lexicon can be made up of **one word** (including **abbreviations**, such as EU, NATO, CEO or FCE) or more than one word (**multi word lexemes or multi word expressions – MWE**).

MWE are word combinations stored in the mental lexicon of the speakers of a language as an indivisible whole. They are a cover term for a considerable variety of lexicon units with different degrees of idiomaticity or structural stability. These fixed expressions are the object of study of **Phraseology**. They include:

- phrasal verbs (verbs plus adverbial particles): *come in*, *look around*, etc.
- prepositional verbs (verbs followed by a preposition and a noun phrase): *look after* (the children), *suffer from* (an infection), etc.
- light verb constructions: *have a look*, *have a go*, *take a shower*, etc.
- complex nominals: *black tea*, *Chief Executive Officer*, *weapons of mass destruction*, etc.
- idioms: *beat about the bush*, *judge a book by its cover*, *a piece of cake*, etc.
- proverbs: *The early bird catches the worm*, *There's no place like home*, etc.
- commonplaces: *I couldn't agree more*, *Great minds think alike*, *One never knows*, etc.
- routine formulas: *Good morning*! *See you*! *So long*! etc.

I. Sag et al.[1] subdivide MWE into three categories:

- fixed, e.g. *by and large*, *every which way*;
- semi-fixed (with possible variation of form or lexical variation), e.g. *car park/ car parks* or *keep it under the carpet/under the rug*;
- syntactically flexible, e.g. *pick up someone/pick someone up*, *keep tabs on/ tabs were kept on*.



According to these authors, the number of MWE in the lexicon of English is no lesser than the number of one-word lexemes. Because MWE are particularly numerous in specialised texts, with the development of professional areas and respective terminology these lexemes constantly increase.

[1] Ivan Sag, Timothy Baldwin, Francis Bond, Ann Copestake, Dan Flickinger 2002. Mutiword Expressions: a Pain in the Neck for NLP. In; Lecture notes in computer science, vol. 2276, pp 1–15.

The Open-choice Principle and the Principle of Idiom

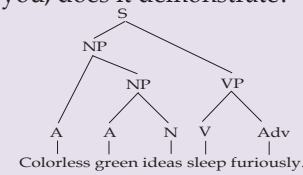
The units of the lexicon, be they one-word lexemes or MWE, combine with other units in accordance with two principles: the Open-choice principle and the Principle of Idiom.

The **Open-choice principle** states the existence of structures with positions to be filled at choice by lexicon units, with a large range of options at each step of the generation. This principle is in harmony with Hockett's Principle of Productivity and the view of language as an open-ended system allowing the production of an unlimited number of different messages.

Taken to an extreme, this principle would allow the appearance of any lexical unit having one of the grammatical markers below in the respective position. Along with fully acceptable strings, such as e.g. *Modern innovative techniques function better*, it would permit the generation of highly improbable ones, as e.g. *Happy sad wardrobes sing adequately*. Blocking such semantically anomalous statements is one of the primary functions of the Lexicon.

The string below is a famous example from Noam Chomsky's 1957 book "Syntactic Structures".

What, according to you, does it demonstrate?

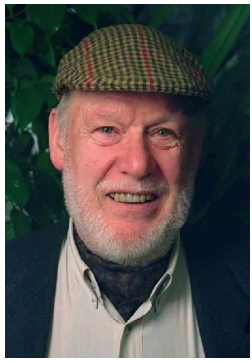


In their use, lexemes are subject to a number of restrictions of a different nature: semantic, syntactic, or simply imposed by convention. This latter restriction was formulated by John McHardy Sinclair (1933–2007) as "**The Principle of Idiom**"[2]. This principle states that a language user has available to him, along with lexemes, a large number of collocations – semi-preconstructed phrases that constitute simple choices even though they might appear to be analyzable into segments.

[2] J. McH Sinclair, 1991. Corpus, Concordance, Collocation. Oxford University Press, p. 110 ff.

An example: According to the Open-choice principle, *manners*, being a noun, can be preceded by a determiner or an adjective-attribute and can be followed by a verbal form or a preposition. The Idiom principle states that while this is true, *manners* nevertheless most often collocates with only a handful of adjective attributes (*easy, good, happy, pleasant, unaffected*) and is most often followed by the link verb *to be*.

Collocation and colligation



The Lexicon is not a simple list of lexemes: it provides information on the meaning of its units and the ways in which they can be used – what their possible forms are or how they can combine with other units. In accordance with the Principle of Idiom, these "prefabricated" unities or **collocates** are favoured by native speakers of a language in preference to alternative, seemingly equivalent combinations. The violation of collocational preference can lead to the unacceptability of otherwise (grammatically and semantically) well-formed combinations.

The Principle of Idiom was extended by other linguists to cover regular co-occurrences of lexemes with grammatical content, syntactic position and function, with patterns of information structure. The cover term used for these types of co-occurrence is **colligation**.

In his book "Lexical Priming: A New Theory of Words and Language"[3], the British linguist Michael Hoey identifies three major **types** of colligation:

- the relation between a lexical item and its grammatical context: the noun *sister* is usually preceded by a personal pronoun (*my, your, his/her*);
- the relation between a lexical item and its syntactic function or syntactic function: English nouns can have many syntactic functions. Typically, however, *manners* appears in subject phrases preceded by a determiner or attribute: *Their manners are not quite equal to his;*
- the relation between a lexical item and its position in discourse – as theme (known information) or rheme (new information).

[3] Michael Hoey 2005: 49–52.

In order to use language correctly, speakers should therefore have access, in their mental lexicon, to a large **volume of information**, including:

- A set of lexemes – typically, between 12 000 and 15 000 for school leavers and around 20 000 for university graduates.

Did you know that?
Shakespeare's active vocabulary was larger than that of any writer – over 24 000 words!
(Homer used only around 9 000).

How rich is you vocabulary in your mother tongue?
How rich do you think is your English vocabulary?

- Knowledge of the meaning of the lexemes, their different senses and their word forms. Note that not all senses of a word make use of the same word forms!

This is the entry for *alcohol* in the Merriam-Webster Dictionary:

1 a: ethanol especially when considered as the intoxicating agent in fermented and distilled liquors
b: drink (such as whiskey or beer) containing ethanol
c: a mixture of ethanol and water that is usually 95 percent ethanol
2: any of a class of organic compounds that are analogous to ethanol and that are hydroxyl derivatives of hydrocarbons.

Find examples of each use. In which of its senses does the lexeme appear with the word form *alcohols*?

- Knowledge of the class of words to which a lexeme belongs, its forms, the grammatical restrictions it imposes on its context: *sleep*, *read* and *buy* all belong to the word class Verb, but they differ in their additional (subcategorisation) features: *sleep* does not take objects (it is intransitive), *read* takes a direct object (it is transitive) and *buy* can take both a direct and an indirect object (it is ditransitive).
- Knowledge of the lexeme's typical collocational and colligational restrictions.

Which of the verbs *miss*, *get*, *do* and *make* collocate with the expressions below?

1. ___ a goal; 2. ___ peace; 3. ___ a home; 4. ___ an appointment; 5. ____ a lesson; 6. ___ the cooking; 7. ____ ready; 8. ___ progress; 9. ____ nothing; 10. ____ an effort; 11. ___ one's best; 12. ___ furniture; 13. ___ the shopping; ___ 14. _____ trouble; 15. ___ someone a favour.

Lexical semantics

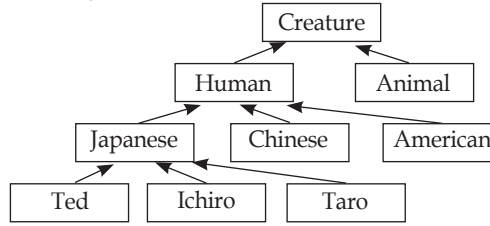
The **lexical meaning of the units of the lexicon and the relations between them** are studied by a field of linguistics called Lexical semantics. It is concerned with the identification of the possible interpretations of lexical items, their formal representation and the relations holding between them.

Semantic features and matrices.

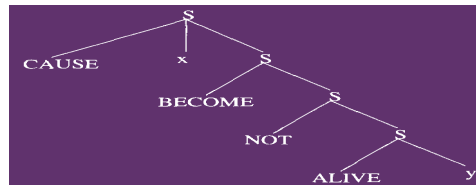
The lexical meaning of a word can be described by decomposing it into constituent parts, a process called **lexical decomposition**, and then comparing it to the constituent makeup of other words.

For example, in its first sense, the noun *girl* is defined in most dictionaries as "a female child from birth to early womanhood". This definition can be reformulated as a set of **semantic features** or **components**: [Female] and [Young]. Because girls are not any kind of young female, but a representative of humans, we could add to these the feature [Human]. The set of semantic features which are sufficient to differentiate a lexical unit from other units form its **semantic matrix**.

Analysing *girl* and *boy*, we end up with similar semantic matrices, which only differ in one semantic feature: [[Human], [Young], [Female]] vs. [[Human], [Young], [Male]]. A more economical way of presenting this opposition would be to list only one of the differing features and provide it with a positive or negative value. *Girl* will then have the feature [+Female], while *boy* will be negatively marked for it: [-Female]. *Woman* can be described as [+Human] (to differentiate it from, say, *doe*:[-Human]). We could, of course, enlarge the matrix by adding the feature [+Mammal] to differentiate *woman* and *doe* from *crocodile*, *ladybird* or *salmon*, or [+Animate] to set all of these apart from *chair*, *mushroom* and *sand*. Such additions are not necessary, though: the presence of the feature [+Human] implies the presence of [+Mammal] and [+Animate]. Semantic features can be presented as hierarchies where the more specific features "inherit" information from the ones standing higher in the hierarchy. A semantic matrix need not contain any features beyond those that cannot be inherited.



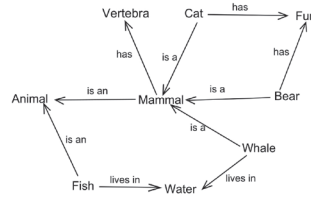
As far back as the 1960s, two American linguists working in a framework called Generative Semantics, George Lakoff (born 24 May 1941) and James McCawley (1938–1999), proposed a set of what they called **atomic predicates** to describe the lexical semantics of verbs. A famous example is McCawley's[4] semantic representation of the verb *to kill*:



Atomic predicates can also be used to describe the semantics of the causative meanings of some verbs:

The soldiers marched.	The commander marched the soldiers
march (x)	march (x,y) or x CAUSE y to march

Semantic components play an important role in helping organise the lexicon into classes and hierarchies and in creating lexical databases (networks of interrelated lexical units) – Cf. a subpart of the lexical data base of the English language known as **Word Net**:



If they are set out so as to be valid cross-linguistically, semantic components can serve as a metalanguage for conceptual representations. These are the basis for the establishment of (more or less) universally valid relations known as semantic networks.

Look at the graph above. What relations hold between:
Cat and Mammal?
Cat and Fur?
Fish and Water?
Why are the lines in the graph directed?
Could a line be drawn between Whale and Bear? What would the relation be? Would it be a directed or non-directed line?

[4] Cf. McCawley 1968

Lexical relations

The relations into which the units of the lexicon can enter are largely (though not solely) based on the makeup of their semantic matrices. The relations that are most important for the establishment of lexical or word net relations are synonymy and hypernymy/hyponymy.

Synonymy is a symmetric relation between two lexicon items based on near identity of their semantic matrices, e.g. *girl* and *lass*.

Troponyms are near synonyms which differ in some additional components, as e.g. the manner of carrying out the same action – for example, different ways of cutting: *slice*, *trim* or *carve*.

Antonymy is, again, a symmetric relation, but here the two units compared differ in one or more of their semantic components, as e.g. *woman* and *man,* [+/Female].

Hyponymy is an asymmetric, "transitive" relation between units/nodes, one of which has additional semantic components standing lower in the semantic hierarchy. *Girl* is a hyponym of both *female* and *human*, which are hyponyms of *animate*. This same transitive series, if reversed, is called **hypernymy**. *Mammal* is the hypernym of *elephant*, *whale* and *man*.

Holonymy and **meronymy** are based on part-whole relations. Thus, as *eyes* are normally situated in a head, *eye* is a meronym of *head*. Inversely, *head* is a holonym of *eye*.

Even if *girl* and *boy* can be viewed as antonyms because of the opposition of the components [+Female] – [- Female], they are nevertheless alike in that they occupy a similar position in the semantic hierarchy. In this respect, they can also be said to be **coordinate terms**. Coordinate terms can form **semantic/lexical fields**.

What semantic relation(s) hold between:
good and bad?
bird and chicken?
cart and bike?
car and automobile?
hand and finger?
sip and gulp?

Consider the following phrases which demonstrate clear deviations form semantic restrictions in the lexicon:

to feel blue
a broken heart
a stench of failure
a bubbly personality

Do you understand them? Why are they possible? What figure of speech are they examples of?

Polysemy and homonymy

Some lexemes have more than one meaning. There usually is a central meaning, from which the others have diverged in different ways. Thus, in one of its meanings, the central one, *crane* is a bird with a long neck. In another of its meanings, it is a piece of equipment with a long rope and chains used to lift heavy objects. There is also a verb *to crane* – to strain one's neck. Lexemes like *crane* which have more than one different yet related meanings are **polysemous**.

There are lexemes which look the same (homographs), sound the same (homophones) or both of these (true homonyms), but this is quite accidental – they have meanings that are different and unrelated. Such lexemes are **homonyms**. The noun *bear* and the verb *to bear* are good examples of homonymy.

In some cases it is not easy to decide whether we are dealing with two meanings of a lexeme or with two different, homonymous lexemes. Etymology does not always help because **semantic shifts** can increase the semantic distance between initially related lexical units, so that they are felt to be homonymous.

Look up the strings listed below in a dictionary and list their different meanings.

Which are the cases of polysemy?
Which are the cases of homonymy?

fluke
bank
wood
check
park
rock
grave
well

What lexical/semantic relations hold between the following lexemes?

house – roof
girl – woman
hen – goose
aisle – isle
gobsmacked – flabbergasted
warn – worn
fair – fare
bore – boar

Quiz 5

1. **'Word form' is a synonym of:**
   a. lexeme
   b. occurrence
   c. allolex.
2. **For the noun *ball*, 'physical object' is a:**
   a. grammatical marker
   b. semantic marker
   c. semantic differentiator.
3. **The MWE *to give a lift* is:**
   a. fixed
   b. semi-fixed
   c. syntactically flexible.
4. **A colligation is the co-occurrence of:**
   a. two lexemes
   b. a lexeme and grammatical content
   c. a syntactic position and a syntactic function.
5. **Near synonyms which differ in some additional components are called:**
   a. hyponyms
   b. meronyms
   c. troponyms.
6. **Homographs are lexemes which:**
   a. sound the same
   b. look the same
   c. have related meaning.
7. **The lexical relation between room and ceiling is one of:**
   a. coordination
   b. holonymy
   c. hypernymy

Further reading

Andrews, Avery D., 1989. Lexical Structure. In: In: Frederick J. Newmeyer (ed.) Linguistics: The Cambridge Survey. I. Linguistic Theory: Foundations.

Cipollone, Nick, Steven Hartman Keiser, Shravan Vasisht (eds.), 1994. Language Files. Materials for an Introduction to Language and Linguistics. Ohio State University Press, Columbus.

Emmoray, Karen D., Victoria A. Fromkin, 1988. The Mental Lexicon. In: Frederick J. Newmeyer (ed.) Linguistics: The Cambridge Survey. III Language: Psychological and Biological Aspects. Cambridge University Press.

Halliday, M. A. K., Collin Yallop, 2007. Lexicology: A Short Introduction. London: Continuum.

Hans, Patrick W., 2007. Lexicology: Critical Concepts in Linguistics. Routledge.

Katz, Jerrold J., Jerry A. Fodor, 1964. The Structure of a Semantic Theory. In: J. A. Fodor, J. J. Katz, (eds.) The Structure of Language. Readings in the philosophy of Language. Englewood Cliffs: Prentice-Hall.

Leech 1981. Geoffrey N. Leech. Semantics: The Study of Meaning (2nd ed.). Harmondsworth: Penguin.

Lyons, John, 1977. Semantics. Cambridge: Cambridge University Press.

McCawley, James, 1978. The Role of Semantics in a Grammar. In: Emmon Bach, R. Harms (eds.) Universals in Linguistic Theory. New York: Holt, Rinehart and Winston, pp. 125–170.

Sinclair, John, 1991. Corpus, Concordance, Collocation. Oxford University Press.

Singleton, David, 2000. Language and the Lexicon: An Introduction.

Lecture 6

The Morphemic Level

Morphology

The unit of the lexical level, the lexeme, is made up of one or more words. Every lexeme and every word are also made up of smaller units, or constituents – phonemes. Phonemes are the smallest discrete building blocks of the units of the lexical level; they are not, however, their *immediate constituents*. There exists a level of structure lying between the phonological and lexical levels – the morphemic level, the units of which are called morphemes.

The morpheme – the immediate constituent of the lexeme – is the smallest unit of language carrying meaning.

The branch of linguistics studying the morphemic structure or lexemes and words is called **Morphology** –a term coined by the German linguist August Schleicher (19.02.1821 – 6.12.1868) – from Greek *morphḗ* (shape, form) and *logos* (science): **the study of shape**. Along with morphemes, morphology also has as its object the grammatical categories marked by word forms, as well as the grammatical classification of lexemes. These are discussed in Lecture 8. Lecture 7 is devoted to morphological patterns of word formation – along with other means of enriching the lexicon.

Although the term 'morphology' itself only dates from the mid-19th century, our records of morphological analysis have a much longer history and start with the 6th or 5th century BCE Indian grammarian Pāṇini, also considered to be the father of linguistics. His treatise on Sanskrit grammar contains 3, 959 rules and is probably the most advanced linguistic work of pre-20th century science. Because of its precise definitions and logical structure, the Dutch linguist Frits Staal compared it to the Turing machine (a mathematical model of computation invented in 1936 by Alan Turing, an English mathematician, logician and computer scientist).

Modern morphology, as part of modern linguistics, begins with F. de Saussure's Course and with the structuralist view of language briefly presented in Lecture 3. In the discussion that follows, several principles of structuralism are applied:

- the definition of language as a system, the elements of which are organised in a structure, a complex set of oppositions;
- the abstract hierarchical organisation of language;
- the asymmetry between the complexity of the sign and of the signifier;
- the relative independence of the synchronic and diachronic perspectives.

The classification of morphemes

Morphemes **can be classified** according to position, lexical/ grammatical content or function.

The morpheme carrying the basic lexical meaning of the lexeme and its forms is the **root.** The root has a central position in the word form and is usually presented (like most other morphemes) thus: *-root–* to show, first, that it is a morpheme, not a lexeme or a paradigmatic form of a lexeme, next – that other morphemes can be attached to its left and/or right.

All lexemes have roots. Some have only a root, i.e. one morpheme. They are said to be **monomorphemic.** This group includes most unchangeable lexemes belonging to the so-called closed classes of the article, conjunction, preposition, particle and interjection. Lexemes belonging to the class of nouns, verbs, adjectives and adverbs are **polymorphemic**: they are made up of one or more roots plus at least one **affix**. Their structural position with respect to the root divides affixes into prefixes, infixes, suffixes or circumfixes.

The **prefix** is a morpheme positioned to the left of the root (and of another prefix)

The **suffix** is a morpheme positioned after the root (and another suffix).

> Consider the lexeme DECENTRALISE. Can you identify its root? How many affixes can you identify? Are they prefixes or suffixes?

Some languages (Arabic, Khmer) have a third type of affix, placed **inside the root** and modifying its lexical or grammatical meaning: the **infix**. English has very few infixes, and they are so marginal that they are not even acknowledged by most works on the subject: a/ the *-iz–* and *-izn–* of hip**-**hop slang: *hizouse* instead of *house*, *shiznit* instead of *shit*; b/ the ironical infix *-ma–* creating pseudo-sophisticated words, such as *edumacation*.

Note that the **linking morphemes (**or **interfixes)** which participate in the formation of complex words are not infixes: they are not inserted into the root, but rather link two roots: Cf. *-o–* in ethnic names (*Anglo-Saxon*), scientific terms of Greek and Latin origin (*thermometer*), the *-s–* in *statesman* (diachronically, a fossilised genitive).

A fourth type of affix is the **circumfix**. It is made up of two parts, which are placed **around** the root or around the root and one or more other affixes. The German past participle (*ge–* …*-t* for regular verbs) is a circumfix – Cf. *spielen* – ***ge****spiel****t***. Similarly, Dutch *spelen* – ***ge****speel****d***. In older forms of English, the present participle was formed using the circumfix '*a– -ing*':

Gather ye rosebuds while ye may
Old time is still a-flying.

Functionally, morphemes are word-formative or form-formative.

Word-formative morphemes carry the lexical (dictionary) meaning. They can serve to form new lexemes and include lexical prefixes and suffixes, most infixes, linking morphemes. Another term for them is **derivational morphemes**. The process of creating new words with the addition of affixes is called **derivation**.

Form-formative morphemes serve to make up the paradigmatic forms of the lexeme. They are the markers of the grammaticalised meanings in a language, being exponents of the members of its grammatical oppositions. To the exception of the initial element of circumfixes and of some infixes, grammatical morphemes are suffixes**.** Another term for these affixes is **inflectional morphemes or inflexions**.

Suffixes, even if they are not inflexions, **have a mixed lexico-grammatical character –** because apart from forming new lexemes, they define the word class (part of speech) to which the lexeme belongs. The suffixes *-able*, *-al* or *-ful* combine with verbal or nominal roots to form lexemes belonging to the class of adjectives; *-ism*, *-ity* and *-ation* form nouns, *-ise* and *-ate* form verbs.

The **stem** (Greek *thema*) is the set of word-formative morphemes of a lexeme. It is obtained when the grammatical morphemes are separated from the form of the lexeme. Depending on the number of morphemes they include, stems can be monomorphemic or polymorphemic.

The term **formative** is used for what remains when the root and the prefixes/infixes are set apart from the suffixal part, be it grammatical or lexical, or both. The term is particularly useful for boundary cases between lexical and inflectional endings – as e.g. in *a crowded room*, where one might have doubts whether the *-ed* is part of a derived adjective or a participial inflexion.[1]

Some authors consider linking morphemes to be "**semi-morphemes**" – because they carry neither lexical nor grammatical meaning. Other semi-morphemes are strings in lexemes which combine with an identifiable morpheme but are themselves neither lexical nor grammatical – such as *cran–* in *cranberry*.

The morpheme is an abstract unit. Just like the phoneme, which can have variants, the morpheme can appear in the form of different **morphs**. Thus, 'caught' consists of the morphs 'cau-' and '-t, 'forgotten' – of the morphs 'forgot-' and '-en'. Both [t] and [n] are allomorphs of the Past Participle morpheme; so too are the [d] of 'sailed' and the [ɪd] of 'waited'. The relation

[1]A common test: if the *-ed* word form can be modified by the adverb *very*, it is an adjective; otherwise, it is a verbal form – compare: *a very crowded room./*a very heated room.*]

between morphs and morphemes is that of **allomorphy.** If morph x is a member of morpheme y, it is an **allomorph** of that morpheme.

Clitics are grammatically independent lexemes which are phonologically dependent on another word and are pronounced like affixes. Some authors view them as a special case of morphemes, functioning at the phrasal or clausal levels.

There are four types of clitics:

A clitic that precedes its host is called a **proclitic** (French: ***j'l'****ai vu*).

A clitic that follows its host is called an **enclitic** (English *I told****ja***, Bulgarian *Видя* ***ли****?*)

A **mesoclitic** appears between the stem of the host and other affixes – Cf. Portuguese *Ela leva-****lo****-ia* ('she take it CONDITIONAL SUFFIX, i.e. *She would take it*).

Finally, an **endoclitic** splits the root apart and is inserted between the two parts. The meaning of endoclitics, unlike that of infixes, is unrelated to that of their host word. Only three languages have endoclitics: Pashto – (an Indo-European language spoken in Afghanistan and Pakistan), Udi (the Lezgic branch of the Northeast Caucasian language family) and Degema (a Nigerian language). The example below is from Pashto:

*wakhist-****el*** *copy mee aw kitaab*
bought notebook CLT-1sg and book

If there are two, three or more clitics in a clause, these tend to appear in clusters – Cf. an example from Bulgarian:

Камен дал	*[* ***ли***	***му***	***я***	***е]***	*книгата?*
Kamen give	INTERR	to-him	it	AUX	book-the

In clitic clusters, each clitic has a well-defined position: the order in which the above clitics appear cannot be changed.

Clitics and clitic clusters can undergo forward (leftward) movement in the clause, known as **clitic movement**:

Някой е взел чантата ***ми.***	*Някой* ***ми*** *е взел чантата.*
Someone AUX taken bag my	Someone my AUX taken bag.

Zero morphemes

Polymorphemic forms of lexemes which contain an inflexional morpheme mark the value of a grammatical category. The morpheme *-ed* in *decided* is a marker of past tense; *-s* in *books* is a marker of plural number.

Grammatical categories are formed by **oppositions of values**: the English category of Tense has two values: Present and Past; the Serbian category of Number has three values: Singular, Dual and Plural. Of these values, one is usually more general, the other/others are specific. While present tense forms can also be used to speak of past or future events, past tense forms refer to past events only. The specific forms have special formal markers and are said to be **marked**:

decide- +	*-ed*
Root	INFL (Past Tense)

The unmarked forms can either have a special marker or coincide in form with the root/stem. In the latter case, we say that they have a zero-inflection:

decide- +	*-0*
Root	INFL (Present Tense)

Accepting the presence of a phonetically null zero inflection is motivated by the following considerations:

- the desirability to match the opposition of values of the grammatical category with an opposition of inflectional markers:

Present tense	Past tense
-0/-s	*-ed*

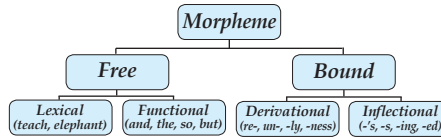
- the impossibility to mark a grammatical meaning with the root morpheme:

WF (lexical)	FF (grammatical)
decide-	*– ed / s/0*

- the invisibility of morphemes at the phrasal or syntactic level.

On the so-called 'bound' and 'free' morphemes

Some linguists speak of **bound** and **free** morphemes. Free morphemes are defined as morphemes which can be integrated into discourse 'without having anything attached to them', as e.g. 'dog' (Cf. Payne 1997), while bound morphemes cannot stand on their own:



This approach, while at first glance offering a simpler view of language structure, creates a number of problems – apart from being incompatible with the structural approach to language:

- If *elephant* is one morpheme, it will not carry grammatical information; it would be a freely functioning lexical morpheme, a root. However, only grammatically formed phrases can function on the syntactic level.
- Alternatively, if *elephant* is a fused lexico-grammatical morpheme, it will be marked for singular number; but then, a plural marker will be incompatible with this form: a single word form cannot be marked for two different values of the same grammatical category. Hence *elephants* would not be an admissible form.

These problems can easily be avoided if a traditional view of morphemes is adopted and they are all analysed as units of the morphological level, equally incapable of free appearance in the syntax.

List the component morpheme(s) of the following words. State how many morphemes they contain and what type of morphemes they are:

Celebrity –
Untrue –
Perfection –
Writing ––
Untrained ––
Disinfectant ––
Classrooms –
Justification –
Activity –
Antisocial –
Decapitate –
Recharges –
Thoughtlessness –
Irrationality –

The hierarchical structure of words

Affixes do not attach to roots or stems indiscriminately. As a rule, affixes attach to roots/stems which are structural parts of a particular part of speech. Thus,*-able* attaches to verbal stems:

suit– + *-able* = *suitable*

(try also with *adjust-*, *break-*, *compare-*, *debate-*)

but not to adjectival or nominal ones:

suite– + *able* = **suitable*

sweet– + *-able* = **sweetable*.

Try also with: house-, beauty-, nice-, cool-, hot-, aloof-

After affixation, the resulting forms also belong to a particular class of words. Adding *–able* to a verbal stem always yields adjectives.

Looking at words with more than one affixal morpheme, such as *unusable*, this categorial markedness of morphemes means that the structure of affixally derived words is not flat, but rather hierarchical, formed in consecutive steps:

un + *use* + *able*

The prefix *un–* in the meaning of 'not' attaches to adjectives and the process yields new adjectives (*unkind, unwise, unhappy*). The suffix *-able* attaches to verbs to yield adjectives. The adjective *unusable* can thus only be formed in two steps:

1/ use_{v}– + *-able* = $-usable_{adj}$

2/ *un–* + $-usable_{adj}$ = $unusable_{adj}$

The structure of *unusable* can be presented in the form of a hierarchy presenting the relative closeness of the bonds between the morphemes:

[A [un-][A [V-use-][-able]]]

In the word *reusable*, however, both *re–* and *-able* attach to verbal roots. Which should attach first? *Re–* cannot attach to *-usable*, because it can only attach to verbs (Cf. *redo, rewrite, reconsider*), while usable is adjectival. It can only attach to *-use*. On the other hand. *-able* can attach to *reuse-*, which is verbal. Thus, the steps in the formation of reusable are:

[A [V [RE-] [-USE-]] [-ABLE]]

Some words are ambiguous and can be analysed as formed in more than one way:

unlockable: a. *un-* + *-lockable* or b. *unlock-* + *-able*.

a. [A [UN-] [A [V –LOCK-] [-ABLE]]]
b. [A [V [UN-] [-LOCK-]] [-ABLE]]

The two *un-* morphemes are different prefixes, homophones. One has the meaning 'undo' and attaches to verbal morphemes; the other has a general negative meaning – 'not' – and attaches to adjectival morphemes.

Some prefixes can become so productive that their combinatorial possibilities can be extended to more than one category. Thus, *pre–* can attach to verbal morphemes to form verbs (*predetermine*), to nominal morphemes to form adjectives (*prewar*), as well as to adjectival morphemes to form new adjectives (*prehistoric*). Most affixes, however, have stricter categorial requirements.

What are the categorial requirements of the suffixes *-al*, *-ise* and *-ation* in *globalisation*?

What are the categorial requirements of the prefixes *un-*, *re-*, *de–* and *pre-*?

Structural types of word forms and morphemes

Word forms are variations within the lexeme. They have a common lexical meaning but differ in their formal markers and grammatical meaning.

For practical purposes, one of the word forms is taken as representative, basic. It is usually referred to as the **base form** of the word, or the **lemma**. It is this word form that appears in dictionaries where the lexical meaning of a word is defined.

In Indo-European languages, most word forms have a common root and differ in their formatives / inflexions. However, **suppletive forms**, i.e. forms of a lexeme having different roots, were typical of the Indo-European protolanguage, and some cases of suppletivism have been preserved. In English, s**uppletivism** can be observed in forms expressing case (*I-me*), number (*I, we*), or person (*I, you, he/she*)

In most Indo-European languages, there are two structural types of word form: synthetic and analytical.

The **synthetic** (simple) form is an accentual and graphic unity consisting of a stem and a formative, where the grammatical meanings are expressed within the formative: *books, nicer, worked*.

The **analytical** (complex) form is an idiomatic combination between one or more auxiliary words and a lexical word, which has the structural status of a word form.

Most English lexemes can be viewed as sets (**paradigms**) of forms, where each form is inflectionally marked for one or more grammatical category. The number of categories depends on the class to which the lexeme belongs. The grammatical categories of the English verb are: Person/Number, Tense, Aspect, Correlation, Mood, Modality, Voice. The form 'kicked' is marked for (Past) Tense: *kick-* + *-ed*. Similarly, the form 'kicks' is marked for the 3rd p. sg, (Present Tense): *kick–* + *-s*. *Kicks* and *kicked* are forms of the verb *to kick*, paradigmatic members of the verbal lexeme *kick*. The category of Aspect is marked with the discontinuous marker: *be–* + *-ing*: *is kicking*. This seeming word group is made up of only one lexical morpheme (the root morpheme *kick-*) and two inflectional markers (for Progressive Aspect and for the Present Tense, 3rd person Singular). The marker for Aspect, in spite of its more complex form, has the same function and status as the marker for Person/Number/Tense. It is an example of a type of marker, called '**discontinuous (or analytical) marker or morpheme'**.

Morphological typology

A typology is a classification according to general type. One of the first proposals for a morphological typology of languages was made in 1989, by Bernard Comrie (born 23 May 1947).

B. Comrie's typology is based on two indexes: synthesis and fusion. Along these indexes, languages form a continuum. On the basis of a combination of these indexes, he sets out three morphological language types: 1/ isolating, 2/ agglutinating and 3/ inflectional or fusional.

The **index of synthesis** has to do with **how many morphemes tend to occur per lexeme**. This index defines a continuum from **isolating languages** at one extreme to highly **polysynthetic languages** at the other.

A strictly isolating language is one in which every lexeme consists of only one morpheme (e.g. Chinese).

A highly polysynthetic language is one in which lexemes tend to consist of several morphemes (e.g. Yup'ik Eskimo):

Tuntussuqatarniksaitengqiggtuq
Tuntu-ssur-qatar-ni-ksaite-ngqiggte-uq
Reindeer-hunt-FUT-say-NEG-again-3SG:IND
He had not yet said again that he was going to hunt reindeer.

The **index of fusion** has to do with the **degree to which units of meaning are 'fused' into single morphological shapes**. A highly **fusional language** is one in which one morpheme can simultaneously mark several categories, e.g. Spanish '*o*' in *hablo* is a marker of the indicative mood, the 3rd person, singular, past tense and perfective aspect. If any of these grammatical values changes, the form will also change. Turkish is a language where each grammatical meaning is expressed by a separate morpheme – it is a highly **agglutinative language**. **Here again, languages form a continuum**.

In English, forms like *anti-dis-establish-ment-ari-an-ism* point to agglutination (but then, such forms are as a rule of Latin origin). On the other hand, the 'strong' forms 'sang', 'thought', 'brought' speak of fusion.

Quiz 6

1. **The morpheme is:**
   a. the immediate constituent of the word/lexeme
   b. the final constituent of the word/lexeme
   c. the immediate constituent of the phrase/clause sentence.
2. **A polymorphemic form is made up of:**
   a. at least one root and at least one affix
   b. two or more roots
   c. two or more affixes.
3. **The form DECENTRALISE is made up of:**
   a. 3 morphemes b. 4 morphemes c. 6 morphemes.
4. **Allomorphs are:**
   a. different morphemes
   b. phonetic variants of a morpheme
   c. spelling variants of a morpheme.
5. **A clitic which splits the root apart is called:**
   a. a proclitic b. a mesoclitic c. an endoclitic.
6. **Zero morphemes are:**
   a. phonetically null b. semantically null c. structurally null.
7. **Forms of a lexeme with different rots are called:**
   a. free morphemes b. lemmas c. suppletives.
8. **A language in which every lexeme consists of one morpheme is called:**
   a. isolating b. fusional c. polysynthetic.
9. **The marker for the Progressive in HAS BEEN RUNNING has the status of:**
   a. a set phrase b. a syntactic structure c. an analytical morpheme
10. **Which of these statements is NOT true?**
    a. Prefixes are not sensitive to the grammatical nature of roots and stems.
    b. Suffixes are sensitive to the grammatical nature of roots and stems.
    c. Prefixes and suffixes are added to roots in a fixed order.

Further reading

Aronoff, Mark., 1994. Morphology by itself. MIT Press, Cambridge.

Beard, Robert., 1995. Lexeme-Morpheme Base Morphology. SUNY Albany Press, Albany.

Carstairs-McCarthy, Andrew, 1992. Current Morphology. Linguistic Theory Guides, Routledge.

Crystal, David, 2003. The Cambridge Encyclopedia of the English Language. Cambridge University Press.

Halpern, Aaron, 1995. On the Placement and Morphology of Clitics. CSLI Publications, Stanford.

Lieber, Rochelle, 2010. Introducing Morphology. Cambridge Introduction to Language and Linguistics. Cambridge University Press.

Molhova, Jana, 1981-94. On the Nature of the Morpheme in Contemporary English. English Studies. The Kliment of Ohrida University of Sofia.

Zwicky, Arnold. 1985a. 'Clitics and Particles.' *Language* 61: 283-305.

Zwicky, Arnold, 1985b. 'How to Describe Inflection.' Proceedings of the Berkeley Linguistics Society 11: 372-386. Berkeley, California.

Lecture 7

Enriching the Lexicon

There are various means of enriching the lexical stock of a language. Some of these means make use of the lexical and structural means of the system of that language, others also involve "borrowing" lexemes or word formation patterns from other languages. In either case, the new lexeme becomes part of the system of the (target) language and answers to, or adapts to, the rules of its phonological, morphological and syntactic system.

Word formation is a means of forming new words by using the linguistic means that are available in a language: already existing in that language lexemes, morphemes and word formation processes.

Most languages make use of one or more of the following means of word formation:

1. Vowel gradation
2. Suprasegmental modification.
3. Reduplication
4. Composition
5. Affixation
6. Back formation
7. Contraction and shortening
8. Blending
9. Conversion

These means of word formation have unequal importance in the systems of different languages and at different stages in the historical development of one particular language.

Vowel gradation is a process of creating new lexemes or forms by changing features of the stem vowels. The inherent vowel of most Proto-Indo-European syllables was a short *e*. One type of vowel gradation – the process whereby an inherent short *e* either changed to become short *o*, long *ē*, long *ō* or became null is called Ablaut. Examples of Ablaut in English are the pairs *strike / stroke* and *sit / set*.

Suprasegmental modification is the signaling of a morphological operation by means of the tone or stress pattern of the word. English makes use of changes in the stress pattern to signal the difference between nouns and verbs. (e.g. *pErmit/permIt*). Already in Primitive Germanic, it was well established for verbs to have stress on the root syllable, and for nouns – on the first syllable. Hence such patterns as *'accent* (noun) – *ac'cent* (verb), *'conduct* (noun) – *con'duct* (verb), etc.

Verbs formed by conversion were sometimes differentiated by **voicing their final consonant**. In some cases, this is reflected in the spelling of the two words (e.g. *advice – advise*, *safe – save*, etc).

New words can be formed by repetition of existing roots or stems. This process of word formation is called **reduplication.**

Many languages – from languages of the Indo-European group to Bantu and Austronesian languages – make use of reduplication, with varying productivity.

Lexemes formed by reduplication are usually made up of a **base** (the segment to be copied) and one, seldom two, **reduplicants**.

The reduplicant can copy the base entirely, as in *fifty-fifty, knock-knock* or *bye-bye*. These are examples of exact or **total reduplication**.

Reduplication can also be **partial**: the reduplicant may differ from the base in one or more phonemes.

In partial reduplication, the reduplicant is usually devoid of meaning. There can however be exceptions, as e.g. *singsong*, where both roots can be viewed as bases, or *boogie-woogie*, where neither root is a meaningful base.

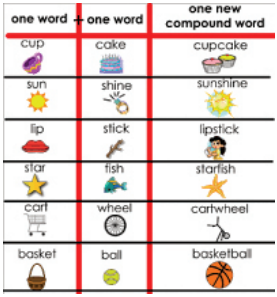
Partial reduplication can be vocalic (ablaut) or rhyming. The English language abounds in examples of both.

Vocalic reduplication can be observed in: *hip-hop, pitter-patter, zig-zag, ding-dong*. Note that the first vowel is almost always a high vowel, while the second one is a low vowel. In most cases, it is the reduplicant that contains the low vowel because it usually follows the base, but this is not always so: in *dilly-dally* the reduplicant precedes the base and thus carries the high vowel.

Rhyming reduplication involves either a change of the initial consonant, as in *hokey-pokey, super-duper, teenie-weenie, walkie-talkie, ragtag,* or the addition of one initial consonant – as in *easy-peasy,* or a consonantal group – as in the very productive American ***shm*-reduplication**: *baby-shmaby, fancy-shmancy*, etc.

Apart from lexical derivation, reduplication can also have grammatical functions – typically, as a marker of plurality or degree.

Composition is one of the most productive means of word formation in modern English. It is a means of forming new words by merging two or more roots into one: *short-sighted, waterfall, cupcake, basketball*, etc. The process of composition results in the creation of compound lexemes, or **compounds**.



Compounds differ from free combinations of syntactically linked lexemes both semantically and structurally. As O. Jespersen[1] put it, in compounds "the meaning of the whole cannot be logically deduced from the meaning of the elements separately". Their meaning tends to be more specialised. The constituents of a compound lexeme have a fixed order, do not allow the insertion of morphemes between the constituents and are inflected as a unity. Normally, it is the last root which carries the morphological markers of the whole: *cupcakes, banknotes, etc*. Stress usually appears on the last root.

The roots in the compound are sometimes linked by a linking morpheme or **interfix**. There are two linking morphemes in English: *-o-*, as in *Anglo-Saxon, thermometer*) and *-s-* (originally a genitive suffix), as in *statesman, tradesman*. Other widespread Indo-European interfixes are *-e-*, *-er-*, *-en-*, *-a-*

Find examples of interfixes in languages other than English.

In agglutinative languages, composition can result in the creation of very long compound forms. English displays agglutinative tendencies in compounds formed by **inner syntax** where the structure of the form reminds of the structure of a phrase or sentence: *forget-me-not*, a *ne'er-do-well*, a *what's-his-name* fellow, *take-away* meals, *hide-and-seek* games, *bring-your-own-bottle* parties, a *sink-or-swim* method of teaching, a *hand-me-down* suit, a *don't-give-a-damn* attitude, etc.

> *To look at Montmorency you would imagine that he was an angel sent upon the earth, for some reason withheld from mankind, in the shape of a small fox-terrier. There is a sort of* **Oh-what-a-wicked- world-this-is-and-how-I- wish-I-could-do-something- to-make-it-better-and-nobler** *expression about Montmorency that has been known to bring the tears into the eyes of pious old ladies and gentlemen.*
>
> From: Gerome K. Gerome *Three Men in a Boat* Chapter 2

[1] O. Jespersen, P. Christophersen, N. Haislund, K. Schibsbye 1954, 2007., p. 137.

Contraction is shortening of a word by omitting some of its elements. Depending on the position of the omitted element, there are three types of contraction:

Aphaeresis is the dropping out of the initial sound(s) of the source lexeme. In English, it can be observed:

- in borrowed words: Cf. French *espace* – English *space*, French *arrière*, English *rear*, etc.
- in words which are often used, mainly in colloquial speech: *bus* from *omnibus*, *van* from *caravan*, *phone* from *telephone,* etc.
- in rapid speech: *taters* for *potatoes*, *brolly* for *umbrella,* etc.

Syncope involves the dropping of a sound or syllable from the middle part of the source lexeme. Many English lexemes are syncopated forms of lexemes borrowed from French – as e.g. *captain* from *capitaine*, *nurture* from *nourriture* or *chimney* from *cheminée*.

While their status in the structure is debatable, the process from which forms like *shan't, won't, doesn't*, etc. originate is clearly syncopation.

Syncopation is, again, the process to which the specific pronunciations of some proper nouns is due – Cf. *Leicester* ['lɛstə] or *Worcester* – ['wʊstə].

Apocope – the dropping of sounds or syllables from the end of the word – is the most frequently used type of contraction in English – Cf. *gym, prof, zoo*, etc. Apocope can act on both parts of compound lexemes to yield: *sci-fi, hi-fi, pro-am, etc*.

Acrostic words or **Acronyms** are the result of a special type of blending – of the initial letters or syllables *RAAF* (Royal Auxiliary Air Force), *FBI, MP*, etc.

Some authors[2] make a distinction between acronyms -- which are pronounced as single words – and **initialisms** (or **alphabetisms**) – pronounced as individual letters.

Globally known examples of acronyms are *laser* (Light Amplification by Stimulated Emission of Radiation) or OPEC (Organization of Petroleum Exporting Countries).

Examples of initialisms are FBI (Federal Bureau of Investigation), CIA (Central Intelligence Agency), DVD (Digital Video Disk).

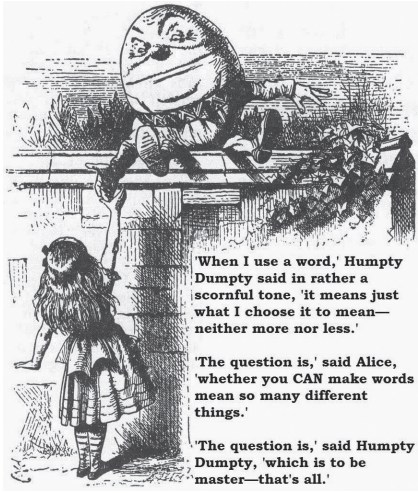
[2] Cf. Crystal, D., 2003: p. 120.

Blending as a process of word formation is a combination of composition and shortening (syncopation). Two roots are blended into one; in the process, sounds from one of the roots, or both, are dropped out.

Blending was a means of word formation as far back as Old English – as evidenced by the blends *lord* from *hlafweard* (loaf ward), *lady* from *hlaf dige* (loaf kneader) or *goodbye* (from *God be with you*).

In Lewis Carroll's *Through the Looking Glass* Alice asks Humpty Dumpty to clarify the meaning of a poem called *Jabberwocky*, beginning and ending with the following verse:

`Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the mome raths outgrabe.



Humpty Dumpty explains that words like *brillig* (four o'clock in the afternoon – the time when you begin broiling things for dinner) or *slithy* (lithe and slimy) are "like a **portmanteau** – there are two meanings packed up into one word".

The 20th century witnessed a rapid increase of blendings in the English language. Some high-frequency lexemes formed by blending are:

flurry (flaw + hurry)
blot (blemish + dot)
smog (smoke + fog)
slender (slight, slim + tender)
twirl (twist + whirl), etc.

Conversion is a means of word formation (of forming nouns out of verbs and verbs out of other parts of speech) whereby a root or stem carrying one generalised grammatical meaning (nominal, verbal, adjectival, etc.) acquires another generalised grammatical meaning, This change of meaning goes hand in hand with a semantic shift, the acquisition of new syntagmatic potential (new possibilities for derivation, composition, etc.) and with a new set of paradigmatic forms. As no formal marking is involved, the base forms of the source and derived word are homonymous. – Cf. *love* (noun) – *love* (verb). Other examples of conversion are:

$go_v \rightarrow go_n$, $say_v \rightarrow say_n$, $talk_v \rightarrow talk_n$

$drum_n \rightarrow drum_v$, $cook_n \rightarrow cook_v$,

$shower_n \rightarrow shower_v$ → $milk_n \rightarrow milk_v$

$water_n$ → $water_v$,

$lavish_{adj} \rightarrow lavish_v$, $black_{adj} \rightarrow black_v$

Because of the scarcity of inflectional markers in English, this language very easily forms conversives. Together with composition, conversion is the most productive means of enriching the lexicon of Modern English.

Conversion is context-dependent, so that each act of conversion realises only one of the lexical meanings of the lexeme. Once a conversive lexeme has become part of the lexicon, however, it can acquire additional meanings as a result of secondary acts of conversion – as is the case with the three meanings of *doctor*:

doctor (n) → doctor (v)

medical specialist, healer → to heal

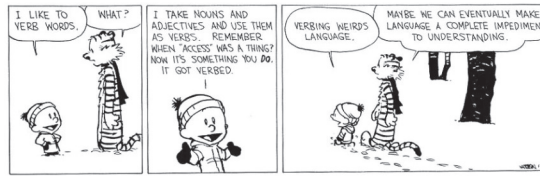
academic degree → award an academic degree

fake coin → to fake.

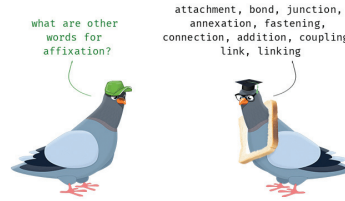
Once a pair of conversive lexemes is established in the lexicon, reconversion is possible:

$cable_N$ (steel or fiber rope) → $cable_V$ (supply or fasten with a cable or cables)

$cable_V$ (send a cablegram) → $cable_N$ (cablegram).



Affixation (or **derivation**) is the formation of new words by means of the addition of **affixes** (prefixes and suffixes) to a root or stem. While this process has been productive at all stages of the development of the Indo-European languages, its productivity varies from one language to another and from one historical stage in the development of a language to another.



Some English affixes are of native, others – of foreign origin. The affixes of foreign origin were not borrowed as word formatives, but as parts of a borrowed word. An affix can only be considered to be adopted in a language after the speakers of the language start using it as a word formative to help in the formation of new lexemes or forms. This becomes possible when many foreign lexemes with the same affix become well established in the language. To take an example, *-cy* (a suffix for the formation of abstract nouns – French *-cie*, Latin *–tia* – first appeared in words borrowed during the same period: *prophecy, primacy, policy, privacy, piracy, infancy, constancy, frequency, diplomacy* etc. After the 16th century, purely English forms started to appear with this suffix: *secrecy, permanency, consistency*, etc.

Back-formation has the appearance of a process which is the inverse of affixation: it is the formation of a seeming root-word from a word which appears to be (but in fact is not) its derivative.

Frequently given examples of back-formation are: *burgle* from *burglar*, *beg* from *beggar*, *edit* form *editor*, *baby-sit* from *baby-sitter.*

The verb *to wing* (1. to fly somewhere; 2. to be sent somewhere very quickly) was formed from the adjective *winged* (which is in turn suffixally derived from the noun *wing*).

In the 1970s, a verb *to therap* appeared, from the noun *therapist*.

Find many more examples of back-formation at:
https://en.wiktionary.org/wiki/Appendix:List_of_English_back-formations

Read the text and do the tasks:

It is very strange, this domination of our intellect by our digestive organs. We cannot work, we cannot think, unless our stomach wills so. It dictates to us our emotions, our passions. After eggs and bacon, it says, "Work!" After beefsteak and porter, it says, "Sleep!" After a cup of tea (two spoonfuls for each cup, and don't let it stand more than three minutes), it says to the brain, "Now, rise, and show your strength. Be eloquent, and deep, and tender; see, with a clear eye, into Nature and into life; spread your white wings of quivering thought, and soar, a god-like spirit, over the whirling world beneath you, up through long lanes of flaming stars to the gates of eternity!"

From Jerome K. Jerome "Three Men In a Boat"

1. List the word formation processes which you can identify in the excerpt.
2. For each process type, list the forms in the text which exemplify it.
3. Provide a morphological analysis of the forms.

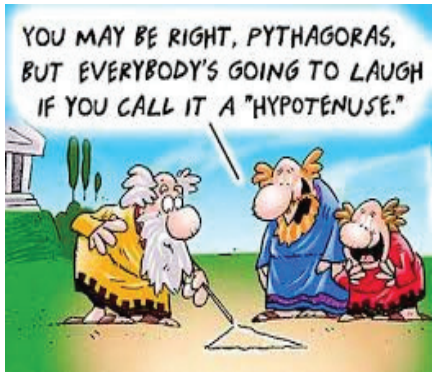
A **neologism** is a newly coined word or term which can be attributed to a particular person or publication and is not yet established in the lexicon. The new coinage is motivated by the necessity to fill a gap in the lexicon with a new signifier, following the emergence of a new signified. A neologism usually takes at least a decade to be established in the lexicon. "Neologism" is itself a relatively new term: it is a borrowing from the French language and was first attested in the second half of the 18th century.

Neologisms may be:

- entirely new creations which do not use the other word formation processes of the language, such as *troll, Jedi, droid, google, noob;*
- newly evolved meanings of existing lexemes – e.g. *sick* (good);
- newly formed compounds: *metrosexual, digital detox*;
- new clippings: app;
  - initialisms: *BFF* (best friends forever);
  - acronyms: *KISS* (keep it short and simple).
- brand names created for specific advertising and PR campaigns that are now used generically, to stand for the product: *hoover, laundromat, Kleenex, escalator, Xerox, zipper* and many others.

The term **protologism** is even more recent, a coinage of the early 2000s referring to a word which has not yet gained sufficient popularity to be considered a neologism, as e.g. *refizzify, googlized or decepticon.*

Look up the meaning of these protologisms and find many more interesting new coinages at: *https://en.wiktionary.org/wiki/Appendix:List_of_protologisms*



Occasionalisms (or **nonce words**) are, as the name indicates, created on a particular occasion; they do not enter the lexicon, or even recur. A famous occasionalism used by Lewis Carroll in his poem *Jabberwocky* is *jabberwock* – meaningless language or nonsensical behavior. Most speakers of English have read *Through the Looking Glass* and know what *jabberwock* means, but do not use it themselves.

B**orrowing** is the process of enriching the lexicon of one language with lexemes belonging to the lexicon of another language.

The borrowed lexemes are called borrowings or **loanwords**. Once they enter the lexicon of a new language, loanwords start adapting to the system of this language. They gradually adapt to the phonological system – changing vowels and consonants, simplifying consonant groups, shifting the accent, etc.; they appear with new paradigmatic forms; they usually also undergo semantic shifts. Thus they become **assimilated** in the new language.

Which are the source languages for the following English lexemes?

ballet, pneumonia, psychology, reason, honour, season

What phonological and grammatical changes have they undergone?

The English lexeme *beef* is a borrowing from the French language. What changes has it undergone? Do some research, if necessary.

When many loanwords with the same affix enter a new language, after a time the affix can be felt to belong to that language and may be used in processes of derivation for the creation of new lexemes, combining with native roots. Such productive affixes in English are the prefixes *re-* and *de-* or the suffixes *-ism* and *-cy*. Conversely, when borrowed roots enter a language, they can combine with native derivational affixes to form new lexemes. Lexemes in which some of the lexical morphemes are native while others are borrowed are called **loan-blends.**

A **loan-translation** or **calque** is the result of an item-for-item translation of the source lexeme. "Loanword" itself is a loan-translation of the German *lehnwort*; *marriage of convenience* is from French, and *long time no see* – from Chinese.

Other examples of calques in English are: *blue-blood*: noble birth — from Spanish *sangre azul*; *Devil's advocate*: one who advocates the opposing side — from Latin *advocatus diaboli; fleamarket*: a place selling secondhand goods, from French *marché aux puces*; *masterpiece*: "a work of outstanding artistry or skill", from Dutch *meesterstuk*; *wisdom tooth*: from Latin *dentes sapientiæ*, in turn from Greek *sophronisteres*, from *sophron* "prudent, self-controlled."

Quiz 7

1. **English makes use of suprasegmental modification to signal the difference between:**
   a. adjectives and nouns
   b. adjectives and verbs
   c. nouns and verbs.
2. ***shm*-reduplication is a type of:**
   a. vocalic reduplication
   b. rhyming reduplication
   c. total reduplication.
3. **An interfix appears between:**
   a. two roots
   b. two affixes
   c. a root and an affix.
4. ***Sonar* (from sound navigation and ranging) is an example of:**
   a. acronymy b. apocope c. syncope.
5. **Portmanteau words are formed by:**
   a. conversion b. affixation c. blending.
6. **The verb *to bottle* was formed by:**
   a. composition b. conversion c. syncopation.
7. **The verb *to back-form* is the result of a process of:**
   a. back-formation
   b. partial reduplication
   c. reconversion.
8. **A nonce word is:**
   a. a neologism
   b. a protologism
   c. an occasionalism.
9. ***Calque* is another term for:**
   a. a loanword
   b. a loan-translation
   c. a loan-blend.

Further reading

Bauer, L., 1983. English Word-formation. Cambridge Textbooks in Linguistics, Cambridge University Press.

Crystal, D., 2003. The Cambridge Encyclopedia of the English Language. Cambridge Univeristy Press.

Jespersen, O., P. Christophersen, N. Haislund, K. Schibsbye 1954, 2007. A Modern English Grammar on Historical Principles. Part VI, Morphology. Routledge, London and New York.

Mair, Ch., 2008. English Linguistics: An Introduction. Tübingen, DE: Gunter Narr.

Marchand, H,. 1969. The Categories and Types of Present-Day English Word-formation. München: Verlag C. H. Beck.

Matthews, P. H., 1974. Morphology: An Introduction to the Theory of Word-structure. London, UK: Cambridge University Press.

Plag, I., 2002. Word-formation in English. Siegen, DE: Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J., 1985. A Comprehensive Grammar of the English Language. New York, NY: Longman.

Ryder, M. E., 1994. Ordered Chaos: the Interpretation of English Noun-noun Compounds. Berkeley, CA: University of California Press.

Scalise, S., Fabregas, A., Forza, F., 2009. Exocentricity in Compounding. NL: John Benjamins.

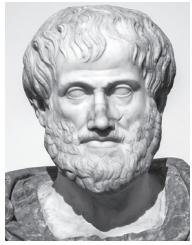
Yule, G., 2006. The Study of Language. Cambridge, UK: Cambridge University Press.

Lecture 8

Categories and Markers

Categories

Categorisation is the process of grouping mental images of objects, phenomena or ideas on the basis of comparison, analogy and perceived similarity/dissimilarity. Our experience of the outside world and of ourselves, the situations in which we find ourselves and the manner in which we reflect them form part of the system of our knowledge of reality, of the manner in which we structure the accessible world in our minds and in our communication with others. This process of systematisation is both reflected in the system of language and supported, enhanced, by it: it is *linguistic* thought in the form of propositions and judgments that is made use of in formulating assertions about reality and setting out groups or **categories**.



The notion of categorisation as the basis of thought and cognition was central to the philosophical works of Plato and Aristotle. In his treatise *Categories* Aristotle (384 BC – 322 BC) presents a division of beings ("the things there are") into categories, based on four predications: genus, peculiar property, definition, accident. Underlying the concept of **definition** is that of **essential predication** – compare the predications: (1) *Mary is a girl* and (2) *Mary is happy*. The second may well be true, but does not define Mary. It is non-essential. Definitions are not only essentially predicated, but must be predicated only of the essence in question. The predication in (1), while essential, can be true for Anne, Peggy, Betty and many others. It is thus not a definition. The predication: (3) *An animal having reason* is generally believed to be a sound definition of man. Essential predications set out, according to Aristotle, ten categories: 1/ substance ("this"): man, bird, house; 2/ quantity: *five meters, fourteen degrees;* 3/ quality: *beautiful, blue;* 4/ relation: *cheaper, longer;* 5/ location: *here, in the garden;* 6/ time: *last month, tomorrow;* 7/ position: *lies, stands;* 8/ having (possession): *is dressed, is equipped;* 9/ action: *writes, runs;* 10/ undergoing: *is hurt, is cold.*

Aristotle's **classical view of categories** – that they should be clearly defined and mutually exclusive – has recently been challenged by a more realistic, **cognitive approach,** first put forward in the 70es by Eleanor Rosch.[1] This approach allows for categories to have **fuzzy boundaries** and to be based on **prototypes** (typical, representative or best examples of a category). Prototypes are considered to be central to concept formation and essential for human development.

[1] Cf. E. Rosch, 1973 and later work by the same author, as well as work by G. Lakoff and R. Langacker in the 80es – Cf. Recommended Literature.

The categories of language

Categorisation is used in linguistics to provide a systematic description of the units of language. In R.L. Trask's words, linguistic categories are "so varied that no general definition is possible; in practice, a category is simply any class of related grammatical objects which someone wants to consider."[2] Still, when speaking of categories in language we usually refer to either:

1/ groups of lexemes, traditionally called "parts of speech" or "word classes" (e.g. nouns, verbs, conjunctions) and their subgroups (e.g. transitive/intransitive verbs) – usually referred to as "**lexical categories**" because they appear with their category and subcategory labels in the lexicon, or to:

2/ structurally marked semantic oppositions within a given conceptual domain (e.g. Number, Person, Aspect, etc.) – usually referred to as "**grammatical categories**" because they appear as oppositions of marked/unmarked forms in the paradigms of lexical categories.

Categories can vary from language to language or from one stage in the development of a language to another. Also, they are in no way "given" or obvious for anyone to observe. Their setting out is based on human thinking and experience and the taxonomy is thus open to change and development. Four criteria (separately or in conjunction) have been used in the setting out and description of word categories: (generalised) meaning, markers, position and function.

So far as we know, the first attempts at the classification of the lexical stock of a language date back to the 6th-5th C. BC and the work of the Sanskrit grammarian Yaska. He defined two main categories of words on the basis of morphological criteria: Inflectable (Nouns and Verbs) and Uninflectable (Pre-verbs and Particles).

In the *Cratylus* dialogue, Plato defined sentences as combinations of Verbs and Nouns (though what he probably had in mind were the syntactic categories of Subject and Predicate and their prototypical exponents). To these, Aristotle added the Article and a class of Conjunctions (also covering Pronouns and Prepositions). This classification was gradually extended to include the Participle as a separate class, Adverbs and Interjections. The Adjective was separated from the class of nouns as late as the mid-18th century.

[2] Cf. R. L. Trask 2007.

Markers

The most general morphological classification of lexemes groups them into two large classes: inflectables and uninflectables. The reason why some lexemes in a language can be inflected (i.e. contain inflectional affixes) is that:

1/ Their meaning is compatible with the expression of different values of a general semantic category. Making use of Aristotle's taxonomy, words expressing "substantivity" can also express different values of the category of "quantity"; words expressing "quality" can also express different values of the category of "relation", etc.).

2/ In the system of that language, members of this particular class of words can appear with morphological markers (inflections) for values of the category.

The taxonomies of Yaska and Aristotle are still applicable to the languages of the Indo-European group, but we now know that there are many other language groups, for which they might not be. Languages can differ in the set of their lexical categories (there are languages, as e.g. Tagalog, which do not differentiate between nouns and verbs[3]), in the set of their grammatical categories (not all languages have the grammatical category of Number – as for instance Khmer, where plurality is expressed with lexical means), in the types and number of the values of the categories, in the nature of the markers. To take an example, the category of Number, which is nearly universal, can have the form of an opposition of one (singular, unmarked)/ more than one (plural); one (singular, unmarked)/two (dual number, as in Slovene)/more than two (plural); one (singular)/two (dual number)/three (trial number, as in many Austronesian languages and the creole language Tok Pisin – but only in the subsystem of Pronouns!)/more than three (plural); collective (unmarked)/singulative (marked) – as in Welsh.

Grammatical categories are formed by binary or tertiary (seldom more than that) oppositions of values. Of the values of a category, one is usually general, the other/others are specific. While present tense forms can be used to speak of past or future events, past tense forms normally refer to past events only. The specific forms usually have special formal markers; they are said to be **marked**. Polymorphemic lexemes which contain an inflectional morpheme mark one of the values of a grammatical category: in (3) the morpheme *-ed* is a marker of past tense. Similarly, *-s* in *books* is a marker of plural number.

[3] *http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.466.1176&rep=rep1&type=pdf.*

(3) decide- + -ed
Root INFL (Past Tense)

The "general" forms either have a special marker or, frequently, coincide phonologically with the root/stem. The root/stem is a lexical morpheme (or lexical morphemes) and cannot express grammatical meaning: in that, it differs from the word form, which is both lexically and grammatically complete. It is in positions of contrast – as in the oppositions of values forming the categories of Tense (Past/Non-Past) or Number (Plural (more than one)/Singular (one)) – that "nothing" has the distinctive value of a zero-marker:

(4) decide- + - ∅
Root INFL (Present Tense)

The contrast between the phonetically null zero inflection and the *-ed* inflection marks the opposition of values of the English grammatical category of Tense:

(5) Present tense Past tense
- ∅/-s -ed.

The introduction of abstract (zero) markers into linguistic analysis, artificial as it may appear, harmonises with a theory of language in which the linguistic sign is abstract in nature and language is an abstract system. Further, it integrates well in a theory of language in which categories are formed by oppositions of values expressed by oppositions of markers. Finally, it is in accord with a theory in which elements entering into oppositions have the same status in the hierarchy – because any opposition is grounded in similarity.

FORMS	STEM	PRESENT THIRD-PERSON SINGULAR	PRESENT PARTICIPLE	PAST TENSE	PAST PARTICIPLE
Inflectional Suffixes		{-s 3d}	{-ING vb}	{-D pt}	{-D pp}
MODELS	show	*shows*	*showing*	*showed*	*showed*
	ring	rings	ringing	rang	rung
	cut	cuts	cutting	cut	cut

If grammatical categories must involve oppositions of values, and if these values receive contrasting markers, then **lexemes belonging to (prototypical members of) inflectable classes cannot be monomorphemic – minimally, they must be formed of a root morpheme and a zero-inflection.**

Most monomorphemic lexemes belong to the class of uninflectables.

Inherent, agreement and relational categories

Grammatical categories can be inherent, relational or of agreement.

Inherent categories are independent of syntactic relations. They can be **inherent proper** (not inflectionally marked) or else **inflectionally marked**. We define a category as inherent if its values follow naturally from the semantics of the class of words which express it.

Examples of inherent proper categories are Gender of nouns in French, German or Bulgarian and verbal Aspect in languages of the Slavonic group. These are categories which hold for the whole lexeme and do not have inflectional markers. Thus, the French lexeme *table* (table) is feminine (f), while *cartable* (backpack) is masculine (m); in Bulgarian, *кола* (car) is feminine, *баща* (father) is masculine, *канапе* (sofa) is neuter. To take another example from Bulgarian, the verb *топя* (dip) is Imperfective while the verb *купя* (buy) is Perfective.

The Aristotelian category of "time" is inherent to the class of words expressing "action" or "undergoing"; the category of "quantity" is inherent to the class of words expressing the category of "substance"; similarly, classes of words with the semantics of "quality" can inherently express "relation". These are all examples of inherent categories marked inflectionally in most Indo-European languages.

Agreement categories are dependent on relations of subordination. Although the category of relation is in no way "inherent" to the class of Adjectives expressing quality, attributive or predicative adjectives in many languages are marked for inherent categories of the noun – typically, Gender or Number: Fr. *une belle fleur* (a beautiful flower) / *la fleur est belle* (the flower is beautiful), *un beau jardin* (a beautiful garden) / *le jardin est beau* (the garden is beautiful). The English adjective does not agree for number with its head noun, but the English verb-predicate agrees with its subject for person/number: *You are happy / Mary is happy*.

Relational categories are also dependent on subordination, but the dependent word form does not "agree" with its head by marking one of its inherent categories. Instead, the syntactic head governs its dependent by imposing on it a form marking a specific syntactic function. The category of Case is a relational category which can have a varying number of values depending of the class of words marked and the language. The category of Case of the English noun has two values (Common and Possessive); Basque nouns have markers for as many as twelve separate values of the category.

Types of markers

There is more than one way in which values of a category can be marked.

Inflection (marking with a grammatical morpheme) is the most widespread type of marker. Inflectional suffixes have word final position, following the stem (e.g. *overgeneralisation**s***). This type of marking is easy to formulate in terms of rules of application and is usually considered to be the *regular* one.

Vocalic change (vowel change or vowel lengthening) was a characteristic feature of Proto-Indo-European. In the system of Germanic languages, it is preserved as a marker of verbal forms (the past tense form and the past participle – as in *sing, sang, sung*). In Greek, vowel gradation marks different values of Case in nouns.

Suppletion is marking by using an etymologically unrelated root. It is a feature of tens of world languages, including those of the Indo-European group. Suppletive forms are used for the forms of the verb of existence or link verb and some other high-frequency verbs (be/is/are/was/were; go/went); for the degrees of comparison of the adjectives *good* (good, better, best) and *bad* (worse, worst); similarly, for the plural of some high-frequency nouns – Cf. Bulgarian *един човек/двама души/много хора.*

> Using a dictionary or a reference book, find more suppletive forms in the languages you are familiar with. Which lexical classes of words are marked? Which grammatical categories are marked?

Grammatical categories can also be marked by special grammatical markers – **form-words**. These can be particles or other function words: primary auxiliaries, modal auxiliaries, analytic markers of degree, etc. (***will*** *come,* ***has*** *written,* ***ще*** *дойде*).

A **change in syntactic function**, usually together with inflection or the use of form words, is used in many languages to mark the Passive or Activo-Passive Voice.

Reduplication is a means of word-formation in Indo-European languages, but there are languages where it has grammatical functions – Cf. the formation of the plural of nouns in Walpiri – the language of an indigenous people living in Northern Australia: *kurdu* (child) → *kurdukurdu* (children) or *kamina* (girl) → *kaminakamina* (girls).

Grammaticalisation

According to Thomas Givon[4], the development of language is from iconicity to arbitrariness and symbolism, from the concrete to the abstract. Such is also the path of development of thought and knowledge. If this is indeed the case, language development is not just a string of successive reorganisations of the system, it is evolution. It is with the process of grammaticalisation that many authors associate the evolution of language.

Grammaticalisation is a process of change (possibly, evolution) from lexical to grammatical forms and from grammatical to even more grammatical forms. The motivation for grammaticalisation is the use of linguistic forms that are specific, easily accessible and clearly differentiated for the purpose of expressing less specific, accessible and clearly distinct meanings. Based on generalisations on modern languages, the theory of grammaticalisation offers an instrument for reconstructing the emergence and development of grammatical forms and constructs. The hypothesis is, first, that at the time of its initial development human language was structurally simpler when compared with most of the languages we can observe today, and second, that the changes in language and the motivation for these changes were the same at earlier and later stages – i.e. that the same processes hold at any stage of language development.

B. Heine and T. Kuteva[5] propose the following parameters of grammaticalisation:

A. Extensions or widening – the emergence of new grammatical meanings when language expressions are used in new contexts.

B. Desemantisation – the loss of semantic content as a result of the use of a unit in a new context. What is lost in the process is the meaning that is incompatible with this new context.

C. Decategorisation – the loss of those morphological and syntactic properties that characterised earlier uses but are not relevant in the new context. Decategorisation often involves the loss of derivational morphology and/or the loss of status of the unit a syntactically visible phrase.

D. Phonetic reduction – the loss of phonetic segments or supersegmental properties.

[4] Cf. T. Givon., 2005.

[5] Cf. B. Heine and T. Kuteva, 2007.

Grammaticalisation leads to a move from concrete to abstract meanings; from relatively independent referential meanings to less referential meanings; from open classes to closed classes; from grammatical forms with an internal morphological structure to unchangeable forms; from longer to shorter grammatical forms. These are the generally acknowledged stages in the development of many grammatical constructs from free word combinations; they could well be part of a universal mechanism of language development.

A clear example of the process of grammaticalisation is the gradual development of analytical verbal forms from descriptive lexico-syntactic structures. In English the perfect in *I have caught a fish* evolved from the free combination of a verb of possession with a noun phrase: *I have (possess) a caught fish*[6]. In French, the majority of verbs form the perfect with *avoir* (to have), but reflexive verbs and a handful of verbs, when used intransitively, require *être* (to be): *J'ai attrapé un poisson. Je me suis habillé. Je suis monté. J'ai monté l'escalier*. Structures with both the verb of being and the verb of having were quite frequent in Old Bulgarian texts. They both underwent processes of grammaticalisation, with the former dominating in the north-eastern dialects (which formed the basis of the literary language) while the latter is still alive in the western dialects – compare *Написал съм пет стихотворения* (I have written ten poems) and *Имам написани пет стихотворения.* (I have ten poems written.)

A number of publications[7] draw attention to the following changes of the English verb in grammaticalised structures marked for Aspect, Correlation or Voice:

1/ phonetic reduction: *They've received the parcel* (but not: **They've a parcel*);

2/ extension of the meaning, with gradual fading of the initial meaning of existence/ characterisation or possession – ending up in:

3/ complete desemantisation;

4/ decategorisation: in the first place, change in the possible set of collocations and colligations: as an auxiliary, the verb forms a unity with a well-defined non-finite verbal form (the latter also undergoing a process of decategorisation).

[6] Cf. M. Mincoff, 1958: 131.

[7] Cf. e.g. M. Stambolieva, 1989; M. Stambolieva, 2008 and others.

The parts of speech

The three criteria traditionally used to set out parts of speech – form, meaning and function, while not resulting in clear-cut categories of the Aristotelian type, nevertheless define classes of lexemes with prototypical members.

Those lexemes which do not inflect and/or do not have more than one form in their paradigms belong to the uninflectable classes of Interjections, Conjunctions, Prepositions and Particles. Many members of the class of Adverbs are uninflectable.

Interjections occupy a detached position in the sentence and signal emotions.

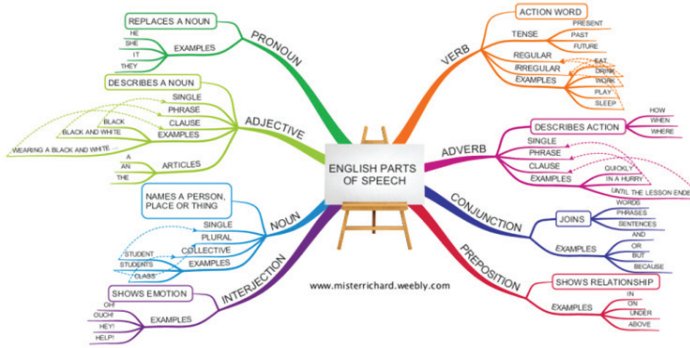
Conjunctions link sentences in a text, clauses within a sentence or phrases with the same syntactic status inside the clause.

Prepositions are syntactic connectives linking subordinate noun phrases or gerunds to other phrases.

Particles are markers of grammatical categories.

Other markers of grammatical categories can have their own paradigms. These are: articles, primary and modal auxiliaries.

Articles are grammatical markers of definiteness/indefiniteness which can have special agreement forms for number, gender or case. Many English grammarians, because they insist on the Latin tradition of eight parts of speech, do not set out articles as a separate class of words but prefer to analyse them as a subclass of adjectives.



Primary auxiliaries are elements of analytical verbal forms; they participate in marking the grammatical categories of Aspect, Correlation and Voice of the verb.

Modal auxiliaries mark the values of the grammatical verbal category Modality.

Inflectable classes include Nouns, Adjectives, Numerals, Pro-forms, Adverbs and Verbs. The categories they mark vary from language to langue. As a general guide:

Nouns are a class of words which express the generalised semantics of substantivity. They can inflect for the inherent categories of Number and for the relational category of Case. Their category of Definiteness can be marked inflectionally or by the article.

A**djectives** express the generalised semantics of quality. They can inflect for one inherent grammatical category – Degree, as well as for Gender, Number or Case and Definiteness – as agreement categories.

Numerals express, as their name suggests, number. Some linguists do not set them out as a separate category, classifying cardinals as nouns and ordinals as adjectives. Numerals however do not express the semantics of quality and do not mark the inherent category of adjectives – Degree.

Pronouns are a varied group of lexemes which do not have reference of their own but acquire it through an antecedent – a referential expression which they substitute (usually for reasons of economy). They can mark Person (an inherent category specific to them), Number and Case.

Adverbs modify predications in terms of manner, reason, time, place, etc. Those which express manner can mark, as an inherent category, Degree. Adverbs do not inflect for agreement or relational categories.

Verbs carry the semantics of state, process or event. They mark the agreement categories of Person and Number and impose on noun phrases, through government, the category of Case. The categories inherent to the verb are Aspect, Tense, Correlation, Modality and Voice. **Aspect** presents the situation as a non-event (Imperfective or Progressive) or an event (Perfective, Non-Progressive). **Tense** places the situation in relation to a reference point (Present, Before-Present (Past), After-Present (Future), etc.) on the time axis. **Correlation** expresses precedence with respect to another situation without introducing additional reference points. In Indo-European languages it is marked by the Perfect. **Voice** (Active, Passive or Activo-Passive) marks changes in the unmarked syntactic realisation of the arguments of the verb-predicate Finally, Modality is formed by oppositions of values relating to the aim of an utterance: the speaker's comment on its truth value (epistemic modality) or the expression of their desire to see the situation change (deontic modality).

Quiz 8

1. **A prototype is:**
   a. an essential predication
   b. a best example
   c. a substance.
2. **Number is:**
   a. a lexical category
   b. a grammatical category
   c. a marker
3. **Particles are:**
   a. inflectable
   b. uninflectable
   c. zero-inflectable.
4. **A grammatical category cannot have:**
   a. one member
   b. two members
   c. four members.
5. **The category of Case of Nouns is:**
   a. inherent
   b. relational
   c. of agreement.
6. **The past tense form of the verb *to go* is formed by:**
   a. vocalic change
   b. suppletion
   c. reduplication.
7. **In the process of grammaticalisation, the loss of meaning is called:**
   a. desemantisation
   b. decategorisation
   c. extension.

Further reading

Gärdenfors, P., 2004. Conceptual Spaces: The Geometry of Thought, MIT Press.

Givon, T., 2002. Bio-linguistics. The Santa Barbara Lectures. Amsterdam, Philadelphia: John Benjamins.

Givon, T., 2005. Context as Other Minds. Amsterdam: Benjamins.

Heine, Bernd and Tania Kuteva, 2007. The Genesis of Grammar. Oxford University Press.

Lakoff, G., 1987. Women, fire and dangerous things: What categories reveal about the mind, University of Chicago Press.

Medin, D. L., Altom, M. W., & Murphy, T. D., 1984. Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(3), 333-352.

Mincoff, M., 1958. An English Grammar. State Publishing House "Science and Art", Sofia.

Rosch, E., 1975. "Cognitive Representations of Semantic Categories", *Journal of Experimental Psychology*: *General*, Vol.104, No.3, pp. 192–233.

Rosch, E.H., 1973: "Natural categories", *Cognitive Psychology* 4, 328–350.

Rosch, E., 1978. "Principles of Categorization", in Rosch, E. & Lloyd, B.B. (eds), Cognition and Categorization, Lawrence Erlbaum Associates, Publishers, (Hillsdale), pp. 27–48.

Rosch, E., 1983. "Prototype Classification and Logical Classification: The Two Systems in Scholnick, E.K. (ed), New Trends in Conceptual Representation: Challenges to Piaget's Theory. Lawrence Erlbaum Associates, Hillsdale," pp. 73–86.

Stambolieva, M., 1989. "On the Semantics of Empty Verbs". *European Journal of Semiotic Studies* 2.

Stambolieva, M., 2008. Building Up Aspect. Peter Lang Academic Publishers.

Taylor, J. R., 2003. Linguistic Categorisation, Oxford University Press.

R.L. Trask, 2008. Language and Linguistics: The Key Concepts, 2nd edition, Edited by Peter Stockwell. New York: Routledge.

Lecture 9

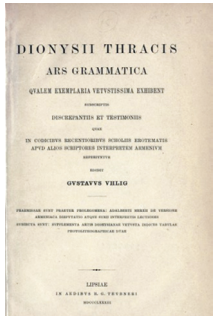
The Syntactic Level

What is syntax?

Syntax is a branch of grammar which studies the rules for building sentences, clauses and word combinations. The term comes from Ancient Greek *σύνταξις* "arrangement", which consists of *σύν*, "together," and *τάξις*, "ordering".

Major problems of syntax are:

- the identification of the units of the syntactic level of language;
- the grammatical nature and structure of the units of the syntactic level and the relations existing between them;
- the structural and functional classification of the units of the syntactic level.


DIONYSII THRACIS
ARS GRAMMATICA
GVSTAVVS VILLIG
LIPSIAE

Syntax only became a dominant branch in linguistic analysis in the course of the 20th century. Before that, works on grammar either predominantly focused on morphological descriptions and classifications or viewed the sentence as structurally identical with the judgment.

Here is what Dionysius Thrax (170-90 BC), considered to be the father of European "traditional grammar" for his work *The Art of Grammar*,[1] ("Téknē grammatiké") had to say about syntax:

> 13. On the Sentence.
>
> A Sentence is a combination of words, whether in prose or in verse, making complete sense. There are eight parts of speech: Noun, Verb, Participle, Article, Pronoun, Preposition, Adverb, and Conjunction. The proper noun, as a species, is subordinate to the noun.

Note that Dionysios, like many syntacticians in the centuries to come, made no distinction between lexical classes and syntactic classes – because he viewed the sentence as nothing more than a meaningful combination of words.

In their 1660 "Grammaire générale et raisonée", Claude Lancelot (1615-95) and Antoine Arnaud (1612-94), two monks from the Port-Royal Abbey in France, introduced the Cartesian view of language as a manifestation of thought, and of syntax as the linguistic form of the judgment.

[1] Cf. *The Grammar of Dionysios Thrax* (Translated from the Greek by Thomas Davidson), Reprinted from the Journal of Speculative Philosophy. St. Louis, MO. Printed by the Studley Co, 321 North Main Street, 1874, p. 8.

Word combinations and syntactic relations

"Got idea. Talk better. Combine words. Make sentences."

The **word combination** is a grammatically well-formed unity of two or more notional words expressing a unified concept and filling one syntactic position in the sentence. Word combinations can be phraseological or free. Phraseological combinations are studied by Lexicology. The object of syntactic analysis are free combinations.

Free word combinations can be structurally simple or complex. Simple word combinations consist of two lexemes, e.g. *nice day*, *walk fast*, etc. Complex word combinations are obtained by enlarging either the whole word combination or one of its parts: *exceptionally nice day*, *exceptionally nice April day*, *walk quite fast*, etc.

Depending on the nature of the relation between their constituent parts, word combinations can be **coordinative**, **subordinative**, or – in the case of some complex unities – both. Examples of word combinations with coordinatively linked parts are: *love and hatred*, *Jack and Jill*, *old but nice*, etc. In subordinatively linked word combinations one of the elements is grammatically dependent on the other.

There are four types of subordinative linking:

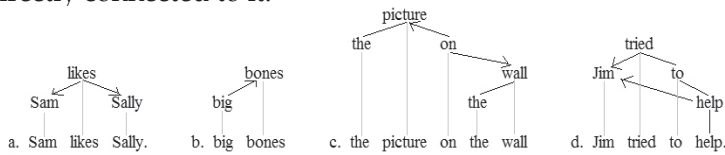
1/ **Agreement**: the subordinate element appears with the same morphological markers as the main element – Cf. agreement for number in: *these houses*.

2/ **Government** (or Case government): the head element requires the subordinate one to appear in a particular (Case) form: *I saw him*.

3/ **Prepositional linking**: *cut with a knife*, *speak to him*, *the branches of the tree*, etc.

4/ **Application**: there are no morphological requirements imposed on the subordinate elements, e.g. *read books*, *good idea*.

One of the dominant models of syntactic analysis, **Dependency Syntax**, is based on relations of dependency (represented below by directed arrows) between the word forms in the sentence. The verb is considered to be the head of the sentence; all other word forms are directly or indirectly connected to it:



The principles of dependency syntax were first presented by the French linguist Lucien Tesnière (1893-1954) in his work "Eléments de syntaxe structurale",[2] where he also introduced basic notions for modern syntactic theory, such as valency and the opposition argument/adjunct.

Predication in logic and in language

Sentences and clauses are not seen by all linguists as strings of word forms linked by coordinative or subordinative bonds.

Traditional linguistics defines the sentence as a grammatically well-structured complete unit having specific intonation and characterized with **predicativity** and **modality**. In other words, the sentence says (asserts, negates, asks, demonstrates) something about someone/something and expresses the attitude of the speaker towards the predicated content. Predicativity becomes a grammatical category when viewed as a syntactic relation, formally expressed as agreement, between the grammatical subject and the grammatical predicate – as the kernel, the structural centre of the sentence. Modality becomes a grammatical category when systematically expressed by means of special modal words or word forms.

The Simple Sentence is built around one predicative centre, i.e. it is made up of one clause. It is a grammatically formed complete unit having specific intonation. The predicative centre of the sentence is the bond between the subject and the predicate. The formal expression of this bond can be the morphological markers for person and number (i.e. agreement of the predicate with the subject) (1) and/or specific word order (2):

(1) John is cleaning his shoes.

(2) John loves Mary. / Mary loves John.

Most sentences consist of two main parts – a subject and a predicate, but this is not necessarily the case:

(3) Darkness. Coming?

Sentences where both the subject and predicate are expressed are called **two-member sentences**. In one-member sentences, either the subject or predicate is missing.

Both one-member and two-member sentences can be **expanded** with additional, **secondary parts of the sentence** (objects, attributes or adverbial modifiers).

[2] Published posthumously in 1959 – Cf. Tesnière 2015. Elements of structural syntax [English translation of Tesnière 1966]. John Benjamins, Amsterdam.

Their communicative aim defines sentences as **declarative** (4), **interrogative** (5), **imperative/optative** (6) and **exclamatory** (7):

(4) Peter is late. I work for a software company.

(5) Are you tired? Does Mary know the address? They are Bulgarian, aren't they?

(6) Give me the keys! / Long live the Republic!

(7) How nice! What a sight! Here you are!

Draw a vertical line to separate the subject and predicate groups.

1. The two women sat chatting.
2. Mary had a new carpet in her parlour.
3. Some of them have already been there.
4. What she wanted was a love too proud and independent to exhibit its joy or its pain.
5. That this charming young woman accepted political life with enthusiasm was a real surprise to me.

Sentences made up of more than one clause can be **compound** (8) – if the clauses are coordinatively linked, **complex** (9) – if the sentence contains a main clause and one or more subordinate clauses, or **complex-compound** (10) – if the clauses are linked by relations of both coordination and subordination:

(8) He walked along the quay and thought about the misery of life.

(9) Who does a thing like that cannot be trusted.

(10) She opened the book that lay on the desk and started reading.

Circle the clauses and define the sentences – Compound, Complex or Complex-Compound

1. I am at this present moment writing in a house situated on the banks of the Hebrus, which runs under my chamber window.
2. There I mused and considered what I would do with this money.
3. I am very tired, nevertheless I shall try to come.
4. Hurry up else we shall be late for dinner.
5. I have two things here that will fit you, and I am sure you want them both.
6. Passengers getting out of the open doors were directed towards the platform.
7. Trasker had been watching her with an inexplicable flush on his face.

Parts of the sentence

There are two main parts of the sentence – the Subject and the Predicate – and several secondary ones – the Object, the Attribute and the Adverbial Modifier.

Traditional syntax defines **the subject** as a major, syntactically independent part of the sentence. It is predicatively characterised by the predicate, which agrees with it in person and number.

The subject is typically a noun, a noun phrase or a pronoun; however, other parts of speech, as well as phrases or clauses, can also function as subject:

Numerals: *The first and the second were the best ones.*

Adjectives: *The poor were offered food and shelter.*

Non-finite verbal forms: *To have loved and lost is better than never to have loved at all. Having and enjoying had only made him long for more having and more enjoying.*

Substantivised parts of speech, e.g. Adverbs: *Once is not enough, twice is too often.*

Word groups: *Cyril to tea was one of their treats.*

Clauses: *That it was the year 1872 meant little to him.*

Although it us one of the main parts of the sentences, the subject of the sentence can sometimes be omitted. Languages like Italian or Bulgarian where the subject, if it is a personal pronoun, can easily be dropped, are called "**pro-drop languages**". English is not a pro-drop language, but it also allows **subject omission** in well-defined cases, as in the second clause of compound sentences (11), or in imperative, optative or exclamatory sentences (12).

(11) *He quietly opened the door and ran into the street.*

(12) *Give me that ball! Long live the Republic! How nice!*

Subject doubling is another widespread feature of natural language. It occurs when the personal pronoun in subject position is "doubled" for greater explicitness by a noun phrase in extraposition – at the beginning or end of the sentence:

(13) *Poor child, she didn't know how to thank us.*

(14) *He is a rare craft, that boy!*

Circle the subject phrase and underline the head of the phrase.

1. Her manner was a trifle too casual.
2. Participating in the preliminary matches gave the sportsmen confidence.
3. Losing the game was a great disappointment.
4. It is a great advantage to have a retentive memory.
5. The happy end of this film can hardly be called its drawback.

Apart from a few cases of omission of the subject, some languages (e.g. English and French) require an explicit subject phrase where pro-drop languages would leave an empty subject position. Because these subjects are felt to be simple gap-fillers, they are called '**dummy subjects'**.

Dummy subjects appear in **impersonal sentences** of non-pro-drop languages. These sentences are of two main types: of "being" and of "having". The former characterize a sensorially accessible part of reality (15, 16) while the latter are statements about reality (or part of it) containing certain entities (17, 18). In the latter case, French preserves the dummy *il*, but with a verb of "having". The Bulgarian equivalents appear with an empty subject position and an impersonal verb of "having":

(15) *It is cold* (*in the room*). *It is early*. FR: *Il est tôt, Il fait froid*.

(16) IT: *È freddo*. BG: *Студено е.*

(17) *There are unicorns* (*in the garden*). FR: *Il y a des licornes* (*dans le jardin*).

(18) BG : *В двора има еднорози*.

Dummy subjects can also appear in the sentence to fill a gap left by the movement of a "heavy" subject (a complex structure or a clause) to a position following the predicate. The dummy subject in such sentences is called "anticipatory" or "preparatory". The English preparatory subject is *it*. French allows both *il* and *ce*. In pro-drop languages, the subject position is empty (19).

(19) *It is a pity that you can't come*. FR*: C'est/Il est dommage que tu ne puisses pas venir*. IT : *È un peccato che tu non possa venire*. BG : *Жалко (е), че не можеш да дойдеш*.

In traditional syntax **the Predicate** is defined as a main part of the sentence presenting a characteristic feature of the Subject. Together with the subject, the Predicate forms the predicative word group, which is the kernel of a sentence.

The Predicate is said to be syntactically dependent on (subordinated to) the Subject because it agrees with the Subject in person and number. Agreement is one of the formal expressions of the relation of subordination.

Some predicates consist of one finite verbal form. They are **Simple Predicates**. Others have a more complex structure and contain (a) the finite form of a link verb and another part of speech, (b) a phase verb or a modal verb and a non-finite form of a full verb, (c) a phase / modal verb, followed by a link verb and a non-verbal phrase. These are **Complex Predicates**.

In simple predicates the verb is semantically complete (a Full Verb) and can function as Predicate on its own. Simple predicates contain a finite form of a Full Verb:

(20) *My mother writes (novels). The man grinned from ear to ear. Come here*!

A subtype of simple predicates are phrasal verbs:

(21) *My grandmother took her hat off and fluffed out her hair.*

Note that the analytical forms of full verbs are simple predicates, even if formally resembling word combinations. In:

(22) *John has been writing for at least an hour.*

has been writing is a finite form of the verb *to write*: the verbal root *writ-* carries the lexical meaning of the form, *be* + V-*ing* is a marker of (Progressive) Aspect, *have* + V-en is a marker for Correlation and *-s* marks the 3rd person singular of the Present Tense. Indo-European languages all have analytical word forms, even those which do not have morphological markers for Aspect or Correlation:

(23) FR: *J'ai fini mon travail*. POL : *Basia będzie czytać książkę*.

Structures with the auxiliary *do* are special cases of analytical forms. They mark negation, interrogation or emphasis:

(24) *I don't love cats. Do you love cats*? *Don't you love cats*? *I do love cats*.

Complex Predicates can be Verbal, Nominal or Mixed (Verbal-Nominal). **Complex Verbal Predicates** contain more than one verbal form. They express Modality, Phase or Iteration.

The **Complex Modal Predicate** is formed with a modal verb and the infinitive of the notional verb (25).

(25) *You should study more. They can't do that! You ought to have a dog, or at least a cat. He had to think carefully before answering that one.*

The phases of a situation – its beginning, middle or final part – can be marked lexically with a small set of verbs called **phase verbs**, e.g. *to begin*, *to start*, *to continue, to keep on, to stop, to cease*. These verbs combine with a full verb or link verb to form **Complex Phase Predicates**:

(26) *They started to walk. He continued to think of it. He had long since ceased to believe in their theories. I stopped smoking a year ago.*

In Slavonic languages phase can be expressed either with an affixally marked verb – Cf. BG. *пожълтея* (turn yellow), *оглушея* (become deaf), *вбеся се* (go mad), *запея* (start singing), *допия* (drink up), etc. or with a complex phase predicate.

Phase is a category closely related to Aspect. Another such category is Iteration: the expression of repetition. The structures formed with *would + infinitive* or *used to+ infinitive* are another type of complex verbal predicate:

(27) *On Sundays, we would go to Granny's place. I used to play the guitar, in those years.*

Complex Nominal Predicates are structural unities of a link verb and a non-verbal part, a **Predicative**. The Predicative can be expressed by an adjective, a noun, a pronoun, a numeral, an adverb, a prepositional phrase or a clause:

(28) *She is pretty. He sounds an exquisite escort. He became a lawyer. The fault wasn't ours. I am only thirty but I feel forty-eight. The war was far away. She was well. The campaign was at its beginning. That was what he did want.*

Language allows blended predicative structures – superimposed simple and complex nominal predications, as in (29). The predicatives appearing, as a result of the blend, after a full verb, are called **supplementary**.

(29) She came first. (from *She came*. and *She was first*.)

Secondary parts of the sentence

The **Object** narrows down or complements the meaning of the predicate.

The **Direct Object** is linked to the head of the predicate (typically, a transitive verb) by application (*I noticed the girl*. BG: *Забелязах момичето*.) or government – it is marked for the Accusative or Oblique case (*I saw him*. BG: *Видях го*.)

The **Indirect Object** is the second object of some transitive verbs of giving or communicating, linked to the verb by apposition or government. In most Indo-European languages, synthetic and analytical alike, the dative case has been preserved in the system of pronouns (In English, the accusative and dative forms coincide and are jointly referred to as 'oblique') (*He left her the house. They told us the news*.). In synthetic languages all noun phrases in this function appear in the dative (RUS: *Он рассказал маме новости. Он сказал моему лучшему другу новости*.).

The **Prepositional Object** is a phrase subordinate to the predicate, introduced by a preposition. The preposition is governed by the head word – in the sense that not any head word can take any preposition. Thus, one typically looks/glances *at* someone or something, sleeps *in* a bed, hopes *for* a miracle, etc.

Attributes and adverbial modifiers have a weaker relation to the predicative word group or the word they modify. In the majority of cases they are not essential to the construction of a grammatically well-formed sentence.

The Attribute can modify any noun (or noun substitute) in the sentence, irrespective of its function. The typical attribute is an adjective (*the scarlet letter*), or a noun in the genitive (*the baby's coat*) but it can also be a pronoun, a numeral, a noun, a non-finite verbal form, a prepositional phrase, a clause: *my* father, *three* books, the *school* house, the road *back*, the back *of the house*, etc.). Whenever structurally possible, the attribute agrees with its head noun. The **Apposition** is a type of attribute typically expressed by a noun or noun phrase and serving to characterise or specify the head word: *Doctor Smith, the river Danube, Mount Everest, uncle Sam,* etc. It is linked to the head noun by application.

The **Adverbial Modifier** completes the meaning of the verb or of the sentence by defining it as to place, extent (including time and space), reason, degree, manner, etc. It is most frequently expressed by an adverb, a prepositional phrase or a clause (*I stopped at five o'clock/at the corner/to rest/because I was tired/abruptly, etc.*)

Word order and syntactic typology

There have been a number of attempts to regroup the languages of the world on the basis not of genetic relations, but of formal characteristics. One of the pioneers of research in language typology, as we have seen (Cf. Lecture 6), is the British linguist Bernard Comrie, who proposed three main morphological language types. Three, again, are the major language types that have been set out for syntax – a classification based on word order: the relative position of the subject, verb-predicate and object. According to this criterion, to a very few exceptions languages fall into the following groups: SOV (e.g. Persian, Hindi, Japanese), VSO (e.g. Welsh, Arabic) and SVO (e.g. English, French, Bulgarian). Ordering patterns VOS, OVS and OSV exist, but are extremely rare.

The relative order of subject, predicate and object is neither diachronically nor synchronically fixed in a language. SOV is the word order pattern used by the majority of world languages. It was the word order of Proto-Indo-European, Ancient Greek and Classical Latin. In their development, many languages of the Indo-European group underwent a SOV-to-SVO shift. In the history of the English language, this shift began towards the end of the Old English period; Middle English (12th century) was a SVO language. In Romance languages, which emerged between the 6th and 9th centuries from Vulgar Latin, SVO was already a widespread word order pattern. SVO is the predominant word order in Old Bulgarian texts (9th – 11th century).

A number of typological distinctions go hand in hand with the OV/VO choice:

- In SOV languages, nouns are followed by postpositions; in SVO and VSO languages, nouns are preceded by prepositions.
- In SOV languages, the Genitive of the possessor precedes the head noun; it follows it in both SVO and VSO languages.
- Where the verb phrase contains an auxiliary, this precedes the main verb in SOV languages and follows it in SVO and VSO languages.
- Attributes expressed by adjectives or relative clauses precede the head noun in SOV languages but follow it in languages of the SVO and VSO groups.

Overall, languages of the OV group can be described as predominantly "head-last", while languages of the VO groups tend to be "head-first" – with a percentage of co-occurrence which excludes simple coincidence.

Quiz 9

1. **When did Syntax become a central branch of linguistics?**
   a. in Ancient Greece b. in the 17th C. c. in the 20th C.
2. **Which relation is not a subtype of subordinative linking?**
   a. coordination b. agreement c. government
3. **In Dependency syntax, the head of the sentence is:**
   a. the subject
   b. the predicate
   c. the finite verbal form
4. **Which of the sentence types below is NOT defined by its communicative aim?**
   a. declarative b. expanded c. optative
5. **How many clauses can a complex sentence contain?**
   a. one b. more than one c. more than two
6. **The clauses in a compound sentence stand in relations of:**
   a. coordination b. subordination c. both of these
7. **Pro-drop languages do not need:**
   a. subject pronouns b. subjects c. dummy subjects
8. **Which of the structures below is NOT a complex predicate?**
   a. two finite verbal forms linked by a conjunction
   b. a link verb and a non-verbal phrase
   c. a modal verb and a non-finite verbal form.
9. **The Apposition is a type of:**
   a. Object
   b. Attribute
   c. Adverbial Modifier
10. **Which of these word order types is the most widespread?**
    a. SVO
    b. SOV
    c. VOS

Further reading

Bechert, Johannes, Giuliano Bernini, Claude Buridant (eds.), 1990. Towards a Typology of European Languages. Empirical Approaches to Language Typology. Berlin: Mouton De Gruyter.

Biber, Douglas, Susan Conrad, and Geoffrey Leech, 2002. Grammar of Spoken and Written English. London: Longman.

Carnie, Andrew, Dan Siddiqi, and Yosuke Sato, 2014. The Routledge Handbook of Syntax. London: Routledge.

Carnie, Andrew, 2013. Syntax. A Generative Introduction. 3rd edition. New Jersey: Wiley-Blackwell.

Kuiper, Koenraad and Jacqui Nokes, 2013. Theories of Syntax: Concepts and case studies. London: Palgrave Macmillan.

Tallerman, Maggie, 2014. Understanding Syntax. 4th edition. London: Routledge.

Van Valin, Robert, 2001. An introduction to syntax. Cambridge: Cambridge University Press.

Culicover, Peter W. and Ray Jackendoff. 2005. Simpler Syntax. Oxford: Oxford University Press.

Pavey, Emma L., 2010. The Structure of Language: An Introduction to Grammatical Analysis. Cambridge: Cambridge University Press.

Sag, Ivan, Thomas Wasow, Emily Bender, 2003. Syntactic Theory: A Formal Introduction. 2nd edition. Stanford: CSLI.

Baltin, Mark and C. Collins. 2000. Handbook of Contemporary Syntactic Theory. New Jersey: Blackwell.

Huddleston, Rodney, Geoffrey Pullum, 2002. Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.

Shopen, Timothy (ed.), 2010. Language Typology and Syntactic Description. I Clause Structure. Cambridge: Cambridge University Press.

Lecture 10

Phrase Structure

Intuitions about structure

Even if, on the surface, a sentence can be seen as an ordered string of words, speakers of a language – be they linguists or not – intuitively feel that the words in this string are grouped into larger unities. Thus, in (1):

(1) My daughter gave the red scarf to her new friend.

'my' and 'daughter' form one group, 'the', 'red' and 'scarf' form another group; 'new' and 'friend' form a third group – together, they combine with 'her' to form a larger group before combining with 'to' to form the group 'to her new friend'. The groups 'the red scarf' and 'to that sweet child' combine with 'gave' to form an even larger constituent of the sentence – the Predicate group of traditional syntax.

A grammatical formalism which captures the intuition that sentences are structured as hierarchies of groups, called constituents or phrases, is a constituent structure, or phrase structure, grammar. Phrase structure can be represented as a set of rewrite rules (2), as a phrase marker (3) or as a bracketed structure (4):

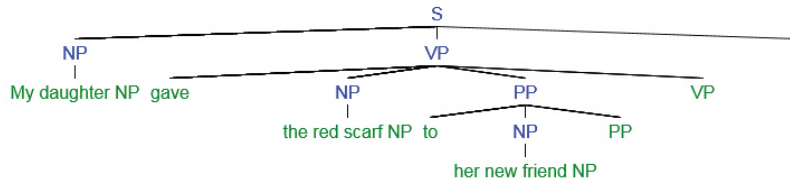
(2) S → NP VP

NP → D (A) N

PP → P NP

VP → V NP PP

(3)



(4) [S [NP My daughter NP] [VP gave [NP the red scarf NP] [PP to [NP her new friend NP] PP] VP] S]

In Phrase Structure syntax, unlike Dependency syntax, words are not the *immediate constituents* of a sentence. They are its *ultimate constituents*, organised hierarchically into consecutive levels of usually (but not necessarily) increasingly larger units called 'phrases'.

Distributional evidence

Evidence in support of the syntactic reality of phrasal constituents comes from different levels of structure. A. Radford[1] proposes a number of diagnostic tests for the constituent status of a word or word group: 1/ whether it permits internal positioning of adverbials; 2/ whether it can undergo movement (preposing or postposing); 3/ whether it can serve as a sentence-fragment; 4/ whether it can undergo ordinary coordination; 5/ whether it can be a shared constituent; 6/ whether it can be replaced by a pro-form; 7/ whether it can undergo ellipsis.

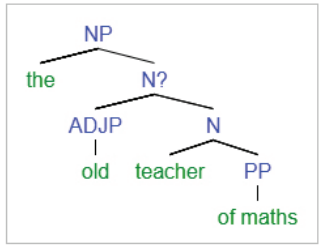
- Adverbial positions : adverbials are inserted between phrases:
  I have [often] been there. He was [badly] injured in the match. Peter has [just] arrived. I [foolishly] forgot to lock the door.
- Preposing: Only a whole phrase can be preposed, Cf.
  *[Your little brother] I can't stand. *Your little, I can't stand brother.*
  *The chairman [put off] the meting. – *Off the meeting, he put.*
- Postposing: (only phrases can appear in post-position).
  We offered [all the worldly goods we had] to the kidnappers.
  We offered to them [all the worldly goods we had].
- Sentence fragments (only phrases can be sentence fragments):
  Where did they go? [Up the road].
  *What argument did he bring up? *Up her late arrivals.*
- Coordination: Only constituents (usually of identical categories) can be coordinated:
  Annie bought [two pens] and [a pencil]. I sent [a letter] and [a postcard].
  We might [go to Plovdiv] and [visit the Fair]. I sent [a letter] and [to Annie].*
- Shared constituent coordination is possible if the shared string can be a constituent of each of the conjuncts:
  *I must, and you had better, [do some homework].*I rang and my brother picked [up Mother]*
- Pronominalisation: Proforms replace phrase-level constituents (of a well-specified level – Cf. below):
  John might go home, and so might Bill. I'll take the red ball and the blue one.
- Ellipsis: Only verb phrases can undergo ellipsis.
  I won't [do the job] – I bet you will [--] if they pay you well.

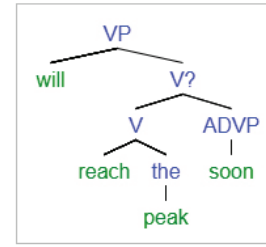
[1] A. Radford. Transformational Grammar. Cambridge Textbooks in Linguistics. Cambridge Cambridge University Press, pp. 69-105.

X-bar theory

It has long been observed that, in spite of considerable variation in the number and type of their ultimate constituents, at a more abstract level phrases display structural similarity. A number of arguments point to the existence of a level of syntactic structure which is intermediate between lexical categories (N, V, A, P – or, in general, X) and their maximal phrasal projections (NP, VP, AP, PP – or, in general, XP). In other words, there are arguments supporting an analysis where phrases like *the old teacher of maths* or *will reach the peak soon* have the following structure:

(5)





The "intermediate" level of structure is set out because, on the one hand, it cannot be equated with the maximal phrase: *I saw [a red butterfly]/ *I saw [red butterfly], We [will reach the peak soon] / *We [reach the peak soon]*; on the other hand, it cannot be equated to the lexical head, either: **We reach.*

The "intermediate" strings are phrases because they answer the diagnostic test for constituency:

- they can be coordinatively linked with other phrases of the same type: *My [dear colleague] and [best friend], will [buy a stamp] and [post the letter];*
- they can be a shared constituent in cases of shared constituent coordination: *He was our first (and beyond doubt our best) [teacher of English];*
- they can undergo pronominalisation with specific pro-forms, e.g. *one* for intermediate level noun phrases and *so* for intermediate level verb phrases: *This [colleague of yours] is more motivated than that [one].*
- they can be preposed: *[Go to school] I won't.*
- intermediate verb phrases are preceded by adverbials: *My friend has* never *[written a letter].*

Complements and adjuncts

In his 1977 book 'X-bar syntax'[2] the American linguist Ray Jackendoff formulated a "Unified Three-Level Hypothesis" of sentence structure, according to which there are exactly three levels of structure for all lexical categories (hence the *X*).

The first level (*Xo*) is that of the lexical head (*N, V, A*, …) of the syntactic phrase, e.g. *teacher, buy, anxious*, etc.

The second level (*X'* or *X-bar*) consists of a head and, possibly, one or more **complements** – compulsory or highly predictable dependents of the head, e.g. *professor of linguistics, read an article, anxious to go*, etc. This level can recursively be expanded with **adjuncts** (optional dependents): *new professor of linguistics with grey hair, buy a book to please mum, really anxious (to go)*, etc.

To form the highest phrasal level (X'', X-double bar or XP), a "completing" constituent must be added: a determiner (article / demonstrative / possessive pronoun) for noun phrases, an auxiliary for verb phrases, a marker of degree for adjectives and adverbs, etc.: *the/my/that (blue) car, has bought a book, most anxious to go*, etc.

Thus the "flat" rewrite rules of early constituent grammar (6) acquire a new, layered form (7):

(6) NP → Det (Adj) N (PP)

(7) NP → Specifier N'
N' → (Adjunct*) N' (Adjunct*)
N' → N_0 Complement(s)

where the star (*) indicates the possibility to recursively add adjuncts to one-bar phrases. Rule (7) also suggests that post-nominal phrases can be either complements (e.g. *a professor of maths*) or adjuncts (e.g. *a professor with grey hair*).

In terms of this X-bar convention specifiers, complements and adjuncts in XPs can be defined as follows:

(8) a. Specifiers expand X-bar (X') into X-double bar (X'')
b. Adjuncts expand X-bar into X-bar.
c. Complements expand X_0 into X-bar

This rule correctly predicts that in cases where both the complement and the adjuncts follow the head, the complement phrases will precede the adjuncts: *a professor of maths with grey hair*.

[2] Cf. Ray Jackendoff, 1977.

Extending X-bar

Syntactic phrases are endocentric, i.e. they are built round a lexical head which defines the overall properties of the phrase. A number of authors have indicated possible extensions of this structure to the description of clauses. In the string:

(9) I must find out $[_{S'}$whether $[_{S}$ this theory makes sense $_{S}]_{S'}]$

the bracketed part is a constituent made up of a clause preceded by a sentential conjunction, called 'complementiser'. In early constituent syntax the string *this theory makes sense* would be labeled *S*; this *S* together with the complementiser would be labeled *S'*:

(10) S' → COMP S

The phrase structure rule underlying this analysis however (first formulated in 1970 by J. Bresnan[3]) does not comply with the X-bar schema. It may of course be that S and S' are not endocentric categories and that X-bar rules only apply to projections of lexical categories. On the other hand, arguments have been put forward in favour of aligning them with the X-bar schema by introducing into the structure an abstract node: Inflection, INFL, or just I. This node would contain features that are basic for all predications: Agreement, as the expression of the predicative bond between the subject and the predicate phrases, Modality and Tense.

Even where clauses have no overt Auxiliary constituent, VP fronting leaves an I trace, demonstrating that sentences and clauses do contain a third node – one that is neither NP nor VP :

(11) $[_{S}\ [_{NP}$ He $_{NP}]\ [_{VP}$ answered/answers the question $_{VP}]_{S}] \rightarrow$
$[_{S}\ [_{VP}$ Answer the question $_{VP}]$, $[_{S}$ he $_{S}]\ [_{I}$ did/does $_{I}]_{S}]$

The fact that 'answer the question' can be moved indicates that it is a constituent, a separate phrase. The fact that the information on tense/ agreement is left behind indicates that it can be separated from the VP constituent, i.e. at a more abstract level of representation, the Tense/ Agreement features are not part of the verb phrase.

[3] Cf. J. W. Bresnan 1970.

Clause types

If the I_0 node is positively marked for Tense, the clause is finite; if it is negatively marked for Tense, the clause is non-finite. Because the finiteness/non-finiteness of a clause is one of its defining features, this is yet another argument in favour of a head status for I_o. VP will then appear in the complement position of I_0, the two forming the intermediate level I′. This phrase will in turn combine with the subject (now reanalysed as the Specifier sister of I′) to form a complete I′′ (IP) projection:

(12) [IP [[Subject (NP)] [I′ [Io] [Predicate (VP)]]]

The I_0 node is thus a structural position hosting Tense and Agreement features. To this position, auxiliaries and some link verbs move upon insertion into the syntactic structure. Proof of this movement is provided by verbal negation in many languages – Cf. the examples from English and French:

(13) a. You $[_{I'}[_{I}$ must $_{I}]$ not $[_{VP}$ read that book$_{VP}]_{I'}]$.

b. Pauline $[_{I'}$ ne $[_{I}$ doit $_{I}]$ pas $[_{VP}$ lire ce livre $_{VP}]_{I'}]$.

In English, verbs taking a clausal complement can subcategorise for either [+ Tense, +Agreement] clauses (14) or [–Tense, –Agreement] clauses (15):

(14) I know [that she is here].

(15) I want [him to come].

The nature of a [+ Tense, +Agreement] subordinate clause is largely determined by a. the presence/absence of a complementiser and b. the nature of the complementiser. Compare:

(16) a. Father wants [John to water the plants]

b. Father asked [whether John would water the plants]

c. Father knows [that John will water the plants]

d. Father suggested [that John should water the plants].

Because, on the one hand, the complementiser determines the type of clause (interrogative, declarative etc.) and, on the other hand, it is a lexical item, it is analysed as the head of S' (now CP):

(17) [CP [Specifier] [C' [Co] [IP]]]

Complementisers do not constitute an open class: the four complementisers that introduce subordinate clauses in English are: *that, if, whether* and *for*. The choice of IP type is determined by the choice of complementiser; *that* and *if* select a finite clause as their complement, *for* selects an infinitival clause, *whether* selects either type:

(18) a. I know [that [this is analysis is correct]].
b. I wonder [if [this is analysis is correct]].
c. I expect [for [the analysis to prove correct]].
d. I wonder [whether [my colleagues abandoned this analysis]].
e. I wonder [whether [to abandon this analysis]].

Two types of complement clauses do not have CP status: Exceptional Clauses and Small Clauses.

Exceptional Clauses typically occur as complements of verbs of saying or of thinking. They are non-finite and have no complementiser. Their subject behaves like the object of the predicate in the main clause: *I believe [John/him to be mistaken*].

Small Clauses are strings of the type *I consider [him very intelligent],* where the I_0 head of this clause contains only Gender/Number agreement features, but no Tense. Its complements are non-verbal projections.

Bulgarian complement causes can be marked as both [+Tense, + Agreement], as in English (19), or [– Tense, + Agreement] – because the *da*-construction dominant translation equivalent of the *to*-infinitive contains agreement features (20). Even Small Clauses, which are verbless, contain agreement features: for Gender and Number (21):

(19) Знам, [че ще дойдеш]. Знам, [че ще дойде].

(20) Искам [0 да дойдеш]. Искам [0 да дойде].

(21) Мислят [го за умен]. Мислят [я за умна]. Мислят [ги за умни].

Syntactic ambiguity is the possibility to interpret a sentence in more than one way and, for that reason, provide for it with more than one possible syntactic analysis.

A simple example of syntactic ambiguity is the phrase

old men and women,

which can be analysed as either:

$[_{NP}\ [_{N'}$ old $[_{N0}$men$_{N0}]\ _{N'}]_{NP}]$ and $[_{NP}\ [_{N'}\ [_{N0}$women$_{N0}]_{N'}]\ _{NP}]$

or:

$[_{NP}\ [_{N'}$ old $[_{N0}$men$_{N0}]$ and $[_{N0}$women$_{N0}]_{N'}]\ _{NP}]$.

Think of alternative interpretations for the following sentences and provide syntactic analyses for them:

Flying kites can be dangerous.

The policeman shot at the man with a gun.

They made a decision on the boat.

Movement

The introduction of CP as an X-bar equivalent of S' not only allows to formulate formal criteria for defining clause types; it offers, further, syntactic positions for moved phrases, as e.g. moved auxiliaries and "WH-words" in interrogative structures.

A declarative sentence or main clause has a lexically non-filled Complementiser (C_0) position:

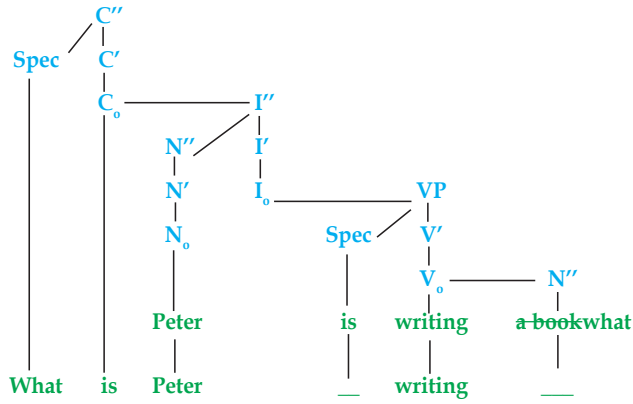
(22) a. CP [C' [C_0 [IP I will adopt this analysis.]]]
b. CP [C' [C_0 [IP Peter can check [whether the analysis is reliable.]]]]

In most English interrogative sentences, a link verb or auxiliary precedes the subject phrase, which in turn can be preceded, for Special Questions, by a "WH-word". The positions in the structure that are available for these phrases are C_0 and C', respectively:

(23) a. CP [C' What [C_0 is [IP Peter writing]]?]
b. CP [C' [C_0 [IP Peter is writing a book.]]]

Most syntactic theories assume that the structure in (23a) is derived from (23b) via movement. The movement of the WH-substitute of the complement *a book* to the Specifier position of CP is a type of NP-movement called 'WH-movement'. The movement of the head of IP (*is*) to the head position of CP (C_0) is called 'I-to-C movement'.

(24)



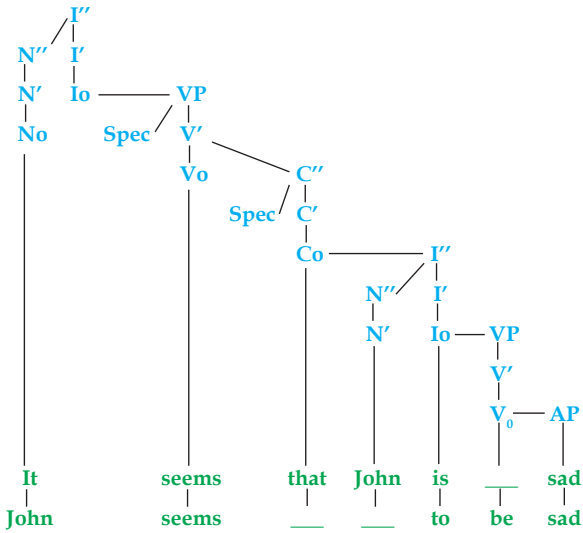
Raising

A Raising structure is one in which the subject or object of a subordinate clause appears in one of the argument positions of the main clause. Most syntactic theories view such structures as special cases of NP movement. The raised constituent (e.g. *John* in (25 b)) is not a semantic argument of the upper clause and typically occupies a position which is either unfilled by a NP, or else filled by a 'dummy' (e.g. *It* in (25)). The appearance of the object of a subordinate clause in the subject position of the main clause is called 'subject-to-subject raising':

(25a) It seems that John is sad.

(25b) John seems to be sad.

(26)



The verbs *to expect* or *to believe* allow so-called subject-to-object raising:

(27) Bill expected that Annie would be late → Bill expected Annie to be late.

Structures demonstrating object-to-subject raising are also possible in English, though infrequent – Cf. the so-called 'tough'-construction:

(28) It is tough for Carol to convince Bill. → Bill is tough (for Carol) to convince.

Quiz 10

1. **A Phrase Structure Grammar is:**
   a. a type of dependency grammar
   b. a set of phrases
   c. a grammatical formalism.
2. **In a Phrase Structure Grammar, the word forms are:**
   a. immediate constituents of the sentence
   b. ultimate constituents of the sentence
   c. both a. and b.
3. **Pro-forms replace:**
   a. word forms b. lemmas c. phrases
4. **X-bar theory is based on:**
   a. the structural variation of phrases
   b. the structural similarity of phrases
   c. the structural identity of phrases
5. **A compulsory dependent of a lexical head is a:**
   a. complement b. adjunct c. direct object
6. **Endocentric phrases are built:**
   a. by combining basic information from all the constituents
   b. round a lexical head
   c. round a head defining the basic properties of the phrase
7. **A Small Clause:**
   a. has only one predicative centre
   b. contains a non-finite verbal form
   c. does not contain a verbal form
8. **Which of the following is NOT an example of raising?**
   a. subject-to-subject b. subject-to-object c. object-to-object

Further reading

Bresnan, J.W., 1970. "On complementizers: towards a syntactic theory of complement types". *Foundations of Language* 6, 297-321.

Chomsky, Noam, 1981. Lectures on Government and Binding. Dordrecht: Foris.

Chomsky, Noam., 1995 The Minimalist Program. The MIT Press, Cambridge, Mass.

Cipollone, Nick, Steven Hartman Keiser, Shravan Vasishth, 1998. Language Files. Materials for an Introduction to Language and Linguistics. Seventh Edition. Ohio State University Press. Columbus.

Haegeman, L., 1991. Introduction to Government and Binding Theory. Blackwell.

Jackendoff, Ray, 1977. X-bar Syntax. A Study of Phrase Structure. Cambridge, Mass., and London, England: MIT Press.

Radford, A., 1992. Transformational Grammar. Cambridge Cambridge University Press.

Stambolieva, Maria, 2007. An Outline of Bulgarian Phrase Structure. In За човека и езика. Sofia: Publishing House of Sofia University "Sv. Kliment Ohridski", 90-100.

Bibliography

Avgustinova, Tania, K. Oliva. The Structure of Bulgarian Verb Complex. Computational Linguistics at the University of Saarland. Report No 4. Feb.1991.

Avgustinova, Tania. Word order and clitics in Bulgarian. Saarbrücken Dissertations in Computational Linguistics and Language Technology. Vol. 5 (1998). Saarbrücken, 1997.

Bally, Charles, Albert Sechehaye (eds.) Ferdinand de Saussure. Course in General Linguistics. McGraw Hill New York, Toronto, London, 1915.

Barwise, J. Situations and small worlds. In: The Situation in Logic, no. 17, CSLI Lecture Notes, 1987, 79–92.

Barwise, J. Perry. The Situation Underground, Stanford Working Papers in Semantics, Vol. 1, eds. J. Barwise and I. Sag, Stanford Cognitive Science Group 1980, Section D, 1980, 1–55.

Bates, Elizabeth, Judith C. Goodman. On the Inseparability of Grammar and the Lexicon: *Evidence from Acquisition, aphasia and real-time processing*. "Language and Cognitive Processes" 12(5/6), 1977. 507–84.

Benveniste, E. Problèmes de linguistique générale. Gallimard, 1981.

Bickerton, D. Language Evolution: A Brief Guide for Linguists. *Lingua* 2005.

Blokh, M.Y. A Course in Theoretical English Grammar. Москва, Высшая школа, 1983.

Bloomfield, Leonard. Language. New York. Henry Holt, 1933.

Borer, Hagit (editor). 1988.*Proceedings of the Seventh West Coast Conference on Formal Linguistics*. Stanford, CA: Stanford Linguistics Association, 1988.

Bresnan, J.W. "On complementizers: towards a syntactic theory of complement types". Foundations of Language 6, 1970, 297–321.

Campbell, L., R. Janda. Introduction: Conceptions of grammaticalization and their problems. Language sciences 23 (2–3), 2001, 93–112.

Carstairs-McCarthy, Andrew. Current Morphology. Linguistic Theory Guides, Routledge, 1992.

Cer, Daniel, *Marie-Catherine de Marneffe*, *Daniel Jurafsky*, and *Christopher D. Manning*. Parsing to Stanford Dependencies: Trade-offs between speed and accuracy. In *7th International Conference on Language Resources and Evaluation* (*LREC 2010*).

Chomsky, Noam. Current Issues in Linguistic Theory. The Hague, 1964.

Chomsky, Noam. Lectures on Government and Binding. Foris, Dordrecht,1981.

Chomsky, Noam. Remarks on Nominalisation. Jacobs, R.A. Rosenbaum, P.S., English Transformational Grammar, 1970, 184–221.

Chomsky, Noam. Syntactic Structures. Mouton, The Hague, 1957.

Chomsky, Noam. The Minimalist Program. The MIT Press, Cambridge, Mass., 1995.

Cipollone, Nick, Steven Hartman Keiser, Shravan Vasishth. Language Files. Materials for an Introduction to Language and Linguistics. Seventh Edition. Ohio State University Press. Columbus, 1998.

Comrie, Bernard. Before Complexity. Hawking and Gell-Mann (eds.), 1992, 193–211.

D'Errico, F., Henshilwood, C., Vanhaeren, M., van Niekerk, K. *Nassarius kraussianus* shell beads from Blombos Cave: evidence for symbolic behavior in the Middle Stone Age. *Journal of Human Evolution* 48 (1), 2005, 3–24.

Davis, Philip W. Modern Theories of Language. Prentice-Hall, Inc., Englewood Cliffs. New Jersey, 1973.

De Swart, Henriette. Verbal Aspect. In R.I.Binnick. The Oxford Handbook of Tense and Aspect. Oxford University Press, 2012, 752–780.

Desclés, J.-P. State, Event, Process and Typology. *General Linguistics*, vol. 29, No. 3, 1990, 159–200.

Devlin, K. Logic and Information. 1991, 49–51.

Erman, B., & Warren, B. The idiom principle and the open choice principle. Text, 20(1), 2000, 29–62.

Fillmore, Charles J. "The Case for Case". In Bach and Harms (ed.): *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1968, 1–88.

Firth, J. R. Modes of meaning. Papers in Linguistics 1934–1951, Oxford, 1957, pp. 190–215.

Fodor, Jerry A. The Language of Thought. Crowell Press, 1975.

Fries, Charles C. The Structure of English, New York: Harcourt, Brace, 1952.

Gentner, Dedre, Susan Goldin-Meadow (ред.) Language in Mind. Advances in the Study of Language and Thought. The MIT press, Cambridge, Massachusetts, London, 2003.

Givon, T. Bio-linguistics. The Santa Barbara Lectures. Amsterdam, Philadelphia, John Benjamins, 2002.

Givon, T. Context as Other Minds. Amsterdam, Benjamins, 2005.

Gruber, J. *Lexical structures in syntax and semantics*. Amsterdam; North Holland, 1965/76.

Haegeman, Liliane. Introduction to Government and Binding Theory. Oxford, Blackwell, 1991.

Halle, Morris (1973), *Prolegomena to a theory of word formation*. Linguistic Inquiry 4, 1973, 3–16.

Hauge, K.R. The Word Order of Predicate Clitics in Bulgarian. Meddelelser Nr 10, Slavonic and Baltic Institute of Oslo University, Oslo, 1976.

Hauser, M.D., Chomsky, N., Fitch, W.T. The faculty of language: what is it, who has it, and how did it evolve? *Science*, 2002, 298, 2002, 1569–79.

Heine, Bernd and Tania Kuteva. The Genesis of Grammar. Oxford University Press, 2007.

Henshilwood, C. Reading the artefacts: gleaning language skills from the Middle Stone Age in Southern Africa. Cradle of Language Conference, Spier Estate, Stellenbosch. Nov. 2006.

Hoey, M. Persuasive rhetoric in linguistics: a stylistic study of some features of the language of Noam Chomsky. In Hunston, S., & Thompson, G. (eds.), Evaluation in text: authorial stance and the construction of discourse (pp. 28–37). Oxford: Oxford University Press, 2000, 28–37.

Hudson, Richard. An Introduction to Word Grammar. Cambridge University Press, 2010.

Hudson, Richard. Language Networks. The New Word Grammar. Oxford University Press, 2007.

Hudson, Richard. Word Grammar. Oxford, OX, England, B. Blackwell, 1984.

Hunston, S. Semantic prosody revisited. *International Journal of Corpus Linguistics*, 12(2), 2007, 249–268.

Jackendoff, R.S. Semantic Interpretation in Generative Grammar. MIT Press, Cambridge, Mass, 1972.

Jackendoff, Ray. Foundations of language: brain, meaning, grammar, evolution. Oxford University Press, 2002.

Jackendoff, Ray. How language helps us think. *Pragmatics and cognition 4*, 1996, 1–34.

Jackendoff, Ray. X-bar Syntax. A Study of Phrase Structure. MIT Press, Cambridge, Mass., and London, England, 1977.

Jespersen, O. The Philosophy of Grammar. The University of Chicago Press, 1965.

Karolak, S. Configurations aspectuelles inchoative et résultative en bulgare: Studia kognitywne No 4. 2001, 31–47.

Katz, Jerrold J., Jerry A. Fodor. The Structure of a Semantic Theory. In: Jerrold J. Katz, Jerry A. Fodor (eds.) The Structure of Language. Readings in the Philosophy of Language. Englewood Cliffs, New Jersey. Prentice-Hall, 1964, 479–518.

King, A. Configuring Topic and Focus in Russian.Doctoral Dissertation, Stanford University, 1992.

Klein, R.G., Edgar, B. The dawn of human culture. New York, John Wiley & Sons, 2002.

Kuczaj, Stan A., II, and Jennifer I. Hendry. Does language help animals think? Dedre Gentner, Susan Goldin-Meadow (ред.) Language in Mind. Advances in the Study of Language and Thought. The MIT press, Cambridge, Massachusetts, London, 2003, 237–276.

Lepeschy, Giulio. A Survey of Structural Linguistics. London, Faber and Faber, 1970.

Lightfoot, D. The Language Lottery. MIT Press, Cambridge, Mass., 1982.

Lucy, J.A. Lucy. Grammatical categories and cognition: a case study of the linguistic relativity hypothesis. New York: Cambridge University Press, 1992.

Lucy, John A. Language diversity and thought: A reformulation of the linguistic relativity hypothesis. Cambridge, UK: Cambridge University Press, 1992.

Luria, A.R. Luria. Cognitive development: Its cultural and social foundations. Harvard University Press, 1982. ISBN 9780674137325

Marchand, Hans. The Categories and Types of Present-Day English Word-Formation. München: Verlag C. H. Beck, 1969.

Mincoff, M. An English Grammar. State Publishing House “Science and Art”, Sofia, 1958.

Molhova, Jana. On the Nature of the Morpheme in Contemporary English. English Studies. The Kliment of Ohrida University of Sofia. 1981.

Mourelatos, A. "Events, Processes and States". Syntax and Semantics. Tense and Aspect 14, 1981, 191–212.

Partee, Barbara H. Partee, Michael Bennett. Toward the Logic of Tense and Aspect in English. Blackwell, Published online 14 Jan 2008. DOI: 10.1002/9780470751305.ch4.

Radford, Andrew. Transformational Grammar. Cambridge University Press, 1988.

Ritz, Marie-Eve Ritz. Perfect Tense and Aspect. In R.I.Binnick. The Oxford Handbook of Tense and Aspect. Oxford University Press, 2012, 881–907.

Rivkin, Julie, Michel Ryan (eds.). Literary Theory: An Anthology. Blackwell Publishing Ltd. (Second Edition), 2004.

Rudin, Catherine. Aspects of Bulgarian Syntax: Complementizers and WH-constructions. Slavica Publishers, Columbus, Ohio, 1986.

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake and Dan Flickinger: *Multiword Expressions: A Pain in the Neck for NLP* (2002) in: Lecture Notes in Computer Science, Vol. 2276, 2002, 1–15.

Saussure, Ferdinand de. Course in General Linguistics. New York, London, Toronto. McGraw-Hill, 1966.

Saussure, Ferdinand de: *https://www.scribd.com/document/25077281/SAUSSURE-Ferdinand-Cours-de-linguistique-generale*, Uploaded by EDisPAL on Jan 12, 2010.

Seuren, Pieter A.M. Western Linguistics. An Historical Introduction. Oxford, Blackwell, 1998.

Sinclair, J. McH. Collocation Reviewed. Italy: Tuscan Word Centre, 2007.

Sinclair, J. McH. Corpus, concordance, collocation. Oxford: Oxford University Press, 1991.

Sinclair, J. McH. The lexical item. In E. Weigand (Ed.), Contrastive Lexical Semantics (Current Issues in Linguistic Theory 171) (pp. 1–24). Amsterdam: John Benjamins, 1998.

Sinclair, J. McH. The search for units of meaning. Textus, 9(1), 75 V106, 1996.

Sinclair, J. McH. Trust the text. London: Routledge, 2004.

Slobin, Dan I. Language and thought online. Dedre Gentner, Susan Goldin-Meadow (ред.) Language in Mind. Advances in the Study of Language and Thought. The MIT press, Cambridge, Massachusetts, London, 2003, 157–92.

Stambolieva, Maria. "On the Semantics of Empty Verbs". European Journal of Semiotic Studies 2, 1989.

Stambolieva, Maria. An Outline of Bulgarian Phrase Structure. In: „За човека и езика". Университетско издателство „Св. Климент Охридски", 2007, 90–100.

Stambolieva, Maria. *Building Up Aspect*. Peter Lang Academic Publishers, 2008.

Stambolieva, Maria. The Scope of Subcategorisation in Bulgarian and English. B: Maria Grozeva-Minkova/Boris Naimushin (ред.) Globalisierung, Interkulturelle Kommunikation und Sprache. Akten des 44. Linguistischen Kolloquiums. Frankfurt am Main, Peter Lang GmbH, 2014.

Talmy, L. Path to realization: A typology of event conflation. *Proceedings of the Berkeley Linguistics Society*, 17, 1991, 480–519.

Thornton, Tim. Wittgenstein on Language and Thought: The Philosophy of Content. Edinburgh University Press, 1998.

Tomasello, Michael, Malinda Carpenter. Shared Intentionality. In: Developmental Science, vol.10, Issue 1, January 2007, pp. 121–125.

Tomasello, Michael. The Origins of Human Communication. The MIT Press, 2008.

Vendler, Z. "Verbs and Times". Linguistics and Philosophy. Cornell Uniersity Press, Ithaca, 1957.

Vygotsky, L.S. Thought and language. Cambridge, MA: MIT Press, 1962.

Wackernagel, Jacob. *Lectures on Syntax: with special reference to Greek, Latin, and Germanic*, (ed. and translated by David Langslow), New York: Oxford University Press, 2009

Williams, E.O. Williams. "On the notions 'lexically related' and 'head of a word'". *Linguistic Inquiry* 12, 1981, 245–74.

Wundt, W.M. [Logik. Eine Untersuchen der Principien der Erkenntniss und der Methoden Wissenschaftlicher Forschung (Logic. An investigation into the principles of knowledge and the methods of scientific research), Enke, Stuttgart 1880 & 1883.

Yokoyama, O. Discourse and Word Order. Amsterdam: John Benjamins, 1986.

Березин, Ф. М., Б. Н. Головин. Общее языкознание. Москва, „Просвещение", 1979.

Бояджиев, Тодор. Словообразуване. В: Съвременен български език. София, „Петър Берон", 1998, 227–276.

Георгиев, Вл. Възникване на нови сложни глаголни форми със спомагателен глагол „имам". Помагало по българска морфология. Глагол. София, 1976.

Засорина, Л. Н. Введение в структурную лингвистику. Москва, Высшая школа, 1974.

Исаченко, А. В. Грамматический строй русского языка. Морфология. Братислава, 1960.

Каращук, П. М. Словообразование английского языка. Москва, Высш. школа, 1977.

Москов, Моско. Езикознание. София, „Народна просвета". 1982.

Осенова, Петя, Кирил Симов. Формална граматика на българския език. Институт за паралелна обработка на информацията, БАН, София 2007.

Осенова, Петя. Граматическо моделиране на българския език (с оглед на автоматичната обработка на естествен език). Парадигма, София, 2016.

Паскалева, Елена П. Българската морфология в анотационните стандарти. В Стамболиева (ред.) „Компютърна лингвистика. Проблеми и перспективи", АНАБЕЛА, кн. 1. 2012, 39–65.

Пенчев, Йордан. Български синтаксис (Управление и свързване). Пловдивско университетско издателство, 1993.

Смирницкий, А. И. Морфология английского явыка. Москва. Изд. литературы на иностранных языках, 1959.

Стамболиева, Мария. Строеж на естествения език. Издателство „Парадигма“, София, 2007.

Стамболиева, Мария. „Класове думи“. Годишник на Департамент Англицистика, НБУ, 2015.

Стамболиева, Мария, Петър Дошков. Три етапа в развитието на трансформационната граматика. ЮЗУ, Благоевград, 1991.

Answers to the Quizzes

Quiz 1:

1c, 2d, 3c, 4c, 5a, 6d, 7d.

Quiz 2:

1c, 2b, 3b, 4c, 5b, 6a, 7c

Quiz 3:

1c, 2c, 3c, 4a, 5a, 6c, 7b

Quiz 4:

1b, 2b, 3a, 4a, 5b, 6c, 7a, 8a, 9b

Quiz 5:

1c, 2b, 3c, 4b, 5c, 6b, 7b

Quiz 6:

1a, 2a, 3b, 4b, 5c, 6a, 7c, 8a, 9c, 10a

Quiz 7:

1c, 2b, 3a, 4a, 5c, 6b, 7a, 8c, 9b

Quiz 8:

1b, 2b, 3b, 4a, 5b, 6b, 7a

Quiz 9:

1c, 2a, 3c, 4b, 5b, 6a, 7c, 8a, 9b, 10b

Quiz 10:

1c, 2b, 3c, 4b, 5a, 6c, 7c, 8c

The textbook ***10 Lectures in Linguistics*** offers a concise and accessible introduction to the field, designed to arouse reader interest and encourage further study.

Major contributors and contributions are presented, with the structuralist and formal approaches highlighted. Topics covered are both general – such as communication, human language development, major theories of language, levels of linguistic analysis – and more specific – as e.g. the structural types of morphemes, the linguistic status of the word, the organisation of the Lexicon, the analysis of phrase structure and syntactic movement.

The clear structure and language of the lessons and the interactive approach adopted by the author make the textbook suitable for a wide range of audiences taking their first steps in the science of language.