

# Introduction to Language

Jirka Hana

October 5, 2020



# Contents

1	Basic Concepts of Linguistics	5
2	Animal Communication	13
3	Phonetics	17
4	Phonology	25
5	Morphology	31
6	Syntax	39
7	Semantics	61
8	Pragmatics	67
9	Linguistic Theories	69
10	An example formal theory of language	71
11	Historical Linguistics: Language Change	73
12	Psycholinguistics: Language and Mind	81



# Chapter 1

## Basic Concepts of Linguistics

### 1.1 Language

**Language** is a system that associates sounds (or gestures) with meanings in a way that uses words and sentences.

**Linguistics** is the scientific study of human language. It tries:

- first, to observe languages and to describe them accurately,
- then, to find generalizations within what has been described,
- finally, to draw conclusions about the general nature of human language.

**Applied linguistics** attempts to make practical use of the knowledge derived from general linguistic research – in order, for example, to:

- improve the ways in which a student's native language is taught
- help people learn foreign language more efficiently
- write better dictionaries
- improve therapy for people with language problems
- search the Internet more efficiently and successfully

Linguistics overlaps and (ideally) cooperates with: psychology, sociology, anthropology, philosophy, logic, mathematics, computer science, speech pathology, acoustics, music, cryptanalysis, etc.

### 1.2 Speech vs. Writing

Speech (or signs in signed language) is primary, writing is secondary.

One of the basic assumptions of linguistics is that the most immediate manifestation of language is speech and not writing. Spoken language encodes thoughts into a physically transmittable form, while writing in turn, encodes spoken language into a physically preservable form.

### 1.2.1 Why it is sometimes claimed that writing is primary

- Written texts tend to be more carefully worded and better organized than spoken texts, they contain fewer errors, hesitations, and incomplete sentences, because writing is usually planned in advance, is subject to fewer time constraints, is proofread, etc.

However: How about instant messaging, quick e-mails?

- Spelling is more uniform across different individuals, places and times using the same language than is pronunciation.

However: *Swau lásku slawjk rúži pěl – Růžinu gewil wonný wzdech. – Gezero hladké w křowjch stinných* [K.H.Mácha: Mág 1836]

However: UK: *tyre, draught, colour, dialogue, penalise, centre, defence, ...*

USA: *tire, draft, color, dialog, penalize, center, defense, ...*

Moreover: Is uniformity the same as primacy?

- Written texts last and can be preserved for a long time.

However: CDs, youtube ...can preserve speech

- Writing styles change much more slowly than speech styles, and so writing seems more “permanent” and “authoritative”.

However: This can be is also disadvantage – writing lags behind the times.

### 1.2.2 Linguists’ reasons for claiming that speech is primary

- Historically, spoken language existed much earlier than writing.  
Writing was most likely invented in Sumer (Mesopotamia, current Iraq) about 5500 years ago. Language probably exists for 40,000 or more.
- There are many societies which only speak their language and do not write it. And no society uses only a written language (with no spoken form).
- We learn to speak before we learn to write.
- Most people say more during one month than they write during their entire lives.
- Writing must be taught, whereas spoken language is acquired automatically.
- Psycholinguistic evidence suggest that the processing and production of written language is overlaid on the spoken language centers in the brain (plus certain other centers).
- Speech contains information that writing lacks – intonation, stress, voice quality ...

## 1.3 Competence & Performance

”What’s one and one?”

”I don’t know”, said Alice, ”I lost count.”

”She can’t do Addition, ” the Red Queen interrupted.

Lewis Carroll, Through the Looking-Glass

**Competence:** the speaker's unconscious knowledge of his language enabling him to produce and understand an infinite number of sentences

**Performance:** the actual usage of competence in communication.

Biological limitation (memory, breathing), hesitations, errors, etc.

- (1) a. The mouse escaped.  
 b. The mouse the cat was chasing escaped.  
 c. The mouse the cat the dog barked at was chasing escaped.

Performance point of view: terrible

Competence point of view: ok

Grammatically fine, but stylistically bad.

- (2) a. The book is good.  
 b. The book on the shelf is good.  
 c. The book on the shelf in the corner is good.  
 d. The book on the shelf in the corner of my house is good.  
 e. The book on the shelf in the corner in the bedroom of my house is good.  
 f. The book on the shelf in the corner in the bedroom of my house in LA is good.

## 1.4 Descriptive vs. Prescriptive Approach to Language

### 1.4.1 Descriptive Approach

- Linguists attempt to *describe* the grammar of the language that exists in the minds of its speakers, i.e. to create a model of speakers' mental grammar.
- The resulting descriptive grammar describes person's basic linguistic knowledge. It explains how it is possible to speak and understand and it summarize what speakers know about the sounds, words, phrases and sentences of their language.
- Creating a descriptive grammar involves observing the language and trying to *discover* the principles or rules that govern it.
- Descriptive rules accept as given the patterns speakers actually use and try to account for them. Descriptive rules allow for different dialects of a language and even variation within one dialect.

### 1.4.2 Prescriptive Approach

- Prescriptivists tell you someone's idea of what is "good" or "bad".
- Prescriptive rules make a value judgment about the correctness of certain utterances and generally try to enforce a single standard. For example:

English:

- Don't split infinitives; don't say: *to easily understand*
- Don't end a sentence with a preposition; don't say *Where are you from?*

Czech:

- Proto je třeba jednoznačně odmítnout a do slušné společnosti nepouštět především hrůzy typu chromozóm či lýzozóm, neboť ty mohou užívat pouze málo gramotné osoby neznalé základu "sōma-sōmatos". Stejně odpudivá a nepříjemná je ale např. i dizertace.

Jaroslav Hořejší: Pláč jazyka českého. 1 Feb 2010

<http://blog.aktualne.cz/blogy/jaroslav-horejsi.php?itemid=8790>

- Don't use *mistička*, use the correct *mištička* [1930's]
- The people who prescriptive grammar *make up* the rules of the grammar.
- They attempt to impose the rules for speaking and writing on people without much regard for what the majority of educated speakers of a language actually say and write.
- So-called prescriptive grammar usually focuses only on a few issues and leaves the rest of a language undescribed (unprescribed?). In fact, from the linguistic point of view, this is not grammar at all.

### 1.4.3 Prescriptivism vs. Descriptivism

In summary: Linguists **describe** language, they do not **prescribe** it.

As a science, linguistics:

- is not in the business of making value judgments about language use.
- studies how language really is used and then attempts to **describe** the facts, in order to analyze and, eventually, explain them.

An Analogy:

- Physicists:
  - don't complain that objects fall to earth
  - simply observe and describe the fact of falling, then try to discover the laws that are behind it.
- Linguists:
  - don't say that people shouldn't use *ain't* or *bysme* 'colloq. would<sub>1pl</sub>'
  - simply observe that some people in certain situations do use *ain't* (without judging, although they do note any systematic correlations of such use with particular groups, regions, situations, styles, etc.)

## 1.5 The parts of Grammar

**Grammar** is a language system, a set of principles (rules) that underlie a language.

**Mental Grammar** – the knowledge of language that allows a person to produce and understand utterances

Grammar can be described as having different parts:

- phonetics

- phonology
- morphology
- syntax
- semantics
- pragmatics

Since linguists study all of these, the terms are also used to refer to subfields of linguistics.

### 1.5.1 Phonetics & Phonology

**Phonetics** – the production and perception of speech sounds as physical entities.

E.g., [v] is pronounced by bringing the lower lip into contact with upper teeth and forcing air out of the mouth while the vocal folds vibrate and nasal cavity is closed off.

**Phonology** – the sound patterns (the sound system of a particular language) and of sounds as abstract entities.

In Czech, a word never ends with a voiced obstruent (e.g., *zuba* [zubu] ‘tooth<sub>gen</sub>’ but *zub* [zup] ‘tooth<sub>nom</sub>’).

In English, a word never starts with [kn] (note that *knife* starts with [n] not [k]), while in German it is possible (e.g., *Knabe* ‘boy’)

In Setswana (a language of southern Africa), a consonant is always followed by a vowel – when the speakers adopted the word *Christmas* from English, they pronounce as *kirisimasi*.

### 1.5.2 Morphology

**Morphology** – the word structure and of systematic relations between words.

**Morpheme** – the building-blocks of words, the smallest linguistic unit which has a meaning or grammatical function.

Words are composed of morphemes (one or more).

*Sing-er-s answer-ed un-kind-ly*

*uč-i-tel-k-u* ‘she-teacher<sub>acc</sub>’

In comparison with many other languages, English has rather simple morphology.

### 1.5.3 Syntax

**Syntax** – phrase and sentence structure

Syntacticians try to discover rules that govern:

word order:	<i>The book is on the table.</i>	<i>*Table book on is the the.</i>
agreement:	<i>I am here.</i>	<i>*I are here.</i>
subject/object forms (cases):	<i>I like her.</i>	<i>*I like she.</i>
etc.		

Note: In linguistics, placing an asterisk (\*) before a sentence marks that sentence as ungrammatical, i.e., not of the kind normally used by most speakers of that language.

### 1.5.4 Semantics

**Semantics** is the literal meaning of sentences, phrases, words and morphemes.

E.g., What is the meaning of the word *vegetable*?

E.g., How does the word order influence meaning of sentence in English? How about Czech?

### 1.5.5 Pragmatics

**Pragmatics** studies language usage, especially how context influences the interpretation of utterances – the same sentence can be used to do different things in different situations.

E.g., *Gee, it's hot in here!* can be used either to state a fact or to get someone open a window.

Simply put: semantics is the literal meaning and pragmatics is the intended meaning.

## 1.6 Arbitrariness

The relation between form and meaning in language can be either:

- **arbitrary (conventional)**, in which case:
    - the meaning is not deducible from the form
    - the form is not deducible from the meaning
    - the connection between the form and meaning must be learned via memorization
  - **nonarbitrary**
    - the meaning is (at least partly) derivable from the form, and vice versa  
E.g., *buzz*<sup>En</sup>, *bzučení*<sup>Cz</sup> – ‘sound of the type made by (the wings of) bees’
- iconicity** – the most extreme example of nonarbitrary form/meaning connection: the form shows a physical correspondence to the meaning and vice versa

Non-language examples:

- arbitrary: traffic lights, warning siren
- nonarbitrary: a “no-smoking” sign (with a crossed-out cigarette), a deer-crossing sign (with a silhouette of a deer)

Language is overwhelmingly arbitrary.

If language were not arbitrary, then:

- different languages would not use different words for the same thing (in fact, there would be just one language), as they obviously do:  
English *tree*, Czech *strom*, French *arbre*, German *Baum*, Japanese *ki*, Korean *namu*.
- word forms would not change over time.  
Old English (before 1100) *hūs* → Modern English *house*

- word meanings would not change over time  
Middle English (before 1500) *girle* ‘child’ → Modern English *girl* ‘girl’  
Middle English *nice* ‘ignorant’ → Modern English *nice* ‘pleasant’  
Old Czech *letadlo* ‘bird’ → Modern Czech *letadlo* ‘airplane’

### 1.6.1 Limited Exceptions: Onomatopoeia and Sound Symbolism

There are two very limited and partial exceptions to the arbitrariness of language:

- **Onomatopoeia** = words whose sound imitates either the sound they denote or a sound associated with something they denote. These words are not entirely arbitrary. However, different languages represent the same natural sounds in slightly different ways (e.g., . English *cock-a-doodle-doo* ≠ Czech *kykyrykŧ*), which shows that they are not completely nonarbitrary, either.
- **Sound symbolism** refers to the very vague, elusive way in which certain sounds “feel” more appropriate for describing certain objects or meanings than do other sounds.
  - the vowels [i] or [ɪ] seem to suggest smallness  
*teensy-weensy, wee, little, Tommy* (vs. *Tom*), *squeak*; **but:** *big*
  - to English speakers, *gl-* suggest brightness:  
*glint glitter, gleam, glow*; **but:** *glove, glue, glum, glop*

### 1.6.2 Why is arbitrariness is an advantage?

- It allows user of a communication system to adopt the most convenient means available for communicating, since it obviates any need for the forms of signs to bear an inherent relationship to their meanings.
- It also makes it much easier for users of a communication system to refer to abstract entities, since it is hard to find a combination that involves an inherent link between a form and an abstract meaning.



## Chapter 2

# Animal Communication

### 2.1 Communication Systems

Human language is just one kind of communication system; not every communication system is a language. Communication can be defined as the intentional transmission of information.

Communication: From Latin *communicare* ‘to share’, from *communis* ‘shared by all’, from *com* ‘together’ + *munia* ‘public duties’

Cf. common, community, communism, communion

All communication systems have the following elements or properties:

- **Signals** (or signs) – perceptible entities by which messages can be sent.
  - **Visual:** gestures, signs, writing
  - **Auditory:** speech, warning cries
  - **Tactile:** braille

This categorization of modalities is made from the point of view of perception, we can also categorize them in terms of production (speech, clapping, and warning sirens are all auditory, but they are produced in very different ways).

- **Semanticity** – signals have meanings. If signals are simply random, no communication is taking place.
- **Pragmatic Function** – messages serve some purpose (e.g., survival, transmission of culture, influencing others’ behavior)

Communication systems have some or all of the following additional properties:

1. **Interchangeability** – a single individual has the ability both to send and to receive messages.  
E.g., the female silkworm moth can only send messages, and the males can only receive messages.
2. **Arbitrariness** – symbols used are arbitrary (conventional).

3. **Cultural Transmission** – the system can and must be learned (it is not completely innate).  
Human language: Children learn the language of their surrounding environment.
4. **Discreteness** – a message is composed of smaller, distinct functional parts.  
No animal communication system has so far been shown to have this property.  
Human language: phonemes → morphemes → words → phrases → sentences
5. **Displacement** – ability to communicate about things or situations not present in space or time.  
Human language: We can talk about yesterday, the surface on Mars, the Lord of the Rings (Sauron), the future, etc.
6. **Productivity** – the system is open, i.e., it can be expanded if necessary.  
Human language: Productivity (a completely new sentence) is allowed by discreteness (through the manipulation of parts).

## 2.2 Bees

- Honeybees communicate the location of food sources to other bees.
  - Three distinct methods (not all species use all of them). All are based on “dances” performed by scouts returning to the hive after having located food sources.
1. **Round dance** – Used when food is less than 7 m (20 ft) from the hive. It communicates:
    - the **existence** of food – by the very performance of the dance;
    - the **quality** of food – by the excitement level of the bee;
    - the **type** of food – by the dust on the bee.
  2. **Sickle dance** – Used when food is 7-20 m (20-60 ft) from the hive. It communicates:
    - the **existence** of food – as round dance
    - the **quality** of food – as round dance
    - the **type** of food – as round dance
    - the **direction** of the food wrt the hive and the sun – by angle of “perpendicular” wrt vertical
    - the **distance** of the food – the farther is the food the faster is the dance
  3. **Tail-wagging dance** – Used when food is more than 20 m (60 ft) from the hive. It communicates:
    - the **existence** of the food
    - the **quality** of the food
    - the **type** of the food

- the **direction** of the food – by the angle of the “straight” part of the dance wrt vertical
- the **distance** of the food – the farther is the food the slower is the dance Opposite strategy than in the sickle dance. Reason: The bees cannot dance so fast

## 2.3 Primate Studies

- The great apes (gorillas, chimpanzees, and orangutans) have very complex communication systems. They communicate with facial expressions, gestures, and calls to express anger, dominance, fear, danger, and the like.
- These communication systems nevertheless lack properties as displacement and productivity (they do not combine their gestures or calls in novel ways to create new meanings)
- The great apes are, however, very intelligent creatures and *Homo Sapiens*' nearest relatives.
- *Can language be taught to apes, even though it does not occur naturally?*

### 2.3.1 Nim Chimpsky

The early experiments showed (by failure) that apes are not capable of producing human speech sounds.

- Taught ASL to prove that a chimp could acquire and display some use of grammar
- Acquired 125 signs at the age of four

#### *Conclusion*

- The apes' uses of signs are very different from human language.
- 40% of the signs were mere repetitions of what the trainer had just signed.
- The signing was always a request for food or social reward.
- There was no evidence that Nim knew any grammar.

Many researches believe the conclusions are not correct.

### 2.3.2 Koko the Gorilla

- One of the longest project of its kind.
- In 1972, the one year old gorilla Koko started learning ASL.
- Now she uses around 1000 sings and can understand around 2000 signs.
- She also invents her own combinations of signs.
- In addition, she supposedly understands spoken English.

*Questions*

- Has Koko acquired human language?
- How long are her signed utterances?
- What are the main topics of her signing?
- Is there any evidence that she really knows the meaning of the signs she uses?
- Is she probably also only repeating and imitating her trainer's signs?

See: <http://www.pbs.org/wnet/nature/koko/> – Koko at PBS

# Chapter 3

## Phonetics

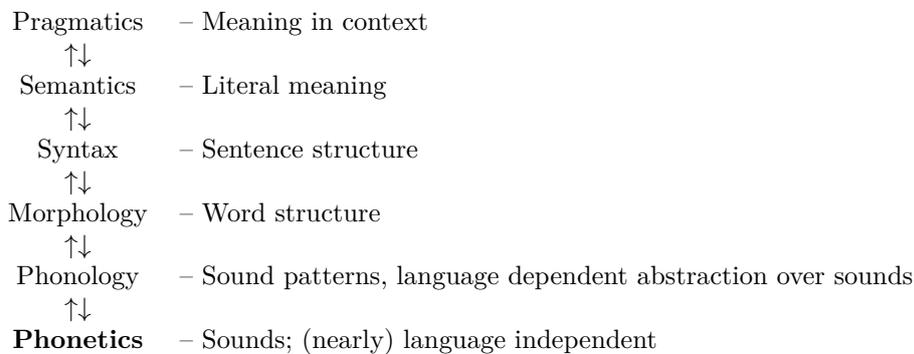
### 3.1 What is Phonetics

**Phonetics** is the study of speech sounds:

- how they are produced,
- how they are perceived,
- what their physical properties are.

The technical word for a speech sound is **phone** (hence, *phonetics*). Cf. *telephone, head-phone, phonograph, homophone*.

Place of phonetics in the language system:



↑ – understanding language expressions; ↓ – producing language expressions

### 3.2 Subfields of Phonetics

**Articulatory Phonetics** – the study of the production of speech sounds.

The oldest form of phonetics.

A typical observation: “The sound at the beginning of the word ‘foot’ is produced by bringing the lower lip into contact with the upper teeth and forcing air out of the mouth.”

**Auditory Phonetics** – the study of the perception of speech sounds.

Related to neurology and cognitive science.

A typical observation: “The sounds [s, ʃ, z, ʒ] are called sibilants because they share the property of sounding like a ‘hiss.’”

**Acoustic Phonetics** – the study of the physical properties of speech sounds.

A relatively new subfield (circa 50 years); uses sophisticated equipment (spectrograph, etc). Related to acoustics (the subfield of physics dealing with sound waves).

A typical observation: “The strongest concentration of acoustic energy in the sound [s] is above 4000 Hz.”

### 3.3 Phonetic Alphabet

Why do we need a new alphabet?

Because: We want to be able to write down how things are pronounced and the traditional Roman alphabet is not good enough for it:

- Words are pronounced differently depending on region, speaker, mood, ...but they are (usually) spelled the same way  
*root* [rut] or [rʊt], *truck* [trʌk] or [tʃrʌk], *strong* [strɔŋ] or [ʃtrɔŋ]
- Words or word forms sounding differently can be spelled the same way  
*read* [rid] vs. [rɛd]
- One sound is spelled many different ways:  
[k]: *king*, *card*, *clique*, *nick*, *chasm*, *exit*  
*reed* vs. *read*; *mě* ‘*I<sub>gen/acc</sub>*’ vs. *mně* ‘*I<sub>dat/loc</sub>*’, *tip* ‘tip’ vs. *typ* ‘type’
- There are many more sounds than there are letters in Latin alphabet  
There are only 6 vowel letters, but English has at least 10 vowel sounds  
*thigh*, *thy*  
There are many other sounds in other languages

What we want is a simple system where every symbol would correspond to exactly one sound.

IPA (International Phonetic Alphabet) - a special alphabet for representing sounds was developed. See: <http://www.langsci.ucl.ac.uk/ipa/index.html>

### 3.4 Describing Czech and English sounds

#### Consonants vs Vowels

- consonants – involve some constriction (closure/narrowing) at some point in the vocal tract
- vowels – do not have constriction; can always be held indefinitely.

### 3.4.1 Describing Consonants

Three-part description of consonants:

- **Voicing** – do vocal folds vibrate?
  - **voiced** – vocal folds vibrate  
[b], [d], [g] [m], [n], [ŋ], [z], [ʒ], etc.
  - **voiceless** – vocal folds are open and do not vibrate  
[p], [t], [k], [s], [ʃ], etc.
- **Manner of Articulation** – degree of the obstruction (narrowing, closure) & closure release type (sudden, slow), etc.
  - **Stops:** made by completely obstructing (“stopping”) the flow of air  
[p/b, t/d, c/ʃ, k/g, ʔ]  
[c]: *tisknout* ‘print’, *tapka* [capka] ‘paw’, Greek: [ceri] ‘candle’  
[ʃ]: *dábel* [ja:bɛl] ‘devil’, *děda* [jɛda] ‘grand-father’  
[ʔ] (glottal stop): *uh oh!* [ʊʔoʊ], *doopravit* [dɔʔɔp..] ‘finish repairing’  
English voiceless stops are either aspirated (*pít* [p<sup>h</sup>it]) or unaspirated (*spit* [spit])
  - **Fricatives:** made by forming a very narrow constriction and forcing air through, producing a hissing turbulent sound because of the friction between the air and the sides of the constriction.  
[f/v, θ/ð, s/z, ʃ/ʒ, x/χ, h/ħ]  
[θ]: *thick, thigh*  
[ð]: *then, thy*  
[ʃ]: *ship, šest* ‘six’  
[ʒ]: *visual, žába* ‘frog’  
[j]: *yes*  
[x]: *chleba* [xlɛba] ‘bread’; German *Bach*  
[χ]: *abych byl* [abɯɣɪl]; Spanish *digno* ‘worthy’  
[h] (voiceless) / [ħ] (voiced)  
Czech glottal fricative is usually voiced, English usually voiceless.
  - **Affricates:** stop immediately followed by a fricative.  
[ts / dz, tʃ / dʒ]  
[ts]: *cihla* [tʃɪfɪla] ‘brick’, German *Zug* ‘train’ (≈ as *ts* in *cats*)  
[dz]: *lecko* [lɛdʒɔdɔ] ‘various people’; sometimes in *Honza* [..ndʒa] x [..nza]  
[tʃ]: *change, český* ‘Czech’;  
[dʒ]: *journal, džbán* ‘pitcher’; sometimes in *manžel* [..ndʒɛl] x [..nʒɛl] ‘husband’  
Often written as [tʃ, dʒ, ...] or simply as [tʃ, dʒ, ...].
  - **Trills**
    - [r] (voiced alveolar trill): *krtek* ‘mole’
    - [r̥] (voiced (.) raised (.) alveolar trill): *dří* [dri] ‘labour/sweat<sub>imp</sub>’
    - [r̥̥] (voiceless (.) raised alveolar trill): *tři* [tri] ‘three’

- **Nasals (nasal stops):** the velum is lowered, air passes also through the nose.  
[m, ŋ, n, ŋ, ɲ]  
[ŋ]: sometimes in *comfort* [kʌŋfɔrt] x [..mf..], *tramvaj* [..ŋv..] x [..mv..] ‘tram’  
[ŋ]: *walking* [..kŋ] (some dialects [..km])  
[ɲ]: *nic* [ɲts] ‘nothing’; Spanish *caña* [kaɲa] ‘cane’ ( $\approx n$  in *annual*)
  - **Liquids:** narrow passage, but not narrow enough to cause friction (a la fricatives).  
[l, ɹ, r]  
[r] (flap): *vitamin* in U.S. English ( $\approx$  short [d]), some *r*’s in Czech
  - **Glides:** almost a vowel, but slightly more constricted. The least constricted type of consonant.  
[j, w]
- Since liquids and nasals are produced with a relatively open passage of air flow, they can be *syllabic*: *bird* [bɹd] (U.S.), *simple* [sɪmpɹl], *reason* [rɪzɹn]; *vlk* [vɹk] ‘wolf’, *krk* [krk] ‘neck’

- **Place of Articulation** – where the main obstruction is made (lips, teeth, velum, etc.)

- **Bilabial** (bi – two, labium – lip): the lips are close together or touching.  
[p/b, m, w]
- **Labiodental** (dental – teeth): the lower lip up is against the upper front teeth.  
[f/v]
- **Interdental** (inter – between): the tip of the tongue is between the front teeth  
[θ/ð]
- **Alveolar:** the tip of the tongue is at the alveolar ridge (the ridge just behind the teeth)  
[t/d, s/z, n, l, r, ɹ, r]
- **Palatal:** the tongue is near the hard palate (hard part of the roof of the mouth).  
[ʃ/ʒ; j; tʃ/dʒ]
- **Velar:** the tongue approaches the velum (soft part of the roof of the mouth)  
[k/g, ŋ]
- **Glottal:** the glottis is the point of constriction.  
[h/ɦ; ʔ]

### 3.4.2 Describing Vowels

Vowels are produced with a mostly open oral tract, so place/manner of articulation (a la consonants) is not useful in describing them.

#### English Vowels

In English, all vowels are usually *voiced*.

- **Tongue height:** high = near the roof of mouth  
Opening of mouth coincides more or less with tongue height.
  - **High:** [i, ɪ, u, ʊ] *leak, lick, luke, look*

- **Mid:** [e, ε, ə, ʌ, ɔ, o] *bait, bet, sofa, but, bought, boat*
- **Low:** [æ, a] *cat, cot*
- **Tongue advancement:** tongue further forward or back in mouth
  - **Front:** [i, ɪ, e, ε, æ] *seek, sick, sake, sec, sack*
  - **Back:** [u, ʊ, o, ɔ, a] *ooze, look, road, paw, dot*
  - **Central:** [ə, ʌ] *sofa, but*
- **Lip rounding:**
  - **rounded:** [u, ʊ, o, ɔ] *food, put, road, caught*
  - **unrounded:** the rest
- **Tenseness:** tongue position of the lax vowels are less extreme.
  - **Tense** – [i] (*deed*), [u] (*loose*)
  - **Lax** – [ɪ] (*did*), [ʊ] (*put*)

Tense vowels are longer than their lax counterparts – *deed* vs. *did*.

**diphthong** – a complex sound consisting of two vowel sounds.

[aɪ] (*right*), [ɔɪ] (*boy*), [eɪ] (*they*), [aʊ] (*laud*), [oʊ] – (*go*)

Note: English [e] occurs only in [eɪ] and [o] only in [oʊ]. Therefore the glide is often omitted in simplified transcriptions, because you know it is always there.

### Czech Vowels

short: [ɪ, ε, a, ʊ, ɔ]

long: [i:, ε:, a:, u:] and borrowed [ɔ:]

diphthongs: [oʊ] and borrowed [aʊ], [ɛʊ]

## 3.5 Some Other Speech Sounds

### 3.5.1 Vowels

#### Rounded Vowels

The only rounded vowels in Czech/English are the back vowels [u, ʊ, o, ɔ].

German and French both have front rounded vowels. They are written as [y] and [œ] in IPA. The high front rounded vowel [y] is pronounced like [i], but with rounded lips.

	Front rounded	Back rounded	Front unrounded
French:	[sy] sue ( <i>I sweat</i> )	[su] sous ( <i>under</i> )	[si] si ( <i>if/yes</i> )
	[nø] noed ( <i>knot</i> )	[no] nos ( <i>our</i> )	[ne] né ( <i>born</i> )

### Nasalized Vowels

- Oral vowels – more common; the nasal passage is closed (the velum is raised).
- Nasal vowel – like oral, but the nasal passage is open; marked by a tilde ([ẽ, ã, ...]).

	Oral		Nasal
French:	[mɛ] <i>mais</i> (but)	[mẽ]	<i>main</i> (hand)
	[lɛ] <i>lait</i> (milk)	[lẽ]	<i>lin</i> (linen)
	[ʃas] <i>chasse</i> (hunt)	[ʃãs]	<i>chance</i> (luck)
	[mɔ] <i>mot</i> (word)	[mõ]	<i>mon</i> (my)

In fact, there are nasal vowels in Czech and English – before nasal consonants, e.g. in *bin* [bĩn]. However, because they do not distinguish meaning (as they do in French), in a simplified transcription, their nasality is usually ignored ([bm]).

### 3.5.2 Consonants

- voiceless uvular stop [q]: Farsi (Persian, Iran)
- voiced bilabial fricative [β]: Spanish *Cuba*
- voiceless labial affricate [pf]: German *Pfennig* (penny)
- non-pulmonic consonants – sounds whose airflow is not dependent on the lungs:
  - clicks – in Khoisan languages and some Bantu languages in Africa
  - implosives
  - ejectives – in many native American languages and Caucasian languages

## 3.6 Narrow vs. Broad Transcription

Depending on the purpose of the transcription, we are either more detailed (the so called narrow transcription) or less (broad transcription). Usually we omit details that can be obtained by using simple and regular rules (e.g. all English word initial voiceless stops are aspirated).

In a really narrow transcription of English we have to capture all of the following (and much more). We ignore it in broad transcriptions:

- **Aspiration:** *pat* [p<sup>h</sup>æt] vs. *spat* [spæt]  
All English word initial voiceless stops are aspirated ([p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>])
- **Flaps:** *put* [pʊt] vs. *putting* [pʊtɪŋ], *ladder* [læɾər] = *latter* [læɾər]  
In American English, /t/ and /d/ are pronounced as [ɾ] between two vowels, where the first one is stressed.
- **Lengthened vowels:** *hat* [hæt] vs. *had* [hæ:d], *beat* [bit] vs. *bead* [bi:d]  
Vowels are slightly longer before voiced consonants.
- **Assimilations** – sounds tend to become similar to their neighbors:
  - *input* often as *imput* [ɪmpʊt], *lean bacon* often as *leam bacon* [lim beɪkən]

- **Nasalized vowels:** *bit* [bɪt] vs. *bin* [bɪ̃n]  
vowels are nasalized before nasals ([n, m, ŋ]).
- **Labiodental nasal:** [ɱ] in *comfort* [kʌmfɔːt]  
/m/ is pronounced as [ɱ] before labiodentals ([f, v]).
- **Deletions:** some sounds are omitted, esp. in fast speech and word-final consonants.  
*past* as *pas*, *lost shoe* [last ʃu] as *losh shoe* [lɑʃ ʃu] ([t] omitted, [s] assimilated to [ʃ])
- etc.

## 3.7 Links

- IPA (charts, sounds)  
<http://www.langsci.ucl.ac.uk/ipa/index.html>
- Customizable vocal tract showing the corresponding IPA symbol  
<http://www.chass.utoronto.ca/~danhall/phonetics/sammy.html>
- Interactive IPA charts linking each symbol to its pronunciation (note that consonants are surrounded by vowels)  
<http://www.yorku.ca/earmstro/ipa/index.html>
- List of various online phonetic resources:  
<http://www.unc.edu/~jlsmith/pht-url.html>



# Chapter 4

## Phonology

What to remember/understand:

Phonotactics, phoneme, [ ] vs. / /, minimal pair, phonological rule, assimilation, dissimilation, insertion, deletion,

### 4.1 What is Phonology

**Phonology:**

- studies how sounds are organized in particular languages
- tries to discover the psychological patterns and underlying organization of sounds shared by native speakers of a certain language.
- abstracts from the physical data provided by phonetics.

### 4.2 Phonotactics

Phonotactics studies what kind of sound patterns (sound combinations) are in a particular language and which are not.

For example, certain languages do not allow consonant clusters. Preference for CV-syllables (consonant+vowel) is fairly common, but some languages are more strict than others. Interesting thing happens with borrowings from other languages:

**Japanese:**

*besuboru* – baseball  
*gorufurendu* – girlfriend

*sutoraiku* – strike  
*arubaito* – job (German *Arbeit*)

**Setswana** (Botswana):

*keresemese* – Christmas      *gelase* – glass      *hafu* – half

**Shona** (Southern Bantu language, Zimbabwe, replacing [l] with [r]): Strictly CV (C even cannot be word final)

<i>turoko</i>	– truck	<i>puruvhu</i>	– proof
<i>furusetopi</i>	– full stop	<i>bhirifi</i>	– brief
<i>sitirecha</i>	– strecher	<i>giramu</i>	– gram
<i>hendibhegi</i>	– handbag	<i>kirimu</i>	– cream
<i>kanduro</i>	– candle	<i>bhirifi</i>	– breaf
<i>chitofu</i>	– stove		

Other examples of phonotactic constraints:

- Word initial stress – Czech, Hungarian, Finish, English (for most words)
- Word final obstruents (stops, fricatives, affricates) are voiceless – Czech, Polish, Russian, German, Dutch and many other languages.

### 4.3 Phonemes

It is sometimes difficult for native speakers of a language to tell the difference between sounds which may be completely distinct for speakers of another language.

- (3) a. English: *pit* [p<sup>h</sup>it] vs. *spit* [spit]  
 b. Hindi: [p<sup>h</sup>u:l] (fruit) vs. [pu:l] (moment)

- English speakers consider [p] and the [p<sup>h</sup>] to be the same sound, despite some irrelevant articulatory details.

For Hindi speakers, the same details are enough to completely differentiate the two sounds, making them as different as [p] and [b] for English speakers.

- In English, [p] and [p<sup>h</sup>] are called variants (allophones) of the same **phoneme** /p/.  
 In Hindi, [p] and [p<sup>h</sup>] are two distinct phonemes – /p/ and /p<sup>h</sup>/

You can think about phonemes as the stuff in your head, and phones as the real stuff you say. You know there is a /p/ in both *pit* and *spit*, but you pronounce [p<sup>h</sup>] in *pit* and [p] in *spit*.

So phonetics studies how sounds really sound, while phonology studies how they sound to speakers of some language.

Note:

- [ ] are used when capturing sounds in phonetics (encoding all the details)
- / / are used when capturing phonemes in phonology (disregarding details not relevant for a particular language)

Two sounds are called variants (allophones) of a single phoneme if:

- speakers of that language consider them to be one sound
- we can predict which one will be pronounced in a certain context (e.g. [p<sup>h</sup>] word initially, [p] otherwise); or we can say the choice is free

- (4) a. Czech: *sílí* [si:li:] ‘gets stronger’ vs. *šílí* [ʃi:li:] ‘is mad’  
 b. English: *see* [si] vs. *she* [ʃi]  
 c. Japanese: *saru* [saru] (monkey), *shiru* [ʃiru] (to know)
- In Czech and English, [s] and [ʃ] are not allophones, because there are words that are distinguished just by these two sounds.
  - In Japanese, [s] and [ʃ] are allophones, because there are no words distinguished just by these two sounds. The context determines which one is used. Before any vowel except [i], only [s] can occur; [ʃ] occurs in all other contexts. Thus in [-aru] we must choose [s] (*saru* ‘monkey’), but in [-iru] we must choose [ʃ] (*shiru* ‘to know’).

Different languages behave phonologically differently – they have different sets of phonemes and phonemes have different variants.

In U.S. English, /t/ has several variants (allophones) – [t], [t<sup>h</sup>], [ɾ] and [ʔ]:

	top	stop	little	kitten
(5) phonology:	/tap/	/stap/	/lɪtl/	/kɪtn/
phonetics:	[t <sup>h</sup> ap]	[stap]	[lɪɾtl]	[kɪʔn]

The best way how to find whether two sounds are two distinct phonemes or just variants of a single phoneme, is to try to find so called **minimal pair** – words that are distinguished only by these phones. If you find it, the sounds are *not* variants of a single phoneme.

[tʃ] vs. [dʒ]: *chin* [tʃɪn] vs. *gin* [dʒɪn]

[p] vs. [p<sup>h</sup>]: There is no such a pair

## 4.4 Phonological Rules

Phonological rules translate phonemes to the real sounds (phones).

- (6) a. /pɪt/ → [p<sup>h</sup>ɪt]  
 b. /spɪt/ → [spɪt]

So we can say:

- (7) a. /p/ → [p<sup>h</sup>] at the beginning of the word  
 b. /p/ → [p] otherwise (rules like this are usually omitted)

We can also state similar rules for /t/ and /k/:

- (8) a. /t/ → [t<sup>h</sup>] at the beginning of the word  
 b. /k/ → [k<sup>h</sup>] at the beginning of the word

However, /p/, /t/ and /k/ are all English voiceless stops, therefore we can write just one general rule:

(9) voiceless stop → aspirated at the beginning of the word

Or in a more “scientific” way:

(10) [-voiced, +stop] → [+aspirated] / # —

Note: # marks word boundary (# — means word initially, — # means word finally)

Languages have many rules like that. Some of them all the speakers share, some are used only by some speakers. Some of them occur always, some only in fast speech, etc.

In these rules we can refer to classes of phonemes like:

- voiced consonants ([b, d, ʒ, g, ð, z, ʒ, n, m, ...]),
- rounded vowels ([u, ʊ, o, ɔ]),
- nasals ([m, n, ŋ, ɲ]),
- sibilants (hissy sounds [s, z, ʃ, ʒ, tʃ, dʒ]),
- etc.

In Czech or German, all word-final obstruents become voiceless:

(11)

a.	/hund/	→ [hunt]	<i>Hund</i>	(dog)
b.	/hunde/	→ [hunde]	<i>Hunde</i>	(dogs)
c.	/tag/	→ [tak]	<i>Tag</i>	(day)
d.	/tage/	→ [tage]	<i>Tage</i>	(days)

The rule:

(12)

In some Spanish dialects, voiced stops change to fricatives if surrounded by vowels:

(13)

a.	/la beya dama/	→ [la βeya ðama]	<i>la bella dama</i>
b.	/la dama es beya/	→ [la ðama es βeya]	<i>la dama es bella</i>
c.	/la baka/	→ [la βaka]	<i>la vaca</i>
d.	/las bakas/	→ [las bakas]	<i>las vacas</i>
e.	/el dwεño/	→ [el dwεño]	<i>el dueño</i>
f.	/el buño/	→ [el buño]	<i>el burro</i>

The rule:

(14)

## 4.5 Kinds Of Phonological Rules

Different languages have different rules, however there are some some typical kinds of rules that are very common:

- **Assimilation** – a process by which a sound becomes more like a nearby sound.
  - voicing assimilation: *kdo* [gdɔ] ‘who’, *vše* [fʃɛ] ‘all’ – a consonant must match the following consonant in voicing.
  - place assimilation:
    - comfort* [mʃf], *tramvaj* [mʃv] – /m/ can assimilate to /f/ → [mʃ] (bilabial → labiodental)
    - input* [mp], *bonbon* [mb] ‘candy’
  - *bit* [ɪ] vs. *bin* [i] – /ɪ/ assimilates to the following /n/ (nasal) → [i]
  - *because you* [bɪkɔʒju] – /z/ can assimilate to /j/ (palatal) → [ʒ]

Reason: easier to pronounce; the assimilation level depends on speakers and situation

- **Dissimilation** – the opposite of assimilation, two nearby sounds become less alike.
  - In Latin, suffix *-alis* changes to *-aris* when it is added to a word containing [l]. These words came into Czech/English as adjectives ending in *-al* or *-ar*.
    - al*: *anecdotal*, *annu-al*; *natur-ální*, *manu-ální*
    - ar*: *angul-ar*, *annul-ar*; *mol-ární*, *plan-ární* (\**plan-ální*)
  - Here the change is even reflected in spelling.

- **Insertion** – a new sound is inserted.
  - *prince* /prms/ → [p<sup>h</sup>rm(t)s]: [t] is sometimes inserted
  - *hamster* /hæmstr/ → [hæm(p)str]: [p] is sometimes inserted

Reason: Difference of timing of various articulators. For example, in *prince* the velum is already positioned for pronouncing [s], but the tongue is in place for [n], and [t] results.

- **Deletion** – a phoneme is not pronounced in certain environments
  - *hřebík* [ɦrɛbi:k] → [rɛbi:k] ‘nail’ (optional)
  - *pojď* [pɔjc] → [pɔc] ‘come’ (optional)
  - *okay* [okey] → [ʔkey] (optional)

Reason: easier and faster to say

- **Metathesis** – two sound (usually adjacent) switch their place.
  - *vel+ryba* ‘whale’ (lit: big fish) [vɛlɹɪba], sometimes [vɛrɪlɪba]
  - obligatory in Leti: /ukar + ppalu/ → [ukrapalu] ‘index finger’, /morut + kdieli/ → [mortukdjeli] ‘very curly hair’

Rules may be **obligatory** (all speakers do it; e.g., final devoicing in Czech or nasalization of vowels in Czech/English) or **optional** (sometimes or some speaker do it; e.g., insertions/deletions above)



# Chapter 5

## Morphology

### 5.1 Basic terminology

- **Morphology** – study of internal structure of words
- **Morpheme** – the smallest linguistic unit which has a meaning or grammatical function. Words are composed of morphemes (one or more). There are some complications with this simple definition.

*sing·er·s, home·work, moon·light, un·kind·ly, talk·s, ten·th, flipp·ed, de·nation·al·iz·ation*

The order of morphemes matters:

*talk·ed* ≠ *\*ed·talk*, *re·write* ≠ *\*write·re*

- **Morph**. The term morpheme is used both to refer to an abstract entity and its concrete realization(s) in speech or writing. When it is needed to maintain the signified and signifier distinction, the term morph is used to refer to the concrete entity, while the term morpheme is reserved for the abstract entity only.
- **Allomorphs** – morphemes having the same function but different form. Unlike the synonyms they usually cannot be replaced one by the other.

(15) a. indefinite article: *an orange* – *a building*

b. plural morpheme: *cat·s* [s] – *dog·s* [z] – *judg·es* [əʃ]

(16) a. *matk·a* ‘mother<sub>nom</sub>’ – *matek* ‘mothers<sub>gen</sub>’ – *matc·e* ‘mother<sub>dat</sub>’ – *matč·in* ‘mother’s’

### 5.2 Classification Of Morphemes

#### 5.2.1 Bound × Free

- **Bound** – cannot appear as a word by itself.  
*-s* (*dog·s*), *-ly* (*quick·ly*), *-ed* (*walk·ed*);  
*-te* (*dělá·te* ‘do<sub>2pl</sub>’), *-y* (*žen·y* ‘women’), *vy-* (*vy·jít* ‘walk out’)

- **Free** – can appear as a word by itself; often can combine with other morphemes too.  
*house* (*house·s*), *walk* (*walk·ed*), *of*, *the*, *or*  
*hrad* ‘castle’, *žen* ‘woman<sub>root</sub> = gen.pl.’, *přes* ‘over’, *nebo* ‘or’

Past tense morpheme is a bound morpheme in English (*-ed*) but a free morpheme in Mandarin Chinese (*le*)

- (17) a. *Ta chi le fan.*  
 He eat past meal.  
 ‘He ate the meal.’  
 b. *Ta chi fan le.*  
 He eat meal past.  
 ‘He ate the meal.’

### 5.2.2 Root × Affix

- **root** – nucleus of the word that affixes attach too.

In English, most of the roots are free. In some languages that is less common (Lithuanian: *Billas Clintonas*).

Compounds contain more than one root: *home·work*; *železo·beton* ‘reinforced concrete’

- **affix** – a morpheme that is not a root; it is always bound
  - **suffix**: *talk·ing*, *quick·ly*; *mal·ý* ‘small<sub>masc.sg.nom</sub>’, *kup·ova·t* ‘buy<sub>imperf</sub>’
  - **prefix**: *un·happy*, *pre·existing*; *do·psat* ‘finish writing’, *nej·méně* ‘least’
  - **infix**: common in Austronesian and Austroasiatic lgs (Tagalog, Khmer)  
 Tagalog: *basa* ‘read’ *b·um·asa* ‘read<sub>past</sub>’ – *sulat* ‘write’ – *s·um·ulat* ‘wrote’  
 very rare in English: *abso·bloody·lutely*,
  - **circumfix**: morpheme having two parts that are placed around a stem  
 Dutch collectives:  
*berg* ‘mountain’ *ge·berg·te* ‘mountains’ \**geberg*, \**bergte*  
*vogel* ‘bird’ *ge·vogel·te* ‘poultry’ \**gevogel*, \**vogelte*  
 Czech *po+...+í*:  
*Vltava* → *Po·vltav·í* ‘Vltava river area’ (\**povltava*, \**vltaví*);  
*Pobaltí*, *pohoří*, *pohraničí*, *potrubí*, *pobřeží*, *polesí*

Suffixes more common than prefixes which are more common than infixes/circumfixes

### 5.2.3 Content × Functional

- **Content** morphemes – carry some semantic content  
*car*, *-able*, *un-*
- **Functional** morphemes – provide grammatical information  
*the*, *and*, *-s* (plural), *-s* (3<sup>rd</sup> sg)  
*jsem* ‘past aux<sub>1sg</sub>’, *-a* ‘gen.sg’ (*měst·a* ‘town<sub>gen</sub>’)

### 5.2.4 Derivation vs. Inflection

- **inflection** – creating various forms of the same word

**lexeme** – an abstract entity; the set of all forms related by inflection (but not derivation).

*table* – *table·s*

*uč·í·m* – *uč·í·š* – *uč·í* – *uč·í·me*

**lemma**: A form from a lexeme chosen by convention (e.g., nom.sg. for nouns, infinitive for verbs) to represent that set.

Also called the canonical/base/dictionary/citation form.

E.g., *break*, *breaks*, *broke*, *broken*, *breaking* have the same lemma *break*

**ending** – inflectional suffix

- **derivation** – creating new words

*slow* – *slow·ly* – *slow·ness*

*uč·í·t* – *uč·í·tel* – *uč·í·tel·ka* – *uč·í·tel·ský* – *uč·í·tel·ova·t* – *vy·uč·ova·t*

Inflection vs. Derivation:

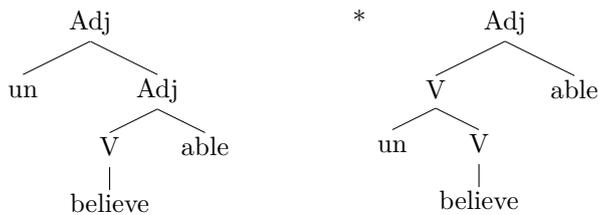
- Derivation tends to affect the meaning of the word, while inflection tends to affect only its syntactic function.
- Derivation tends to be more irregular – there are more gaps, the meaning is more idiosyncratic and less compositional.
- However, the boundary between derivation and inflection is often fuzzy and unclear.

## 5.3 Structure of words

Structure of words can be captured in a similar way as structure of sentences.

(18) unbelievable =

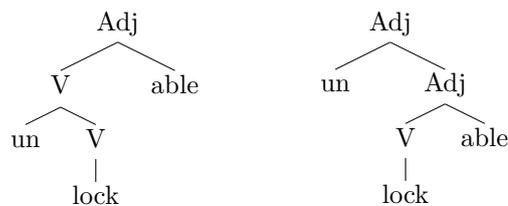
un + (believ + able), not \*(un + believe) + able



Some words can be ambiguous:

(19) unlockable =

un + (lock + able) or (un + lock) + able



## 5.4 Morphological processes

- **Concatenation** (adding continuous affixes) – the most common process  
Often phonological changes on morpheme boundaries.
- **Reduplication** – part of the word or the entire word is doubled:
  - Tagalog: *basa* ‘read’ – *ba·basa* ‘will read’; *sulat* ‘write’ – *su·sulat* ‘will write’
  - Afrikaans: *amper* ‘nearly’ – *amper·amper* ‘very nearly’; *dik* ‘thick’ – *dik·dik* ‘very thick’
  - Indonesian: *oraj* ‘man’ – *oraj·oraj* ‘all sorts of men’ (Cf. *orangutan*)
  - Samoan:
 

<i>alofa</i>	‘love <sub>Sg</sub> ’	<i>a·lo·lofa</i>	‘love <sub>Pl</sub> ’
<i>galue</i>	‘work <sub>Sg</sub> ’	<i>ga·lu·lue</i>	‘work <sub>Pl</sub> ’
<i>la:poʔa</i>	‘to be large <sub>Sg</sub> ’	<i>la:·po·poʔa</i>	‘to be large <sub>Pl</sub> ’
<i>tamoʔe</i>	‘run <sub>Sg</sub> ’	<i>ta·mo·moʔe</i>	‘run <sub>Pl</sub> ’
  - English: *humpty·dumpty*
  - American English (borrowed from Yiddish): *baby·schmaby*, *pizza·schmizza*
- **Templates** – both root and affix  
Both the roots and affixes are discontinuous. Only Semitic lgs (Arabic, Hebrew). A root (3 or 4 consonants, e.g., *l-m-d* – ‘learn’) is interleaved with a (mostly) vocalic pattern
  - Hebrew:
 

<i>lomed</i>	‘learn <sub>masc</sub> ’	<i>shatak</i>	‘be-quiet <sub>pres.masc</sub> ’
<i>lamad</i>	‘learnt <sub>masc.sg.3rd</sub> ’	<i>shatak</i>	‘was-quiet <sub>masc.sg.3rd</sub> ’
<i>limed</i>	‘taught <sub>masc.sg.3rd</sub> ’	<i>shitek</i>	‘made-sb-to-be-quiet <sub>masc.sg.3rd</sub> ’
<i>lumad</i>	‘was-taught <sub>masc.sg.3rd</sub> ’	<i>shutak</i>	‘was-made-to-be-quiet <sub>masc.sg.3rd</sub> ’
- **Morpheme internal changes** (apophony, ablaut) – the word changes internally
  - English: *sing* – *sang* – *sung*, *man* – *men*, *goose* – *geese* (not productive anymore)
  - German: *Mann* ‘man’ – *Männ·chen* ‘small man’, *Hund* ‘dog’ – *Hünd·chen* ‘small dog’
  - Czech: *kráva* – *krav*, *nés·t* – *nes·u* – *nos·ím*
- **Subtraction (Deletion)**: some material is deleted to create another form
  - Papago (a native American language in Arizona) imperfective → perfective
 

<i>him</i>	‘walking <sub>imperf</sub> ’	→	<i>hi</i>	‘walking <sub>perf</sub> ’
<i>hihim</i>	‘walking <sub>pl.imperf</sub> ’	→	<i>hihi</i>	‘walking <sub>pl.perf</sub> ’
  - French, feminine adjective → masculine adj. (much less clear)
 

<i>grande</i> [grād]	‘big <sub>f</sub> ’	→	<i>grand</i> [grā]	‘big <sub>m</sub> ’
<i>fausse</i> [fos]	‘false <sub>f</sub> ’	→	<i>faux</i> [fo]	‘false <sub>m</sub> ’
- **Suppletion** – ‘irregular’ relation between the words. Hopefully quite rare.
  - Czech: *být* – *jsem*, *jít* – *šla*, *dobrý* – *lepší*
  - English: *be* – *am* – *is* – *was*, *go* – *went*, *good* – *better*

## 5.5 Word formation

- **Affixation** – words are formed by adding affixes.

English:

- V + *-able* → Adj: *predict-able*
- V + *-er* → N: *sing-er*
- *un-* + A → A: *un-productive*
- V + *-en* → V: *deep-en, thick-en*

Czech:

- N + *-ov-ý* → Adj: *motor-ový*
- V + *-tel* → N: *spisova-tel, stavi-tel*
- N + *-ova-t* → V: *pan-ova-t, parazit-ova-t,*

- **Compounding** – words are formed by combining two or more words.

English:

- Adj + Adj → Adj: *bitter-sweet*
- N + N → N: *rain-bow, Internet Security Association Key Management Protocol*
- V + N → V: *pick-pocket*
- P + V → V: *over-do*

Czech:

- N + N → N: *maso-žravec* ‘carnivore’, *vzducho-loď* ‘airship’
- A + N → N: *černo-zem* ‘black soil’, *plno-vous* ‘beard’
- A(dv) + A → A: *star-o-česk-ý* ‘Old Czech’, *tmav-o-modr-ý* ‘dark blue’

German: *Donau-dampf-schiff-fahrts-gesellschafts-kapitän*

- **Acronyms** – like abbreviations, but acts as a normal word

*laser* – *light amplification by simulated emission of radiation*  
*radar* – *radio detecting and ranging*  
*Čedok* – *Česká dopravní kancelář*

- **Blending** – parts of two different words are combined

*breakfast + lunch* → *brunch*  
*smoke + fog* → *smog*  
*motor + hotel* → *motel*

- **Clipping** – longer words are shortened

*doc(tor)*, *prof(essional)*, *lab(oratory)*, *ad(vertisement)*, *dorm(itory)*, *exam(ination)*  
*auto(mobil)*

## 5.6 Morphological Types Of Languages

Two basic morphological types of languages:

- **Analytic** (isolating) languages – have only free morphemes, sentences are sequences of single-morpheme words.
- **Synthetic** languages – both free and bound morphemes. Affixes are added to roots.

### 5.6.1 Analytic languages

Analytic languages have only free morphemes, sentences are sequences of single-morpheme words.

(20) Vietnamese:

khi tôi đến nhà bạn tôi, chúng tôi bắt đầu làm bài (Comrie 1989)  
 when I come house friend I, PLURAL I begin do lesson

‘When I came to my friend’s house, we began to do lessons.’

### 5.6.2 Synthetic languages

Synthetic languages have both free and bound morphemes.

Has further subtypes:

- **Agglutinating** – each morpheme has a single function, it is easy to separate them.  
 E.g., Uralic lgs (Estonian, Finnish, Hungarian), Turkish, Basque, Dravidian lgs (Tamil, Kannada, Telugu), Esperanto

Turkish:

	sg.	pl.	
nom.	ev	ev·ler	‘house’
gen.	ev·in	ev·ler·in	
dat.	ev·e	ev·ler·e	
acc.	ev·i	ev·ler·i	
loc.	ev·de	ev·ler·de	
ins.	ev·den	ev·ler·den	

- **Fusional** – like agglutinating, but affixes tend to “fuse together”, one affix has more than one function. Common homonymy of inflectional affixes.

*matk·a* ‘mother’ – *a* means the word is a noun, feminine, singular, nominative.

E.g., Slavic, Romance languages, Greek

(21) Homonymy of the *a* ending in Czech:

form	lemma	gloss		category
měst-a	město	town	NS2	noun neut sg gen
			NP1 (5)	noun neut pl nom (voc)
			NP4	noun neut pl acc
tém-a	téma	theme	NS1 (5)	noun neut sg nom (voc)
			NS4	noun neut sg acc
žen-a	žena	woman	FS1	noun fem sg nom
pán-a	pán	man	MS2	noun masc anim sg gen
			MS4	noun masc anim sg acc
ostrov-a	ostrov	island	IS2	noun masc inanim sg gen
předsed-a	předseda	president	MS1	noun masc anim sg nom
vidě-l-a	vidět	see		verb past fem sg
				verb past neut pl
vidě-n-a				verb passive fem sg
				verb passive neut pl
vid-a				verb transgressive masc sg
dv-a	dv-a	two		numeral masc sg nom
				numeral masc sg acc

(22) Ending *-e* and noun cases in Czech:

case	form	lemma	gender	gloss
nom	kuř-e	kuře	neuter	chicken
gen	muž-e	muž	masc.anim.	man
dat	mouš-e	moucha	feminine	fly
acc	muž-e	muž	masc.anim.	man
voc	pan-e	pán	masc.anim.	mister
loc	mouš-e	moucha	feminine	fly
inst	–	–		

- **Polysynthetic:** extremely complex, many roots and affixes combine together, often one word corresponds to a whole sentence in other languages.

*angyaghllangyugtuq* – 'he wants to acquire a big boat' (Eskimo)

*palyamunurringkutjamunurtu* – 's/he definitely did not become bad' (W Aus.)

Sora – LF, p. 132

### 5.6.3 Morphology in real languages

Czech – mostly fusional, but also other properties:

- analytic: future and past tense, conditional, prepositions, ...
- agglutinating: prefixes/suffixes; *vidě·n·a* 'seen<sub>fem.sg</sub>' -n- – passive, -a – fem+sg

English – originally fusional, but now both analytic properties (future morpheme *will*, perfective morpheme *have*, etc. are separate words) and synthetic properties (plural (*·s*), etc. are bound morphemes)

(23) The degree of synthesis of some languages (Haspelmath 2002):

Language	Ration of morphemes per word
Greenlandic Eskimo	3.72
Sanskrit	2.59
Swahili	2.55
Old English	2.12
Lezgian	1.93
German	1.92
Modern English	1.68
Vietnamese	1.06



# Chapter 6

## Syntax

What to remember and understand:

Syntax, difference between syntax and semantics, open/closed class words, all word classes (and be able to distinguish them based on morphology and syntax)

Subject, object, case, agreement.

### 6.1 What is Syntax?

**Syntax** – the part of linguistics that studies sentence structure:

- word order:  
*I want these books.*  
*\*want these I books.*
- agreement – subject and verb, determiner and noun, ... often must agree:  
*He wants this book.*  
*\*He want this book.*  
*I want these books.*  
*\*I want this books.*
- How many complements, which prepositions and forms (cases):  
*I give Mary a book.*  
*\*I see Mary a book.*  
*I see her.*  
*\*I see she.*
- hierarchical structure – what modifies what  
*We need more (intelligent leaders).* (more of intelligent leaders)  
*We need (more intelligent) leaders.* (leaders that are more intelligent)
- etc.

Syntax is not about meaning! Sentences can have no sense and still be grammatically correct:

*Colorless green ideas sleep furiously.* – nonsense, but grammatically correct

*\*Sleep ideas colorless furiously green.* – grammatically incorrect

Syntax: From Greek *syntaxis* from *syn* (together) + *taxis* (arrangement).

Cf. *symphony, synonym, synthesis; taxonomy, tactics*

## 6.2 Parts of Speech

- Words in a language behave differently from each other.
- But not each word is entirely different from all other words in that language.

⇒ Words can be categorized into **parts of speech** (lexical categories, word classes) based on their morphological, syntactic and semantic properties.

Note that there is a certain amount of arbitrariness in any such classification. For example, should *my* be classified as a pronoun or as a determiner, should numerals/participles/auxiliary verbs be a separate category. If *he* is a pronoun, should *do* be a pro-verb?

Open versus closed classes:

- **Open class** – new and new items are added to the class over the time – nouns, verbs, adjectives, and adverbs.
- **Closed class** – contains small number of words, new items are added very rarely – determiners, pronouns, prepositions, conjunctions

### 6.2.1 Open Classes

#### Nouns (N)

- morphology – most nouns form plural  
*cat-s, house-s, kiss-es, men, sheep*
- syntax – usually has a determiner (except proper names like *John*) and can be modified by an adjective:  
Determiner (Adjective) \_\_\_  
*a cat, many kisses, few men, several sheep*  
*a small cat, many exciting kisses, few clever men, several bored sheep*
- semantics – name of a person, thing or place.  
But: problem with abstract nouns (*beauty, anger, aspect*) and actions (*a thump*).

#### Verbs (V)

- morphology – form third person, past tense, past participle, present participle:  
*walk walks walked walked walking*  
*go goes went gone going*  
*buy buys bought bought buying*  
*run runs ran run running*
- syntax:
  1. can combine with an auxiliary: Aux \_\_\_  
*will go, have seen, should run, must leave, is swimming*
  2. can be modified by an adverb: Adverb \_\_\_ or \_\_\_ Adverb  
*usually sleep, read carefully*

- semantics – usually describes an action, a process or a state of being  
But: problem with some verbs (*know, remember*)  
But: hard to distinguish from nouns describing actions (*a thump*)

### Adjectives (A, Adj)

- morphology:
  1. form comparative and superlative forms:  
*cool – cooler – coolest, successful – more successful – most successful, good – better – best*
  2. many can be changed into adverbs by the *-ly* suffix:  
*sad – sadly, funny – funnily, nice – nicely, beautiful – beautifully*
- syntax:
  1. Can modify a noun: Determiner \_\_\_ N  
*a tall man, a cool day, a wonderful trip*
  2. Can be modified by an adverb: Adverb \_\_\_  
*very clever, extremely clever, unusually hot*
- semantics – usually describes a quality or attribute

### Adverbs (Adv)

*quickly, soon, morally, today, here, very, before*

- morphology – often formed from adjectives by the *-ly* suffix.  
But: *good – well, fast – fast, friendly – friendly, eastward*
- syntax – can modify verbs (hence ad-verbs), adjectives & often other adverbs.
  1. V \_\_\_ or \_\_\_ V  
*stop abruptly, usually eat*
  2. \_\_\_ Adj  
*amazingly cheap, very bad* (very cannot modify a verb), *quite nice*
  3. \_\_\_ Adv  
*very quickly, quite soon*

### Summary of open classes

	Typical Morphology	Typical Syntax	Typical Semantics
Noun	plural <i>house – houses</i>	D (Adj) ___ <i>the big house</i>	thing, person, place
Verb	tenses, ... <i>walk – walked</i>	combines with an Aux <i>would walk</i>	action
Adj	comparative, superlative <i>big – bigger – biggest</i>	D ___ N <i>the big house</i>	quality, property
Adverb	often has <i>-ly</i> suffix <i>really</i> , but: <i>well</i>	modifies V, Adj, Adv <i>a really big house</i>	manner, degree, ...

### 6.2.2 Closed Classes

#### Determiners (D, Det)

articles (*a, the*), quantifiers (*many, any, all, several*), possessives (*my, your, his, her*)

Syntax – come before nouns: — (Adj) N

#### Auxiliary verbs (Aux)

*will, may, must, shall, would, can, have*

Syntax:

1. Is followed by a verb: — V  
*It will rain. You must be quiet.*
2. Is negated directly: — *not*  
*He cannot swim. She would not come.*  
*\*He doesn't can swim. \*She doesn't would come.*

Normal verbs vs. Auxiliary verbs:

	Verbs	Auxiliary verbs
negation	needs aux <i>do</i> <i>I don't want it.</i> <i>*I want not it.</i>	directly <i>*I don't will come.</i> <i>I will not come.</i>
question	needs aux <i>do</i> <i>Do you know it?</i> <i>Know you it?</i>	inversion <i>*Do you will come?</i> <i>Will you come?</i>
agreement	yes <i>He knows it.</i> <i>*He know it.</i>	no <i>*He wills come</i> <i>He will come</i>

*Be, have* and *do* are sometimes normal verbs and sometimes auxiliary verbs. (*I don't have it.* vs. *I have not seen it.*)

#### Pronouns (Pron)

Words that stand for a noun or a whole noun phrase.

*I, you, he, she, it, we, they, me, him, her, us, them*

Note: It makes sense to classify possessives (traditionally called possessive pronouns) as determiners. Syntactically, pronouns and possessives behave differently – pronouns act as nouns, but possessives modify nouns:

pronoun: *I run.* – *\*My run.*

possessive: *John likes my house.* – *\*John likes I house.*

pronoun: Based on Latin *pro* (*for*) + *noun*

#### Prepositions (P)

*in, on, about, with, at, to, of, under*

Syntax: stand before noun phrases (see later, simply NP = Det (Adj) N)

Semantics: usually express spacial, temporary, etc. relations.

*on the table, with nice colors, about mammals*

### Conjunctions (Conj)

*and, or, but, ...*

Syntax: connect two words or phrases on the same level

1. N — N (*women and men*)
2. V — V (*run or walk*)
3. Adj — Adj (*warm but rainy*)
4. S — S (*I will talk and he will write.*)
5. etc.

### 6.2.3 Words belonging to more than one lexical category

Some words belong to several categories:

- They have similar meaning:
  - *She talks very much.* vs. *She is giving three talks.*
  - *It's cold.* vs. *I got a cold.*
- They have completely different meanings:
  - *I can't bear the noise.* vs. *There is a bear in the wood.*
  - *tree barks* vs. *the dog barks*

Virtually all English nouns can be turned into verbs.

## 6.3 Phrases, Constituents & Phrase Structure Rules

### 6.3.1 Describing Noun Phrases

In English, a noun phrase a determiner followed by a noun, or determiner followed by an adjective followed by a noun, or a single noun, or ...

To save words, we can use the so called Phrase Structure Rules capture this:

- (24)
- a. NP → Det N     *the cat*
  - b. NP → Det A N   *those noisy cats*
  - c. NP → N         *cats*
  - d. NP → A N       *noisy cats*

We can mark optional subphrases with parentheses and save even more words:

- (25) NP → (Det) (A) N   *cats, noisy cats, the cat, those noisy cats*

A phrase structure rule tell us two things:

- Which smaller phrases (Det, A, N) use to build a bigger phrase (NP).
- How to order the smaller phrases – the rule (25) allows *noisy cats*, but not *cats noisy*

$$(26) \quad \overbrace{\underbrace{\text{those}}^{\text{Det}} \underbrace{\text{noisy}}^{\text{A}} \underbrace{\text{cats}}^{\text{N}}}^{\text{NP}}$$

In addition, a pronoun can be a noun phrase:

$$(27) \quad \text{NP} \rightarrow \text{Pron} \quad \textit{she, you, ...}$$

### 6.3.2 Describing Prepositional phrases

In English, preposition is usually followed by a noun phrase (let's ignore the prepositions at the end of the sentence).

$$(28) \quad \text{PP} \rightarrow \text{P NP} \quad \textit{about those noisy cats}$$

Now we can put that together and say things like:

$$(29) \quad \overbrace{\underbrace{\text{about}}^{\text{P}} \overbrace{\underbrace{\text{those}}^{\text{Det}} \underbrace{\text{noisy}}^{\text{A}} \underbrace{\text{cats}}^{\text{N}}}^{\text{NP}}}^{\text{PP}}$$

### 6.3.3 Describing Sentences

In English, a sentence consists of a subject (usually a noun phrase) followed by a verb which is sometimes followed by an object (another noun phrase), prepositional phrases etc.

- (30) a. *Alphons slept.* – Subject + V  
 b. *Alphons saw his dog.* – Subject + V + Object  
 c. *Alphons asked for a beer.*  
 d. *Alphons begged beer from his dog.*
- (31) a.  $S \rightarrow \text{NP V}$  – *Alphons slept*  
 b.  $S \rightarrow \text{NP V NP}$  – *Alphons saw his dog*  
 c.  $S \rightarrow \text{NP V PP}$  – *Alphons asked for a beer*  
 d.  $S \rightarrow \text{NP V NP PP}$  – *Alphons asked his dog for a beer*

We can abbreviate these rules as:<sup>1</sup>

<sup>1</sup>Of course, we ignored many other sentences like:

- (i) a.  $S \rightarrow \text{NP V NP NP}$  – *Alphons offered his dog some beer*  
 b.  $S \rightarrow \text{NP V PP PP}$  – *Alphons argued with his dog about beer*  
 c.  $S \rightarrow \text{NP V NP InfP}$  – *Alphons persuaded his dog to buy some beer*  
 d.  $S \rightarrow \text{NP V NP that S}$  – *Alphons persuaded his dog that it would be wise to bring beer*

Linguists often distinguish between sentences and verb phrases (VP). A verb phrase is a sentence without a subject (e.g. *saw his dog*). Then you have to describe sentence in two steps: First,  $S \rightarrow \text{NP VP}$  and then  $\text{VP} \rightarrow \text{V (NP) (PP)}$ .

(32)  $S \rightarrow NP V (NP) (PP)$

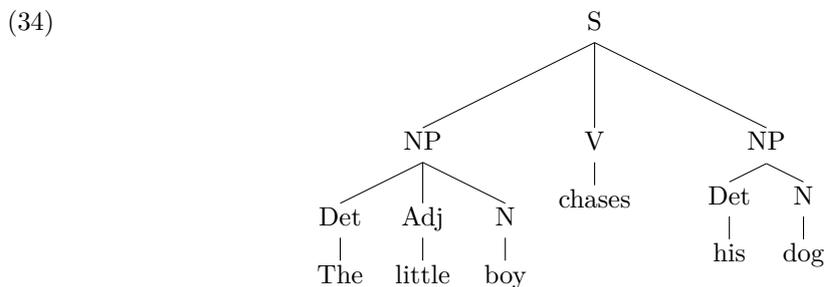
This rule says: Sentence is a noun phrase followed by a verb and possibly some other noun phrase and/or prepositional phrase. For example:

(33)  $\overbrace{\underbrace{Alphons}_{NP} \underbrace{saw}_{V} \underbrace{his\ dog}_{NP}}^S$

### 6.3.4 Phrase structure trees

- Phrases are created from other phrases or words.
- Sentence is the biggest phrase.

We can depict the fact that a sentence is built from smaller parts by a diagram:



A tree diagram represents several aspects of “how words are put together” in a sentence:

- the order of the words in a sentence.
- the word class (Part of Speech) of each word.
- the hierarchical structure of a sentence – the grouping of words into phrases, and the grouping of phrases into larger phrases.
- the centers of phrases that other words group around (e.g. N in NP, V in S)

### 6.3.5 Phrase/Constituent

Intuitively, a **constituent** (phrase) is a group of words which “belong together” in a sentence. They are usually coherent by themselves (i.e., when taken out of the context of the sentence) and make a coherent contribution to the meaning of the sentence as a whole.

- (35) a. The dog ate the bone.  
 b. The president of the company likes to see big profits.  
 c. My stupid kid brother told my mom about my F in algebra.

Always relative to a given sentence. What is a constituent in one sentence is not necessarily a constituent in another sentence.

- (36) a. I reminded *the president of the company* policy regarding smoking which he himself had instated.  
 b. When she was *told*, *my mom* broke into uncontrolled sobbing.

### Constituency Tests

- Ability to stand alone
- Constituents can be replaced by pro-forms (pronouns, *do*)
  - (37) a. HE ate the bone.
  - b. The dog ate IT.
  - c. HE likes to see THEM.
  - d. The president likes to see SUCH profits.
  - e. My stupid kid brother DID.
- (38) a. \* The dog it/such/that bone.
- b. \* The that/such/did company likes to see big profits.
- Movement – (some) constituents can be moved, non-constituents cannot.
  - (39) a. The bone ,  
          the dog ate.
  - b. Big profits ,  
          the president of the company likes to see.
- Clefts – only constituents can form cleft sentences
  - (40) a. What the dog did was [eat the bone].
  - b. What the dog ate was [the bone].
  - c. It is [the president of the company] who likes to see big profits.

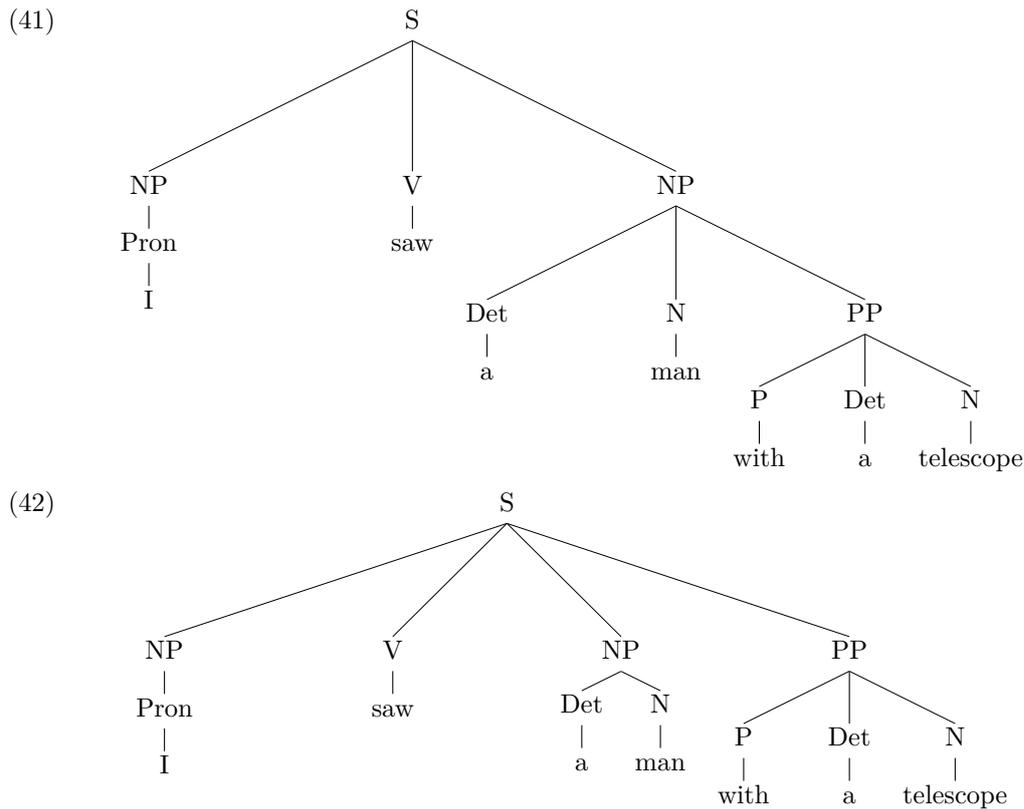
**Not always clear. Some strings pass some tests but not all.**

## 6.4 Ambiguity

Ambiguity:

- syntactical – more than one possible structure for the same string of words.  
*I saw a man with a telescope.* (Who has the telescope, me or the man?)  
*We need more intelligent leaders.* (*need more* or *more intelligent*?)
- morphological – a form has more than one morphological interpretations  
*Napětí vyvolalo zdražení.*
- lexical (homonymity) – a word has more than one meaning.  
*Did you see the bat?*  
*Where is the bank?*

All languages have expressions which have more than one possible interpretation.



Another example:

(43) Old men and women are exempt from the new tax.

Sometimes, world knowledge can help you to select the right interpretation:

- (44) a. I saw a policeman with a gun.  
 b. I saw a dog with a telescope.

Syntactically, these examples are ambiguous, however your knowledge of the world helps you to choose the most probable interpretation.

## 6.5 Characteristics of Phrase Structure Rules

A simple grammar:

- (45) a.  $S \rightarrow NP V (NP) (PP)$       d.  $N \rightarrow \{\text{dog, boy, ...}\}, P \rightarrow \{\text{on, in, ...}\}, \dots$   
 b.  $NP \rightarrow (Det) (A) N (PP)$   
 c.  $PP \rightarrow P NP$

This grammar describes a simple language (similar to English). It has several characteristics, which it shares with grammars of real languages:

- **Generativity:**

It does not *list* the sentences of the language, it describes the way how to *build* them. This is important, since languages contain infinite number of sentences.

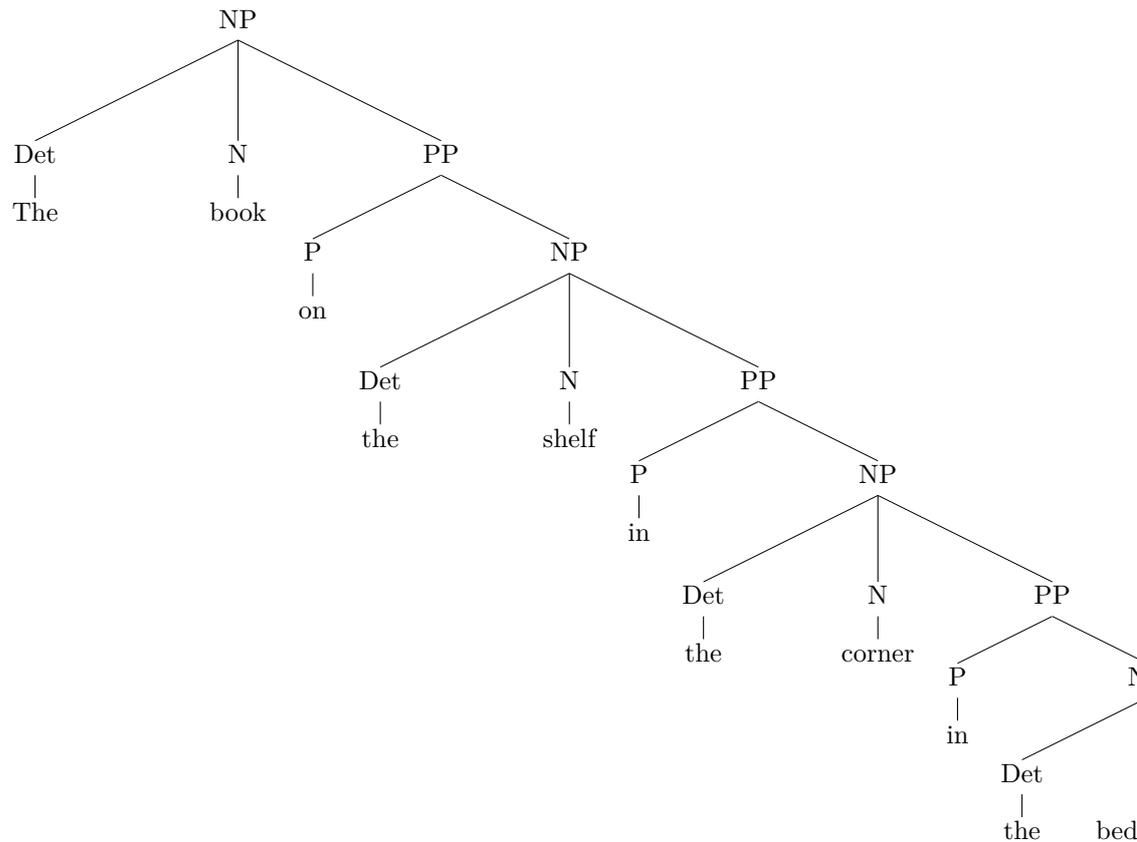
- **Ambiguity:**

Some sentences can be build in more than one way (starting with the S rule and ending with the words in the sentence) These sentences have more than one syntactic structure – they are syntactically ambiguous (for example the *telescope* sentence). This also shows that sentences are more than just simple strings of words.

- **Infinite Recursion:**

The grammar is recursive and thus allows to produce an infinite number of sentences using a finite (very small) number of rules.

(46)



### Questions

Some things to think about:

- Is it similarly easy to write a simple grammar of Czech?
- How would one capture agreement? (In English?, In Czech?)
- How is Phrase Structure Grammar different from Dependency Grammar (the type of syntax taught at Czech schools)

### 6.5.1 Formal grammars and Chomsky hierarchy (a refresher)

Formally, a grammar is:  $G = [N, T, P, S]$  where  $P \subseteq (N \cup T)^+ \rightarrow (N \cup T)^*$ ;  $S \in N$

Chomsky hierarchy of grammars:

	Languages	Complexity	Automaton	Type of rules
0	Recursively enumerable	Undecidable	Turing machine	any
1	Context-sensitive	Exponential	Linear-bounded	$\alpha A \beta \rightarrow \alpha \gamma \beta$
2	Context-free	Polynomial	pushdown	$A \rightarrow \gamma$
3	Regular	Linear	Finite state	$A \rightarrow a aB Ba$

## 6.6 Valency

### 6.6.1 Subject & Object/Complement

- (47) a. **John** snores loudly.  
 b. **John** wrote a long homework.  
 c. **A long homework** was written by John.

Roughly: **Subject** is the active participant in the active sentence.

- (48) a. **John** writes a letter.  
 b. However: **John** underwent torture (at the hands of the terrorists).

Note: We are defining a syntactic notion using semantics. It works in most of the cases, however not always (e.g. 2b)

Objects (complements) are the other participants.

Most of the active sentences can be transformed into passive sentences – an object becomes the subject, the subject becomes the by-PP or can be omitted.

- (49) a. **John** loves Mary. (active)  
 b. **Mary** is loved by John. (passive)  
 c. **Mary** is loved. (passive)

All English sentences (except imperative) have a **subject**. Many (but not all) have an object. In some languages (e.g. Spanish), the subject can be omitted if it is understandable from the context.

The pronoun in tag questions refers to the subject:

- (50) a. *The boy wrote a book, didn't he?*  
 b. *The girl wrote a book, didn't she?*  
 c. *John underwent torture at the hands of the terrorists, didn't he?*  
 d. *\*John underwent torture at the hands of the terrorists, didn't they?*

### 6.6.2 Transitive and intransitive verbs

- **Intransitive** verb – a verb with a subject and no objects: *sleep, snore*

(51) a. *John snores.*  
 b. \**John snores a book.*

- **Transitive** verb – a verb with a subject and an object: *buy, brush, write, catch*

(52) a. *John buys a candy.*  
 b. \**John buys.*

Some verbs are both intransitive and transitive: *dance* – *John dances* × *John dances samba*.

- **Ditransitive** verbs – a subclass of transitive verbs, take two objects (direct & indirect).

(53) a. *John give a book to his friend.*  
 b. *John gives his friend a nice book.*  
 c. *A nice book is given to Mary by John.*  
 d. *Mary is given a nice book by John.*

### 6.6.3 Form

Typical realization of subjects and objects:

- Noun phrase (NP)

(54) a. *The cat caught the mouse.*  
 b. *He goes to Chicago.*

- A sentence (S)

(55) a. *That I lost the tickets annoys me.*  
 b. *I know that this is true.*

### 6.6.4 How to know what is what?

#### Word order

English:

- declarative sentences – subject goes before the verb, object after it:  
*A small dog chases a big cat.*                      Subj V Obj
- interrogative sentences – subject after the auxiliary verb, object after the verb:  
*Does a small dog chase a big cat?*                      Aux Subj V Obj

### Case

**Case** – morphological marking of a word suggesting its syntactic function. Usually, words have a special suffix at the end indicating the case.

English: very sporadic, only some pronouns distinguish it:

- (56) a. **He** sees Mary.  
 b. Mary sees him.

If a language has cases, then this is the typical situation:

- **nominative** case is used for subjects (e.g. Latin *deus* God-subject)
- **accusative** case is used for (direct) objects (e.g. Latin *deum* God-object)

Some languages have more cases (Old English – 4, German – 4, Latin – 6, Finnish – 15) serving generally the same function as English prepositions.

Czech (the same word order, different cases, different meaning):

- (57) a. *Mal-á kočk-a vidí Honz-u.*                      Subj V Obj  
           Small-nom cat-nom sees John-acc                nom    acc  
           The small cat sees John.  
 b. *Mal-ou kočk-u vidí Honz-a.*                      Obj V Subj  
           Small-acc cat-acc sees John-nom                acc    nom  
           John sees the small cat.

The suffix *a* in *kočka* says it is nominative, therefore it is the subject. The suffix *u* in *Honzu* says it is accusative, therefore it is the object.

### Agreement

In many languages, subject and verb agree – they share certain morphological properties (number, person, gender, etc.)

English: Limited only to the 3<sup>rd</sup> person singular and everything else distinction:

- (58) a. *I write a letter.*  
 b. \**I writes a letter.*

Spanish:

	Singular		Plural	
1	<i>yo habl-o</i>	I speak	<i>nosotros habl-amos</i>	we speak
2	<i>tú habl-as</i>	you speak	<i>vosotros habl-áis</i>	you (guys) speak
3	<i>él habl-a</i>	he speak	<i>ellos habl-an</i>	they speak

Russian:

	Singular		Plural	
1	<i>ya vizh-u</i>	I see	<i>my vid-im</i>	we see
2	<i>ty vid-ish'</i>	you see	<i>vy vid-ite</i>	you (guys) see
3	<i>on vid-it</i>	he sees	<i>oni vid-yat</i>	they see
	<i>on vid-el</i>	he saw	<i>ona vid-ela</i>	she saw

Some languages (Bantu in Africa, etc.) have object-verb agreement.

### 6.6.5 Adjuncts

- (59) a. John eats [often].  
 b. John eats [loudly].  
 c. John eats [in the morning].  
 d. John eats [when he gets hungry].  
 e. John eats [in a restaurant] [on Sunday].

- Can combine with any verb (mostly)
- Can be repeated
- Have meaning on their own
- Usually expressed by adverbs, PPs

## 6.7 Syntax vs. Semantics

Compare:

- (60) a. *I ate a cake.*  
 b. *I entered a room.*

The sentences have the same syntactic structure – both *a cake* and *a room* are (direct) objects, but the cake disappeared after I ate it, while the room did not after I entered it.

Compare:

- (61) a. *I took my friend to Tesco.*  
 b. *I took D5 to Tesco.*

Both *my friend* and *D5* are (direct) objects, but ...

## 6.8 Valency

### 6.8.1 Grammatical Roles

- **subject** (podmět). Roughly: **Subject** is the active participant in the active sentence.

(62) a. *John writes a letter.*

b. However: *John underwent torture (at the hands of the terrorists).*

Note: We are defining a syntactic notion using semantics. It works in most of the cases, however not always.

- **object** (předmět)
- **complement** – various meanings:
  - object – used here
  - object or subject, i.e., non-adjunct
  - (doplňěk) modifier of both the verb and an object/subject.

(63) The task seems hard. – *hard* is subject complement

(64) I consider the task hard. – *hard* is object complement

- **subcategorization, (surface-)valency (frame)** – list of complements of a word (usually a verb)
- **adjunct** (přísloušné určeni)

(65) a. John eats [often].

b. John eats [loudly].

c. John eats [in the morning].

d. John eats [when he gets hungry].

e. John eats [in a restaurant] [on Sunday].

In some theories, these notions are primitive (undefined), in some theories they are defined in terms of other notions. For example, subject can be defined structurally as the NP in  $S \rightarrow NP VP$ .

#### Adjunct versus Complement

complements	adjuncts
combine with particular verbs	can combine with any verb (mostly)
cannot be repeated	can be repeated
their meaning is determined by the verb	have meaning on their own
usually NPs, PPs	usually AdvPs, PPs

#### Transitive and intransitive verbs

- **Intransitive** verb – a verb with a subject and no objects: *sleep, snore*

(66) a. *John snores.*

b. \**John snores a book.*

- **Transitive** verb – a verb with a subject and an object: *buy, brush, write, catch*

- (67) a. **John** *buys* a candy.  
 b. \***John** *buys*.

Some verbs are both intransitive and transitive: *dance* – *John dances* × *John dances samba*.

- **Ditransitive** verbs – a subclass of transitive verbs, take two objects (direct & indirect).

- (68) a. **John** *give* a book to his friend.  
 b. **John** *gives* his friend a nice book.  
 c. **A nice book** *is given* to Mary by John.  
 d. **Mary** *is given* a nice book by John.

### How to know what is what?

**Word order** English:

- declarative sentences – subject goes before the verb, object after it:  
**A small dog** *chases* a big cat.                      Subj V Obj
- interrogative sentences – subject after the auxiliary verb, object after the verb:  
*Does* **a small dog** *chase* a big cat?                      Aux Subj V Obj

**Case** **Case** – morphological marking of a word suggesting its syntactic function.

English: very sporadic, only some pronouns distinguish it:

- (69) a. **He** *sees* *Mary*.  
 b. *Mary* *sees* him.

If a language has cases, then this is the typical situation:

- **nominative** case is used for subjects (e.g. Latin *deus* God-subject)
- **accusative** case is used for (direct) objects (e.g. Latin *deum* God-object)

Some languages have more cases (Old English – 4, German – 4, Latin – 6, Finnish – 15) serving generally the same function as English prepositions.

Czech (the same word order, different cases, different meaning):

- (70) a. *Mal-á*    *kočk-a*    *vidí* *Honz-u*.                      Subj V Obj  
           Small-nom cat-nom sees John-acc                      nom    acc  
           The small cat sees John.
- b. *Mal-ou*    *kočk-u*    *vidí* *Honz-a*.                      Obj V Subj  
           Small-acc cat-acc sees John-nom                      acc    nom  
           John sees the small cat.

The suffix *a* in *kočka* says it is nominative, therefore it is the subject. The suffix *u* in *Honzu* says it is accusative, therefore it is the object.

**Agreement** In many languages, subject and verb agree – they share certain morphological properties (number, person, gender, etc.)

English: Limited only to the 3rd person singular and everything else distinction:

- (71) a. *I write a letter.*  
 b. \**I writes a letter.*

Spanish:

	Singular		Plural	
1	<i>yo habl-o</i>	I speak	<i>nosotros habl-amos</i>	we speak
2	<i>tú habl-as</i>	you speak	<i>vosotros habl-áis</i>	you (guys) speak
3	<i>él habl-a</i>	he speak	<i>ellos habl-an</i>	they speak

Russian:

	Singular		Plural	
1	<i>ya vizh-u</i>	I see	<i>my vid-im</i>	we see
2	<i>ty vid-ish'</i>	you see	<i>vy vid-ite</i>	you (guys) see
3	<i>on vid-it</i>	he sees	<i>oni vid-yat</i>	they see
	<i>on vid-el</i>	he saw	<i>ona vid-ela</i>	she saw

Some languages (Bantu in Africa, etc.) have object-verb agreement.

## 6.8.2 Deep Valency

The concepts of this section are on the interface of syntax and semantics. Some theories are grounded more in syntax, some more in semantics.

### Syntax vs. Semantics

Compare:

- (72) a. I ate a cake.  
 b. I entered a room.

The sentences have the same syntactic structure – both *a cake* and *a room* are (direct) objects, but the cake disappeared after I ate it, while the room did not after I entered it.

Compare:

- (73) a. I took my friend to Brno.  
 b. I took D1 to Brno.

Both *my friend* and *D1* (the freeway) are direct objects, but ...

Note: Below, we focus on verbs only.

### The Mainstream U.S. view of Deep Valency

Sources: Charles Fillmore (1968, The Case for Case), David Dowty (1979, Word meaning and Montague grammar), Ray Jackendoff (1983, Semantics and cognition), etc.

- (74) a. The janitor opens the door with a key. (key – instrument)  
 b. The key opens the door. (key – instrument)

**Thematic role** – semantic counterpart of grammatical function:

agent	one who deliberately does the action
cause	mindlessly performs the action
experiencer	has sensory or mental experience
patient	thing that the action happens to
theme	thing or being that is in a state/location
source	origin of a change in location/possession
goal/recipient	endpoint of a change in location/possession
instrument	the means of accomplishing the action
etc	

The exact set is theory dependent. For example, some theories do not distinguish between patient and theme. Some theories allow a single constituent to have multiple roles, some don't.

- (75) a. John ate the Cheerios with a spoon.  
 AGENT PATIENT INSTRUMENT
- b. John fell.  
 THEME
- c. John is polite.  
 THEME
- d. John heard the explosion.  
 EXPERIENCER THEME
- e. John gave a book Mary.  
 AGENT/SOURCE THEME GOAL
- f. John caught the flu.  
 GOAL THEME

Important: The roles reflect how the speaker structures the reality, not necessarily the reality itself.

- (76) a. John broke the window. (John – AGENT)  
 b. The hammer broke the window. (the hammer – INSTRUMENT)  
 c. The storm broke the window. (the storm – CAUSER)

Going too far? Is it still linguistics? What if John was coerced/hypnotised?  
 However, Agents and Causers sometimes behave grammatically differently:

- (77) a. The pressure/explosion/Will's banging broke the window.  
 b. John broke the window.  
 c. The window broke from the pressure/explosion/Will's banging.  
 d. \*The window broke from John.

### Deep Valency a la Prague

Less semantic, more syntactic.

Main reference: Panevová (1980): *Formy a funkce ve výstavbě české věty*.

Sources: Tesnière, Pauliny, Fillmore.

Modifiers (doplnění) are classified as:

- **actant, inner participant** (aktant, vnitřní participant)
  - Combines with specific verbs; the verbs must be listed
  - Cannot be repeated

There are five actants: Actor, Addressee, Origo, Effect

- **free modifier** (volné určení)
  - Can combine with any verb
  - Can be repeated

About 40 types: when, from when, how long, where, where from, how, instrument, comparison, reason, ...

**Valency frame** of a word (usually a verb) – actants and obligatory free modifiers combining with that verb.

- (78) a. Matka dala dítěti jablko.  
 actor patient addressee  
 mother gave to-child apple  
 ‘The mother gave her child an apple.’
- b. V Praze se zasedání konají na filozofické fakultě.  
 place actor place  
 in Prague refl meetings take place at the Faculty of Arts.  
 In Prague, the meetings take place at the Faculty of Arts.
- c. \*Jan potkal.  
 actor (Patient)  
 lit: ‘Jan met.’
- d. \*Jan se chová.  
 actor (Manner/Comparison)  
 lit: ‘Jan behaves.’ (a different meaning of behave)

### Dialog text

- (79) A: My friends arrived.  
 B: Where? Where from?  
 A: #No idea. No idea

Inner participants might be optional

- (80) a. Matka předělala (dětem) loutku z kašpárka na čerta.  
 Act (Adr) Pat Orig Eff  
 mother transformed for-children puppet from-joker to-devil  
 ‘The mother transformed the puppet from the Joker to the Devil (for her children).’
- b. Šéf vyměnil sekretářku.

**Role shifting**

1. Act is occupied, if one or more actants.
2. Pat is occupied, if two or more actants.
3. Eff/Adr/Orig are semantically assigned if 3 or more actants.

$$\text{Act} \leftarrow \text{Pat} \leftarrow \begin{cases} \text{Adr} \\ \text{Eff} \\ \text{Orig} \end{cases}$$

Formally decided:

- (81) a. Kniha vyšla. Komín kouří.  
 b. oslovit někoho  
 c. vykopat jámu, vyspět v muže (z jinocha)  
 d. vyvinout se z něčeho
- (82) a. John broke the vase. – John always actor  
 b. read/write/copy/damage/burn a book – book always patient

Keep in mind that we are talking about (deep) syntax, not semantics.

**6.8.3 Accusative and Ergative languages**

The argument of an intransitive verb is marked (morphologically and/or syntactically) in the same ways as:

- subject of transitive verbs – accusative languages (Czech, English, ...)
- object of transitive verbs – ergative languages (Basque, Eskimo-Aleut lgs, Tibetan, ...)

- (83) Czech – accusative lg
- a. Honz-a přijel.  
 John-nom has arrived  
 ‘John has arrived’
- b. Honz-a viděl Mari-i.  
 John-nom saw Marie-acc  
 ‘John saw Mary.’
- (84) Basque – ergative lg
- a. Gizon-a etorri da.  
 man-**abs** has arrived  
 ‘The man has arrived’
- b. Gizon-ak mutil-a ikusi du.  
 man-erg boy-**abs** saw  
 ‘The man saw the boy.’

**Traces of ergativity in English**

Ergative verbs:

- (85) a. John broke the vase.  
b. The vase broke.  
c. John opened the door.  
d. The door opened.



# Chapter 7

## Semantics

What to remember and understand:

Semantics, Word meanings (pros and cons for different candidates), Synonym/Antonym, Subordinate/Superordinate term, Meronym/Holonym, Homonym. Compositionality.

### 7.1 What is Semantics

Semantics is the part of linguistics that studies meaning in language:

- the meanings of words
- how word meanings combine to give the meaning of a sentence

It is very close to pragmatics and the border is often uncertain. Somebody even considers pragmatics part of semantics.

Generally:

- Semantics deals with literal meaning.
- Pragmatics deals with the intended meaning, with the usage of language, with language in context, etc.

### 7.2 The Meaning of Words

Every word has some meaning (some only in context).

What is a meaning of a word? Consider some candidates:

- **Dictionary definition**

transmission:

- 1a. The act or process of transmitting.
- 1b. The fact of being transmitted.
2. Something, such as a message, that is transmitted.
3. An automotive assembly of gears and associated parts by which power is transmitted from the engine to a driving axle. Also called gearbox.
4. The sending of a signal, picture, or other information from a transmitter.

**But:**

- to understand meaning of one word, you have to know the meanings of other words.
- definitions often circular

• **Mental image****But:**

- Different people have different mental images of the same words.
- Usually it is an image of some prototype.
- It is hard to find mental images for some words (*the, he, aspect*).

• **Reference** – The meaning of a word is the actual thing in the real world. Works for *chair, house*, etc.**But:**

- *tooth fairy, unicorn*, abstract nouns (*history*), adjectives (*big*), etc.
- different expressions with the same reference  
*Václav Klaus* and *the current Czech president* refer to the same object, so in this view, they should have the same meaning. But:
  1. If they had the same meanings, one should be replaceable by the other.  
*Peter does not know that Václav Klaus is the current Czech president.* – ok  
*Peter does not know that Václav Klaus is Václav Klaus.* – strange
  2. Intuitively, these expressions have different meaning.

• **Parallel worlds** – works, but quite complicated and not very intuitive. One world says how things really are, the other worlds say how things could be. Meaning of a word is the collection of its references in all worlds. In some worlds, *Václav Klaus* is the current Czech president, in some he is not. Therefore in this theory, *Václav Klaus* and *the current Czech president* have different meanings, as we want.

## 7.2.1 Relations between words

- **synonym** – a word having the same or nearly the same meaning as another word.  
e.g. *gearbox – transmission, choice – selection, complex – complicated;*  
*zde – tady* ‘here’, *ekonomika – hospodářství* ‘economy’ – *pěkný, kromě – mimo* ‘except’  
synonym: From Greek: *syn* (same, together) + *onoma* (name).  
Cf. *synergy, symbiosis, symphony; name, nominate, anonymous, pseudonym*
- **antonym** – a word having the opposite meaning as another word.  
e.g. *happy* vs. *unhappy, heavy* vs. *light, long* vs. *short*  
antonym: From Greek: *anti* (opposite, against) + *nomen* (name).  
Cf. *antibiotic, Antichrist, antiviral*

- **subordinate word** (hyponym) – a word whose meaning is included in the meaning of another word.

e.g. *red* < *color* (*red* is subordinate to *color*), *pony* < *horse* < *animal*, *tulip* < *flower*

This is not whole-part relationship, so *page* is *not* subordinate term of *book*.

hyponym: From Greek: *hyppo* (under) + *nomen* (name).

Cf. *hypothermia*, *hypoallergenic*, *hypoglycemia*, *hypothesis*, *hypothalamus*

- **superordinate word** (hypernym) – a word whose meaning includes meaning of another word. If X is superordinate to Y then Y is subordinate to X.

e.g. *color* > *red*, *black*, *green*, ...; *animal* > *horse*, *tiger*, ...

hypernym: From Greek: *hyper* (over, beyond) + *nomen* (name).

Cf. *hyperactive*, *hypertension*, *hypertext*, *hyperbole*

- **meronym** – denotes part/member of something.

e.g. *page* is a meronym of *book*, *wheel*, *engine*, ... are meronyms of *car*

meronym: From Greek: *meros* (part) + *nomen* (name).

Cf. *polymer*, *merit*

- **holonym** – opposite of meronym

holonym: From Greek: *holon* (whole) + *nomen* (name).

Cf. *hologram*, *holistic*, *catholic*, *holocaust*

- **homonym** – a word having a different meaning, but the same form as another word.

e.g. *bear* (verb) vs. *bear* (noun, animal), *bank* (institution) vs. *bank* (river bank)

*kolej* ‘rail’ vs. *kolej* ‘dormitory’, *jí* ‘eats’ vs. *jí* ‘she<sub>dat</sub>’, *ženou* ‘woman<sub>inst</sub>’ vs. *ženou* ‘hurry/herd<sub>3pl</sub>’

It is possible to distinguish homophones (same sound) and homographs (same spelling)

homonym: From Greek: *homos* (same) + *nomen* (name). *homogeneous*, *homosexual*. Not related to *homo sapiens*.

Homonymy and synonymy – complementary notions

- synonyms – the same meaning, different forms.
- homonyms – the same form, different meanings.

Semantic network (Wordnet) – a special kind of dictionary – words are organized into a network, where different relations connect synonyms, different antonyms, etc. Useful for translators or computational linguistics.

## 7.3 The Meaning of Sentences

Knowing the meaning of words is not enough to understand the meaning of a sentence:

- Two sentences with the same words can mean something different:

(86) a. *A tourist eats a tiger.*

b. *A tiger eats a tourist.*

- Two sentences with some of the words different can mean the same:

- (87) a. *A tiger eats a tourist.*  
 b. *A tourist is eaten by a tiger.*

You must know that in English:

- In active sentences, the one who does something precedes the verb.

- (88) *A tiger eats a tourist.*  
 eater            eaten

- In passive sentences, the one who does something follows the verb and has a preposition *by*.

- (89) *A tourist is eaten by a tiger.*  
 eaten            eater

There are many rules like that that help you to get the meaning of the sentence out of the meaning of the words. You know them for your native language unconsciously, but you have to learn them for foreign languages.

**The Principle of Compositionality:** The meaning of an expression is a function of the meanings of its parts and of the way they are syntactically combined.

Note: Each of the key terms in the principle of compositionality is a theory-dependent term.

**Entailment:**  $\phi \models \psi$  iff (if  $\phi$  is true then  $\psi$  must be true)

Basic ideas of Montague grammar

1. Syntactic categories correspond to semantic types. E.g. S to Proposition, N to property of entities, V to properties of events
2. Lexicon specifies the semantics and syntax of basic expressions.
3. Syntactic rules correspond to semantic rules (e.g. function application may correspond to function application)

## 7.4 Other things about semantics

### 7.4.1 Things modern semantics studies

- tense and aspect – different languages use different tenses/aspects to express the same thing.
- ellipsis – *John read the paper before Bill did.*
- determiners – different languages use determiners in different way
- relation between semantics, pragmatics and intonation
- etc.

It is closely related to formal logic.

### 7.4.2 Examples of Applications

Where semantics can be helpful:

- In the creation of dictionaries and thesauruses – both necessary for improving knowledge of your native language or in learning foreign languages.
- In teaching computers to process language:
  - Web and library searches – When trying to repair a car transmission we do not want web sites about information transmission.
  - Machine translation – We probably want to translate *bank* differently in *investment bank* and in *West Bank*.
  - Automatic creation of manuals, etc. – You describe the semantics of the text, and then a program creates corresponding texts in several languages.

### 7.4.3 If you want to know more

- English Wordnet at Princeton: <http://www.cogsci.princeton.edu/~wn/>  
Search – <http://www.cogsci.princeton.edu/cgi-bin/webwn> – First search for some word (*vegetable*) then search for synonyms, antonyms, hypernyms, etc.
- The Global WordNet Association: <http://www.globalwordnet.org>  
Wordnets in various languages.
- B. Partee's course in formal semantics: [http://people.umass.edu/partee/MGU\\_2007/MGU07\\_formal\\_semantics.htm](http://people.umass.edu/partee/MGU_2007/MGU07_formal_semantics.htm)



## Chapter 8

# Pragmatics



## Chapter 9

# Linguistic Theories



## Chapter 10

# An example formal theory of language



# Chapter 11

## Historical Linguistics: Language Change

### 11.1 What is Historical Linguistics?

Historical linguistics studies:

- how languages *change* over time
- how languages are *related*.

Main theses:

- All languages change over time. The change is relatively slow.
- Languages do not change randomly.
- There are many patterns and generalizations we find out when examining the histories of various languages.

Note: Because devices for recording sounds have only been around for about a century, the vast majority of data used for historical linguistics is textual. This is only a matter of circumstance, however – as discussed before, the *spoken* language is primary; we only analyze the *written* language if we have no other options.

### 11.2 Two Approaches to the Study of Language

- **synchronic** – Study of language at one point in time (usually “now” – the modern form of a language).

synchronic: From Greek: *syn* (same, together) + *chronos* (time)

Cf. synchronize, synonym, symphony, synergy; chronicle, chronology, chronic

- **diachronic** – studies historic development of a language or languages. (compares the language with itself at different stages of its development)

diachronic: From Greek: *dia* (thru, across) + *chronos* (time)

Cf. diagonal, diarrhea (thru + flow), diagram, diadem (bound across), dialect (across + speak), dialog (across + speak), diameter (across + measure), diaspora

### 11.3 Development of English

1. Old English (OE, 450-1100)
2. Middle English (MidE, 1100-1450)
3. Early Modern English (1450-1700)
4. Modern English (ModE, 1700-present)

Major changes

1. Old English to Middle English
  - (a) Loss of /x/ (usually written as *gh*, German preserved it)
 

English:	<i>right</i> [rait]	<i>night</i> [nait]	<i>neighbor</i> [neibər]
German:	<i>rechts</i> [rext <sup>s</sup> ]	<i>nacht</i> [naxt]	<i>nachbar</i> [naxbar]
  - (b) Some allophones became distinct phonemes  
[f]/[v], [ð]/[θ], etc.
  - (c) Loss of many final vowels > loss of case endings
  - (d) Word order became more rigid
  - (e) England under Norman (French) rule > about 10K French borrowings
  - (f) etc.
2. Middle English to Early Modern English
  - (a) The Great vowel shift – only [i] and [e] unaffected.  
One vowel “pushed” another to a different place in the vowel chart.  
e.g. ɔ > o > u > au > ɔ  
[ur] > [aʊr] *our*, [namə] > [neim] *name*, [wid] > [waɪd] *wide*, etc.
  - (b) Simplification of some initial consonant clusters:  
[kn] > [n] (*know*, *knee*, *knight*), [hr] > [r] (*hring* > *ring*), [wr] > [r] (*write*, *wrist*)

### 11.4 Development of Czech

1. Old Slavonic – no direct sources, reconstructed language
2. Old Church Slavonic (staroslověňština) – Literary language based on the Thessaloniki Slavic dialect, created Saints Cyril and Methodius in 800's.
3. Proto-Czech (1000-1150)
  - (a) V j V > V: (*dobraja* > *dobrá*)
  - (b) yers (ъ and ѣ; reduced vowels): yer > 0 (odd yers from the end) / e (even yers)  
*okъno* > *okno* but *okъnъ* > *oken*
  - (c) nasal vowels disappeared
4. Old Czech (1150-1250)
 

No continuous texts, mostly single words or glosses in Latin texts.

In comparison with modern Czech, Old Czech has: a full system of palatal and hard consonants, dual number, simple past tenses (aorist, imperfect), etc.

- (a) [a/æ] > [jɛ] (ě) after palatal/palatalized consonants. This led to differentiation between hard and soft paradigms
  - (b) *g* > *h* (*gora* > *hora* ‘mountain’)
5. 1300-1400’s
- First longer written texts
- (a) Depalatalization – Old Czech had a full system of palatal and hard consonants, most of this is lost
  - (b) Difference between *i* and *y* neutralized
  - (c) Beginnings of: *ý* > *ej* (*být* > *bejt* ‘to be’) and *ú* > *ou* (*súd* > *soud* ‘court’)
  - (d) Beginnings of protetic *v* before initial *o*: *vo**kno* ‘window’
  - (e) Simple past tenses disappeared
  - (f) Standardization of orthography (probably by Jan Hus), diacritics replaces digraphs. E.g., *rz*, *rrz*, *rs*, *rzs*, *rzss*, ... > *dotr*
6. Humanistic period
- Kralická bible published, several Czech grammars published
- (a) Dual number disappears except for certain nouns and agreeing attributes (*s dlouhýma rukama* ‘with long arms’, *dvě stě* ‘200’, ...)
7. Baroque period
- Czech is replaced by German in many situations.
- (a) *é* > *í*: *řeci* > *řici* ‘say’, *dobré mléko* > *dobrý mlíko* ‘good milk’
  - (b) dual endings used instrumental plural (*-ama*)
8. The National Renaissance
- Creation/resurrection of Literary Czech on the basis of humanistic Czech ignoring 200 years of development and resulting in split between Literary and Common Czech.

## 11.5 Kinds of Language Change

Languages undergo changes at all linguistic levels: phonetic, phonemic, morphological, syntactic, semantic and pragmatic

1. Phonetic:
 

Old English had [y] (as [i], but rounded) and [x], Modern English has none of them  
 [y] was replaced by [i] or [ai]: [pyt] > [pɪt] *pít*, [my:s] > [mis] > [mais] *míce*.  
 Reduction of yers in Proto-Czech: *дньсь* > *dnes* ‘today’
2. Phonemic:
 

In Old English [v] and [f] were allophones (variants) of one phoneme. In Modern English they are two distinct phonemes.
3. Morphological:
 

In OE, nouns had case endings (nominative, genitive, dative, accusative)

## 4. Syntactics:

In OE, all questions could be formed by inverting the subject and the verb.

In ModE, inversion possible only with auxiliaries, other verbs use *do*.

Old Czech had several simple past tenses, Modern Czech has only analytical preterite, some forms of aorist (one of the simple past tenses) are used as conditional auxiliary (*bych*)

## 5. Semantic &amp; lexical:

In OE, *girl* referred to young men and women.

Czech: *pivo* any drink > beer

- (90) Shakespeare's time: *What thinkest thou?*  
 Modern English: *What do you think?*

- syntactic change – no auxiliary *do*
- morphological change – verb endings
- lexical change – *thou* is now obsolete.

## 11.6 Why Do Languages Change?

Languages change for a variety of reasons, for example:

- **Economy:** Speakers tend to use as little energy as possible to reach the goals of communication.
- **Analogy:** One part of the system (lexicon, morphology, ...) or even a single word or rule is modified to be more like other parts of the system.
- **Change of context** (society, culture, place)
- **Language split** – A language may split into two or more languages if the speakers become separated into two or more groups with little or no contact.  
 Latin > French, Provençal, Spanish, Catalan, Portuguese, Romanian.  
 English > British, American, Indian, South African, Australian, ...
- **Borrowing** between language in contact (because of trade, etc.)  
 Mostly vocabulary. Less frequently in phonetics, phonology; sometimes also syntax, morphology and semantics.

English borrowed words from many languages (often via other languages), e.g., French (*design, court, table*), Latin (*deficit, sponsor*), Scandinavian lgs. (*they, law*), German (*kindergarten, noodle*), Spanish (*canyon, tornado*), Italian (*isolate*), Greek (*comedy, theater*), Native American lgs. (*chipmunk, kayak*), South Asian (*sandal*), Dutch (*cruise*), Arabic (*algorithm, giraffe, alcohol*), Chinese (*tea*), Russian (*czar*), Czech (*robot, howitzer*).

Czech borrowings: English (*gól, software*), German (*flaška, láhev*), Russian (*samovar*), Latin (*košile, norma, cirkus, minuta*), French (*toaleta*), Spanish (*armáda*), Italian (*banka,* ), Greek (*kytara, symbol*), Turkish (*jogurt*), Arabic (*algebra, alkohol*), Chinese (*čaj 'tea'*), ....

## 11.7 Attitudes Towards Language Change

People complained about deterioration of language forever.

There were many attempts to resist language change, defending it against:

- “invasion of barbaric” terms from other languages
- preventing “vulgar speech from corrupting” the language as a whole.

Many countries have/had language academies – institutions organized to attempt to regulate, stop, or even reverse language change.

Academies & government

- can force newspapers and book publishers to conform to specific guidelines,
- but they cannot control how people speak, and that’s where the language change originates.

You cannot prevent language change with any amount of force – it is an inherent part of every living human language.

## 11.8 Language families

We can classify languages by their origin into a tree similar to family trees.

Most of the European and some Asian languages evolved from a common ancestor called Proto-Indo-European.

Some of the subfamilies of **Indo-European** languages.

- Slavic  
Czech, Russian, Polish, Croatian, Bulgarian, etc.
- Italic languages – Latin and its descendants (Romance languages)  
Spanish, Portuguese, French, Italian, Romanian, etc.
- Germanic languages  
English, German, Dutch, Afrikaans, Yiddish, Norwegian, Swedish, Danish, etc.
- Celtic  
Irish, Scots Gaelic, Welsh
- Indo-Iranian  
Sanskrit, Hindi, Urdu, Romany (Gypsy), Pashto
- Hellenic  
Greek

Except Indo-European there are many other language families, for example:

- Sino-Tibetan languages  
Mandarin (Chinese), Cantonese (Chinese), Tibetan, Thai

- Uralic languages  
Hungarian, Finnish
- Afro-asiatic  
Semitic languages (Arabic, Hebrew, Maltese), Somali, Egyptian (Ancient Egypt), etc.

There is about 7000 languages, originating probably in a small number of proto-languages.

For detailed classifications:

- <http://www.ethnologue.com>
- <http://www.krysstal.com/langfams.html>

## 11.9 Sound Change

**A Sound Change** is a phonological process which has been accepted by all speakers of a language (or dialect).

Sound changes tend to spread from speaker to speaker gradually in a wave-like pattern until they are uniformly used by all speakers in a linguistic community. They do not spontaneously occur everywhere in a language.

### When:

1. **Conditioned** – the change happens only in certain phonetic environment  
e.g. all [s] in front of a vowel change into [ʃ]
2. **Unconditioned** – the change happens regardless of the phonetic environment  
e.g. all [s] change into [ʃ]

### What:

1. **Assimilation:** Two sounds become more like each other when they are near or touching.  
e.g. [wulfas] > [wulvas] ([f] is voiceless, both [l] & [v] are voiced)
2. **Dissimilation:** Two sounds become less like each other when near or touching.  
e.g. [fifθ] > [fift] *fifth* (both [f] & [θ] are fricatives, [t] is a stop)
3. **Deletion:** Sounds are deleted in certain environments.  
e.g. MidE [knɪxt] > ModE [nɑt] *knight* ([k] and [x] deleted)
4. **Insertion:** Sounds are inserted in certain environments.  
e.g. [æθlit] > [æθəlɪt] *athlete*
5. **Monophthongization:** Diphthongs become monophthongs.  
e.g. MidE [rɪuələ] > ModE [rul] *rule*
6. **Diphthongization:** Monophthongs become diphthongs  
e.g. MidE [u] > ModE [au]: [hus] > [haus] *house*, [maus] *mouse*

7. **Metathesis:** The order of sounds change.

e.g. *ask* > *aks*

8. **Raising/Lowering:** The position of the tongue becomes either higher or lower when producing certain sounds.

e.g. [metə] > [mit] *meat*

9. **Backing/Fronting:** The position of the tongue becomes either more forward or more back.

e.g. [pɑθ] > [pæθ] *path*

Note: Always look at pronunciation not spelling:

- Pronunciation can change without spelling changing  
e.g. English *name* used to be pronounced [na:mə] but spelling did not change.
- Spelling can change without pronunciation changing  
Turkish switched it's whole alphabet in 1928 without changing pronunciation  
Russian modified it's spelling in 1917 without changing the pronunciation

## 11.10 Comparative reconstruction

How do we know about supposed mother languages if there are usually no speakers of these languages and frequently no texts?

We **reconstruct** vanished languages by **comparing** their descendant languages.



## Chapter 12

# Psycholinguistics: Language and Mind