Roviny analýzy jazyka. Fonetika

Aleš Horák

E-mail: `hales@fi.muni.cz`
`http://nlp.fi.muni.cz/poc_lingv/`

Obsah:

- Roviny analýzy jazyka
- Fonetika a fonologie

Struktura jazyka

Struktura jazyka zahrnuje informace o:

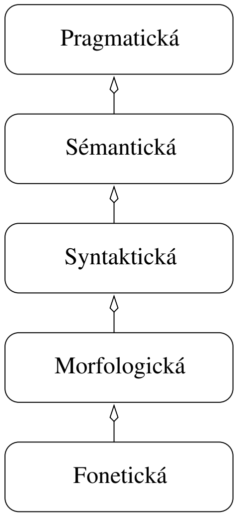
- co jsou **slova** (slovní tvary a jejich složky – morfémy)
- jak se slova (větné složky) kombinují do **vět**
- co slova označují, jaké jsou jejich **lexikální významy**
- jak se **význam věty** skládá z významů slov a slovních spojení (větných složek)

zpracování jazyka dále potřebuje:

- obecnou (encyklopedickou) **znalost světa** (ontologie)
- **inferenční mechanismus**
- znalost **komunikační situace**

Roviny analýzy jazyka

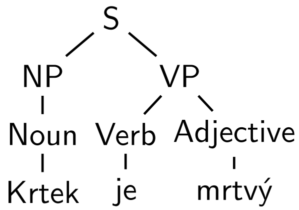
znalosti struktury jazyka jsou propojeny **hierarchicky**



jazykové **roviny**:

- **fonetická**
- **morfologická**
- **syntaktická**
- **sémantická**
- **pragmatická**
- kontextová
- znalost základní ontologie
- jazykové metaznalosti

Roviny analýzy jazyka – příklad

rovina analýzy	příklad
pragmatická	$\neg Na\_živu(Krtek_1, T_3)$ $Unavený(Krtek_1, T_3)$
sémantická	$\neg Na\_živu(Krtek, Teď)$ $Unavený(Krtek, Teď)$
syntaktická	
morfologická	Krtek–Noun1MS, je–Verb3MP, mrtvý–Adjective1MS
fonetická	[k r t e k j e m r t v i:]
povrchová	“Krtek je mrtvý.”

Roviny analýzy jazyka – pokrač.

- fonetická – postihuje vztahy mezi zvuky používanými v (mluveném) jazyce, jejich skládání do slabik a slov
  foném – nejmenší jednotka jazyka, která může **odlišit** význam nadřazených jednotek

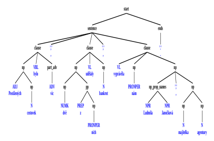
  kosit/*nosit* fonémy *k* a *n* odlišují dvě slova

  často odpovídají *znakům* → vždy ale označují *zvuky*
- morfologická – interní struktura slov, skládání slov z menších jednotek
  morfém – nejmenší jednotka, která může **nést** význam

pří-lež-it-	pří	– prefix (*blízko*)
-ost-n-ými:	lež	– lexikální kořen (*ležet*)
	it	– adjektivní derivační sufix (*ten, který*)
	ost	– substantivní derivační sufix (*ta skutečnost, že*)
	n	– adjektivní derivační sufix (*charakteristický pro*)
	ými	– gramatický afix (*instrumentál plurálu*)

Roviny analýzy jazyka – pokrač.

- syntaktická – struktura větných frází
  popisuje, jak vypadá **gramaticky správná věta**, většinou pomocí **pravidel gramatiky**
  syntaktický analyzátor – nástroj, který analyzuje vstup na základě gramatiky
  na výstup dává různé info, např. derivační stromy



- sémantická – význam výrazů přirozeného jazyka a jejich kombinací
  hodně závisí na zvolené **sémantické reprezentaci**
  logická analýza věty – strukturní část sémantické analýzy
- pragmatická – zkoumá vztah mezi výrazy přirozeného jazyka a **kontextem**
  často se do ní řadí znalost **komunikační situace**, **základní ontologie** a **jazykových metaznalostí**

Fonetika a fonologie

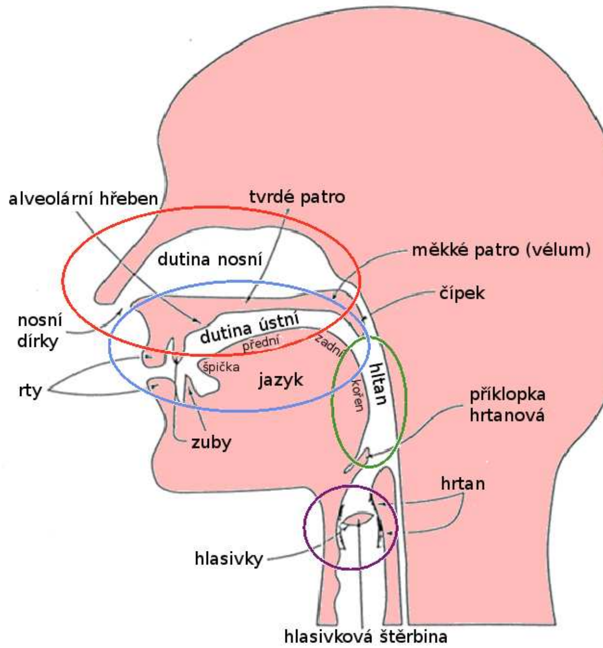
Fonetika:

- studuje **produkci**, **přenos** a **příjem** jazykových zvuků
- má klíčový význam např. pro oblast automatického **rozpoznávání** a **syntézy řeči**
- není tradičně chápána jako součást gramatiky jazyka

Fonologie:

- **fonologický systém** jazykových zvuků v *určitém jazyce*
- pracuje s **gramatikou** řečových zvuků
- pomocí gramatických pravidel popisuje historické změny i současné alternace

Kde vznikají jazykové zvuky?



Členění řečového proudu

řečový proud:

- nejsou mezery mezi slovy
- nejsou žádné izolované zvuky
- přesto všechny jazyky pracují s lingvistickými jednotkami jako separátními

orofón – fráze, které zní stejně/podobně, ale mají jiný obsah

It's not easy to recognize speech.
It's not easy to wreck a nice beach.

Fonetické jednotky

- foném (*phoneme*)
  - základní jednotka **zvukového systému** jazyka
  - foném je *abstraktní věc*, konkretizuje se pomocí *fónů* (viz dále)
  - např. v **češtině** – 37 fonémů:

    `a, a:, b, ts, tS, d, d', dz, dZ, e, e:, f, g, h\, x, i, i:, j, k, l, m, n, n', o, o:, p, r, r', s, S, t, t', u, u:, v, z, Z`
- fón (*phone*)
  - **řečový zvuk** z hlediska jeho **fyzikálních charakteristik** (zvuková vlna určitého tvaru)
  - bez zařazení k zvukovému systému jazyka
  - jeden **foném** odpovídá **množině** fónů
  - **alofón** určitého fonému = jeden z množiny fónů tohoto fonému např. `nosit`, `banka`

Fonetická transkripce

- jeden z nejpoužívanějších **nástrojů fonetiky**
- **převod** řečového proudu do oddělených, lingvisticky významných **symbolických jednotek**
- používá se standardních **fonetických abeced** (viz dále)
- **široká** × **úzká** (broad/narrow) transkripce = převod *do fonémů/fónů*
- důvody pro tento převod:
  - nedostatečnost písmenného zápisu
    - jedno písmeno → různý zvuk `vypít [v] / vpustit [f]`
    - jeden zvuk → různá písmena `chovat [x] / shánět [x]`
  - mezijazykové variace v písmenném zápisu
    - 'k' → 'c' v latinském `canis`, 'ch' v italském `Chianti`
    - 'č' → 'ch' v anglickém `cheat`, 'ci' v italském `ciao`
  - jeden foném může být zaznamenán více písmeny

    např. 'f': → 'f' v českém `fyzika`
    → 'gh' v anglickém `laugh`
    → 'ph' v řeckém `philosophia`

Fonetické abecedy IPA a SAMPA

IPA:

- *International Phonetic Alphabet*
- vznikla v roce 1886 v Paříži, od té doby mnoho revizí (poslední 1996)
- speciální znak pro vyjádření každého **fónu**
- mezinárodně **standardní zápis** – jsou k dispozici tabulky a fonty
- *Unicode* – speciální IPA znaky v rozsahu U+0250–02AD

SAMPA:

- *Speech Assessment Methods Phonetic Alphabet*
- vznikla v projektu SAM (Speech Assessment Methods) v letech 1987–89
- **strojově čitelná** fonetická abeceda
- `http://www.phon.ucl.ac.uk/home/sampa/`

IPA – souhlásky

v **americké angličtině** – *pulmonické* i *nepulmonické*

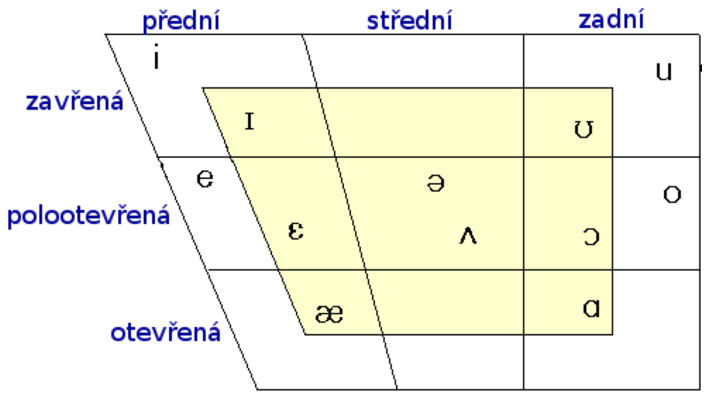
	labiála	labio-dentála	dentála	alveolára	palatála	velára	glotála
ploziva	p b			t d		k g	
frikativa		f v	θ ð	s z	ʃ ʒ		h
afrikáta					tʃ dʒ		
nazála	m			n		ŋ	
aproximanta laterální				l			
retroflexní				r			
koartikulovaná	w				j		

IPA – souhlásky ve slovech

- p plate, piece, spin, capital, stop, tramp
- t trip, time, winter, retire, wait, front
- k kite, climb, character, rocket, back, sink
- b bill, brush, sober, ramble, sob, bulb
- d dark, drive, redden, ponder, head, hard
- g go, grease, rigor, anger, log, iceberg
- m man, mile, remorse, ample, climb, harm
- n nice, know, enough, cunning, sign, burn
- ŋ finger, singer, drunk, rang, thing
- θ thank, three, ether, panther, path, birth
- ð then, these, feather, breathe
- f fit, fly, effort, perform, enough, Ralph
- v very, view, every, prevail, love, starve
- s ceiling, slim, psychology, Pacific, nasty, pass
- z zoo, zipper, hazard, prison, cares, breeze
- ʃ shore, sugar, nation, rash, Porche
- ʒ (genre), visual, measure, decision, massage
- h hat, who, ahead, perhaps
- tʃ China, cheap, ritual, teaching, beach, punch
- dʒ jump, pidgeon, reject, individual, ridge, engine
- l light, look, pillow, applaud, salt, ball, girl
- r real, row, around, part, care, hear
- w wind, was, await, swim, queen
- j yes, use, beyond, beauty, punitive

IPA – samohlásky

v **americké angličtině**

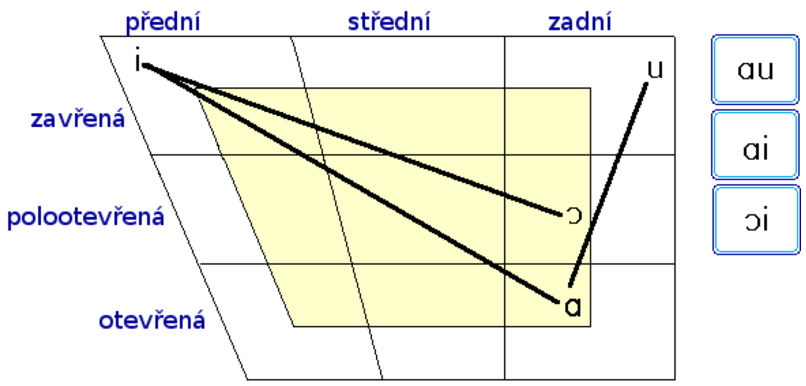


IPA – samohlásky ve slovech

- i heed, beat, believe, people, scary
- ɪ hid, bit, injure, resist, finish
- e hate, bait, great, they, say, neighbor
- ɛ head, bet, friend, says, guest
- æ had, bat, laugh, calf, language
- u food, boot, pool, through, who, sewer
- ʊ hood, book, pull, put, would
- o hole, boat, sew, know, so
- ɔ bought, law, wrong, stalk
- ɑ pot, "la", stocking, father, rob
- ə above, around, sofa, police
- ʌ bus, rush, under, other

IPA – dvojhlásky

v **americké angličtině**



ai find, high, aisle, quiet, ride
au house, crown, around, flower, how
ɔi boy, enjoy, Freud, avoid, join

Prozodie

- tzv. **suprasegmentální rysy**
- popisuje řečový proud spolu s přepisem do fonémů
- vyjádřena pomocí dalších **lokálních fyzikálních charakteristik** výsledné zvukové vlny:
  - **délka** fonému
  - **intonace** věty – vzor pro hladinu **základní frekvence** (*pitch*)
  - **tón** – v některých (tzv. **tónových**) jazycích určuje význam
  - **přízvuk** – v **přízvukových jazycích** ovlivňuje délku, hlasitost a tón slov
- kvalitní výpočet prozodie = **přirozenost** syntetizované řeči

Text-to-Speech systémy

- **syntéza řeči** – převod psaného textu na (digitální) zvuk
- TTS, *Text-to-Speech*
- dvě hlavní části
  1. **jazykový modul**, NLP modul
     vstup = text
     výstup = fonémy + prozodická informace
     označována také jako TTP, *Text-to-Phoneme*
  2. **modul zpracování signálu**, DSP (Digital Signal Processing) modul
     vstup = výstup z NLP modulu
     výstup = zvukový soubor

Příklady dat pro českou syntézu s MBROLA

- pravidla pro přepis do fonémů

```
CLASS SA   [aáeéěiíoóuúůyý]  # samohlásky
CLASS ZPS  [bdd'gvzžhCČ]     # znělé párové souhlásky
CLASS NPS  [ptt'kfsšHcč]     # neznělé párové souhlásky
[[ dě ]] → d' e
[[ b ]] (_|NPS|ZPS_) → p
[[ p ]] ZPS → b
```

- vstup pro MBROLu – text "shání tě též muž"

```
_ 200 0 132        i: 93 0 114       S 81 0 114
z 57 0 115         t' 27 0 120       m 43 0 120
h 45               e 50 0 114        u 61
a: 137             t 31 0 120        S 110
n' 75 0 132        e: 102            #
```

- zvuková databáze cz2 – 37 fonémů, 1442 difónů
  nutné ručně "nařezat" všechny difóny

Příklady TTS systémů

- české
  - **Epos** – z 90. let, Karlova univerzita a ČAV, nejlepší český open source
  - **Demosthenes** – FI MU Brno, laboratoř LSD
    slabiková syntéza, základní prozodie
  - **ARTIC** (ARtificial Talker In Czech) – ZČU Plzeň, **DEMO**
    obsahuje i "Talking head" vizuální část
  - **CS-Voice 97** – komerční, Frog Systems, pro Windows
- zahraniční
  - **Festival** – z Edinburghu, GPL, hodně jazyků, projet Festival Czech
  - **MBROLA** – difónová syntéza MBR-PSOLA, řeší DSP část
    Mikuláš Piňos, DP 2000 – česká DB pro MBROLu, `text2phone` v Perlu
  - mnohé další – **HADIFIX**, **SVOX**, **Bell Labs**, **AT&T**, ...