

# FONOTAKTIKA

- Část fonologie zabývající se syntagmatickými vlastnostmi fonologických segmentálních jednotek (segmentů)

Dva druhy jazykových prvků:



Jazyková promluva má dvě stránky (= dvě úrovně popisu):

### **Segmentální „taktický“**

- promluvu lze rozdělit na posloupnost segmentů
- segmenty mají víceméně pevné hranice (= jsou diskrétní)
- segment může být rozdělen na posloupnost menších segmentů

### **Suprasegmentální (prozodický) „para-taktický“**

- jevy, které se vážou na soubory segmentů
- nemusí mít jasně stanovitelné hranice, mohou mít kontinuální (nediskrétní) povahu
- obvykle nejdou rozdělit na menší části

## VZTAHY MEZI JAZYKOVÝMI JEDNOTKAMI

paradigmatické	×	syntagmatické
bud'/anebo	×	a zároveň
neuspořádané	×	uspořádané
(symetrické)		(symetrické)

[lɛs]

[pɛs]

[jɛs]

# NEUSPOŘÁDANÉ × USPOŘÁDANÉ SYSTÉMY

	Neuspořádaný systém	Uspořádaný systém	Para-taktický	
Figury	<b>Fonématica</b>	<b>Fonotaktika</b>	<b>Prozódie (para-fonotaktika)</b>	Fonologie
Znaky	<b>Morfologie</b>	<b>Syntax</b>	<b>Větná úroveň (para-syntax)</b>	Gramatika

Segmentální úroveň	Suprasegment. úroveň
--------------------	----------------------

Fonotaktika = „fonologická syntax“

## ÚKOLY A ZÁJMY FONOTAKTIKY

- Popsat pravidelnosti a omezení v **distribuci** (výskytu) fonémů a v jejich vzájemných **kombinacích** v konkrétních jazycích
- Vysvětlit, proč tyto pravidelnosti a omezení existují (artikulační, percepční, historické důvody, příznakovost)
- Srovnat různé jazyky a odhalit, co mají společného a čím se liší + otázka, zda to vůbec jde

## Konkrétní fonotaktické problémy:

- Určit, které kombinace jsou (i) možné/přípustné, (ii) nemožné/nepřípustné a (iii) potenciálně možné
- Stanovit frekvenci výskytu různých kombinací a struktur
- Popsat strukturu jednotek vyššího řádu (než segment) (slabiky, morfému, slova, syntagmatu atd.)
- Určit, jak jednotky vyššího řádu dělíme na menší jednotky („problém slabikování“)

- Korelovat fonotaktické vlastnosti s jinými vlastnostmi a určit, jaké vlastnosti to mohou být
  - a) fonologické vlastnosti
    - fonematika (velikost inventáře fonémů)
    - prozódie (tóny, přízvuk)

kompenzace nebo podpoření komplexnosti?

		Slabičná stavba			Celkem jazyků
		složitá	středně slož.	jednoduchá	
Systém tónů	složitý (3 a víc tónů)	10,4 %	75,3 %	14,3 %	77
	jednoduchý (2 tóny)	19,3 %	63,0 %	17,6 %	119
	žádný	40,7 %	49,1%	10,2 %	275
Celkem jazyků		143	268	60	471

- b) gramatické/sémantické/lexikální vlastnosti  
např. typy jazyků (flektivní, aglutinační atd.), slovní druh  
(ohebný × neohebný), kořen × afix
  
- c) historické (diachronní) vlastnosti  
domácí × přejaté výrazy, přejatá slova podle stáří
  
- d) kvantitativní vlastnosti  
např. frekvence slov, ale také třeba počet mluvčích



## e) mimojazykové vlastnosti

geografie:

jižní a střední Amerika → jednoduchá slabičná

struktura

ekologické faktory:

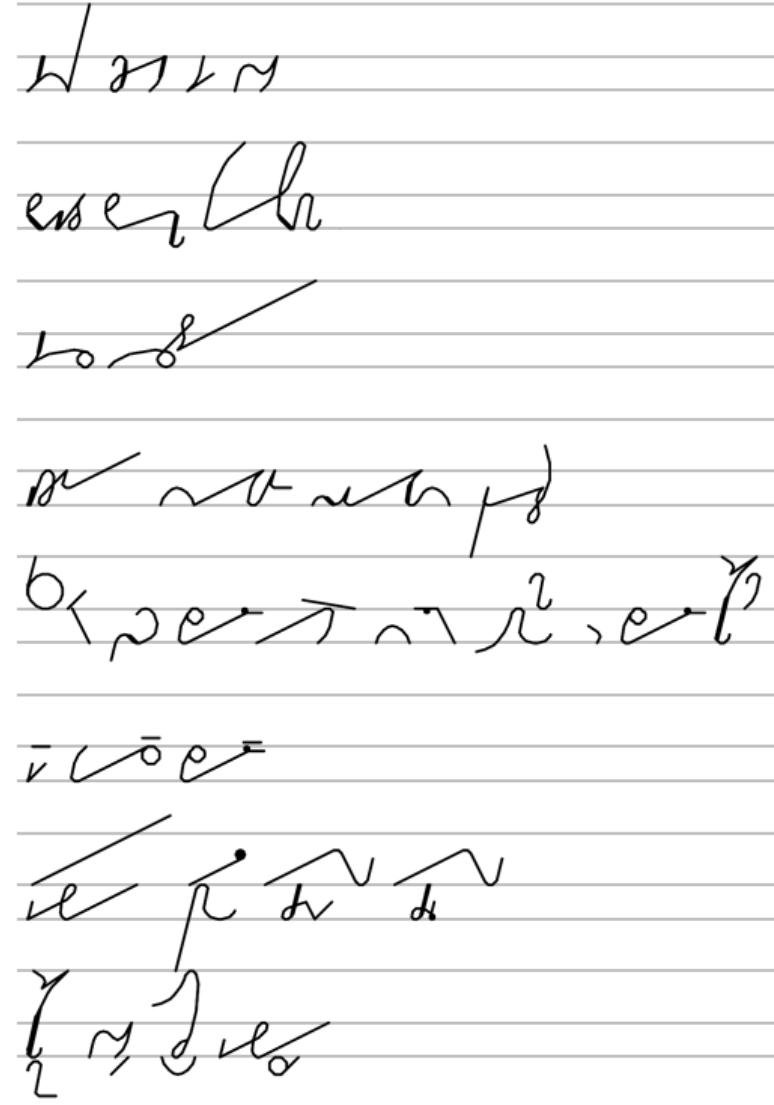
např. teplé klima → více slabik CV

větší hustota zalesnění → jednodušší slabičná

struktura

## DŮLEŽITOST A VYUŽITÍ FONOTAKTIKY V PRAKTICKÉM ŽIVOTĚ

- Porozumění jazykům
- Tvorba nových slov (např. názvů výrobků)
- Kódování, šifrování včetně rozluštění písem starých jazyků
- Automatická identifikace (detekce) jazyků nebo typů slov
- Výuka cizích jazyků
- Tvorba těsnopisu



## **METODY A PŘÍSTUPY UŽÍVANÉ VE FONOTAKTICKÝCH ANALÝZÁCH**

- Sběr dat a ověřování hypotéz: korpusy × experimenty
- Vyhodnocení a prezentace dat: kvalitativní (grafy, schémata, výčty...) × kvantitativní (statistika, četnost, frekvence...)
- Platnost závěrů a teorií: univerzální × jazykově specifická pravidla a zobecnění

## DATABÁZE S FONOTAKTICKÝMI INFORMACEMI

### a) Srovnávací (mezijazykové)

#### **ULSID (UCLA Lexical and Syllabic Inventory Database)**

- nedostupná, ale popsána v Rousset, Isabelle. 2004. *Structures syllabiques et lexicales des langues du monde*. Nepublikovaná disertace. <https://tel.archives-ouvertes.fr/tel-00250154/document>
- 14 (resp. 15) jazyků, obsahuje informace o slabičné struktuře
- navazuje na UPSID (*UCLA Phonological Segment Inventory Database*, viz <http://web.phonetik.uni-frankfurt.de/upsid.html>, 451 jazyků, Ian Maddieson)

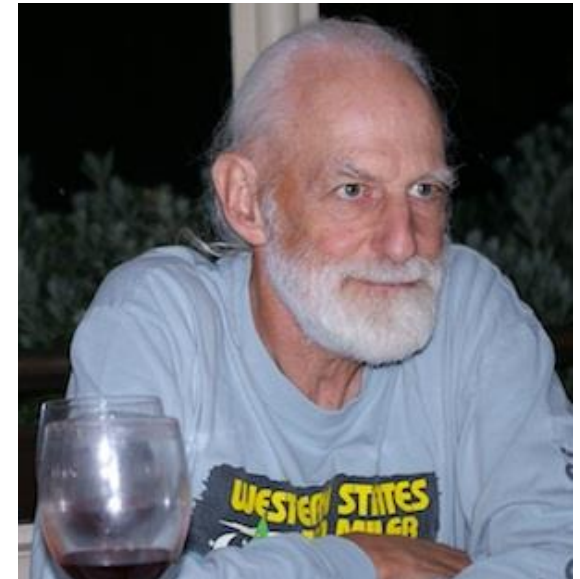
## WALS Online (World Atlas of Language Structures)

- <https://wals.info>
- informace o slabičné struktuře 486 jazyků (Ian Maddieson)
- vzorek 100 jazyků analyzován v Gordon, Matthew. 2016. *Phonological Typology*. Oxford University Press.

## LaPSyD (Lyon-Albuquerque Phonological Systems Database)

- <http://www.lapsyd.ddl.cnrs.fr/lapsyd/index.php>
- v současnosti 625 jazyků včetně češtiny
- informace o inventáři fonémů a slabičné struktuře
- základ tvoří Maddiesonova databáze

Ian Maddieson (\*1942–) →



## World Phonotactics Database (2013)

- informace o fonotaktice více než 2 000 jazyků
- již nedostupná, ale některé údaje jsou v Gordon, Matthew. 2016. *Phonological Typology*. Oxford University Press.

## Phonotaction 1.0

- informace o slabičné stavbě 516 jazyků („lektů“) na území Eurasie
- jaké segmenty a shluky se mohou vyskytovat v préture a kodě slabiky
- informace o slabičném jádru (jednoduché × komplexní vokály × slabičné konsonanty)
- <https://github.com/ianjoo/Phonotacticon>
- [https://assets.researchsquare.com/files/rs-3269302/v1\\_covered\\_a36f6a0f-09aa-4c57-bdee-09f672f46e38.pdf?c=1700483869](https://assets.researchsquare.com/files/rs-3269302/v1_covered_a36f6a0f-09aa-4c57-bdee-09f672f46e38.pdf?c=1700483869)

## *b) Konkrétní jazyky (příklady)*

### **Phonological Corpus of Czech (2016–)**

- <http://www.ujc.cas.cz/phword>
- více než 460 000 českých slov ve fonetické a fonologické transkripci rozdělených na slabiky
- lexikální část: slova z hlavních výkladových slovníků češtiny + databáze výrazů propriální a terminologické povahy (názvy obcí a jejich částí, pomístní jména, nejfrekventovanější křestní jména a příjmení, názvy hradů a zámků, zoologické a botanické názvy)
- součástí dvě databáze: *Fonologická databáze domácích slov* (cca 33 tis. slov) a *Fonologická databáze přejatých slov* (cca 36 tis.)
- textová část: texty (beletrie) + slovní tvary z Českého nár. korpusu

## Celex 2

- databáze slovní zásoby angličtiny (52 tis.), němčiny (51 tis.) a nizozemštiny (124 tis.) ve fonologické transkripci (včetně suprasegmentální rysů)
- slova rozdělena na slabiky
- obsahuje i morfologické a syntaktické informace

## GreekLex2

- databáze slovní zásoby moderní řečtiny (35 tis.) ve fonologické transkripci
- slova rozdělena na slabiky
- obsahuje informace i o frekvenci slabik

podobné databáze existují pro další jazyky (rumunština, švédština, norština, islandština, italština, farština)



# Semestrální úkol

## STATISTICKÁ FONOTAKTICKÁ ANALÝZA CV A VC KOMBINACÍ V NEINDOEVROPSKÝCH JAZYCÍCH

dílčí úkoly:

- 1) najít reprezentativní databáze 2 ne-IE jazyků obsahující fonetickou / fonologickou transkripci slov (popř. morfémů)  
**do 12. 3.**

vyhledávač:

„[jazyk] lexical database pronunciation“

„[jazyk] vocabulary phonological/phonetic transcription“

- 2) upravit databáze (např. změny  $\widehat{/tʃou/} \rightarrow /čö/$ )  
princip „jeden segment = jeden znak (může obsahovat diakritiku)  
**do 2. 4.**
  
- 3) vypočítat poměr mezi doloženou a očekávanou frekvencí  
**do 23. 4.**
  
- 4) interpretovat výsledky, srovnat oba jazyky, vyvodit s nich nějaké závěry a konfrontovat je s tím, co je o CV a VC kombinacích napříč jazyky známo  
**do 7. 5.**