

CORE023

Čeština a jazyková typologie

Markéta Ziková (ÚČJ FF MUNI)

12. 10. 2022

Jazyková typologie

=> zkoumání strukturních vlastností jazyka

=> identifikace opakujících se strukturních vzorců napříč jazyky

=> reprezentace těchto vzorců

Čeština

tradovaná tvrzení o češtině:

=> „je výjimečná tím, že má hodně pádů“

=> „je výjimečná tím, že má hlásku ř“

...

Čeština a pády

Sklo se rozbilo.

nominativ

Pět běžců nedoběhlo.

genitiv

Dal to Petrovi.

dativ

Potkali jsme sousedku.

akuzativ

Běž pryč, ty otrapo!

vokativ

Ležel v rybníce.

lokál/lokativ

Napsal to tužkou.

instrumentál

Typologie pádů

The World Atlas of Language Structures ([WALS](#))

=> vzorek 261 jazyků

=> sledovaný parametr: [počet morfologicky vyjádřených pádů](#)

Struktura přednášky

=> syntaktická typologie: seřazení slov ve jmenné frázi

ty tři bílé myši

=> fonologická typologie: seřazení souhlásek/konsonantů na začátku slov

trup vs. rtuť

Syntax

=> hlavní složky věty: **NP** (nominal phrase) a **VP** (verbal phrase)

Zuřivý pes bez náhubku pokousal pět dětí.

[Zuřivý pes bez náhubku_{NP}] [pokousal [pět dětí_{NP}]_{VP}]

Jmenná fráze

ty	tři	bílé	myši
those	three	white	mice
Dem	Num	A	N

=> Greenberg (1963): Universal 20

=> generalizace o seřazení ukazovacích zájmen (Dem), číslovek (Num), přídavných jmen (Adj) a jmen (N) na rozsáhlém vzorku jazyků

Cinque (2004)

D3AN	D3NA	DN3A	ND3A
*3DAN	*3DNA	*3NDA	*N3DA
*AD3N	*ADN3	AND3	NAD3
*AD3N	DAN3	DNA3	NDA3
*3ADN	3AND	3NAD	N3AD
*A3DN	*A3ND	AN3D	NA3D

=> logicky možné kombinace: 24

=> doložené kombinace: 14

*nedoložené vzorce

nejfrekventovanější vzorce

D3AN (*ty tři bílé myši*)

čeština, angličtina, korejština

NA3D (*myši bílé tři ty*)

thajština, jorubština, maorština

Univerzální struktura

=> univerzální hierarchická struktura: Dem >> Num >> A >> N

=> odpovídá algoritmu budování významové struktury

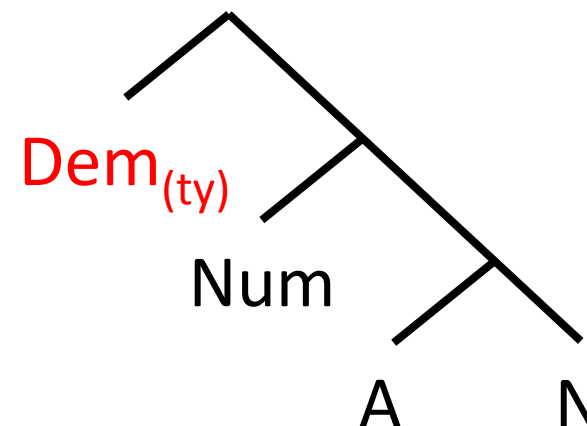
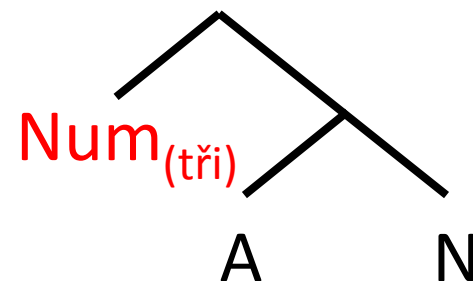
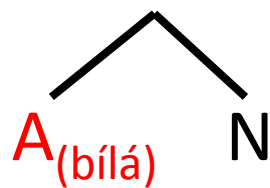
1. věc označím

2. charakterizuju

3. spočítám

4. ukážu na ni

$N_{(myš)}$

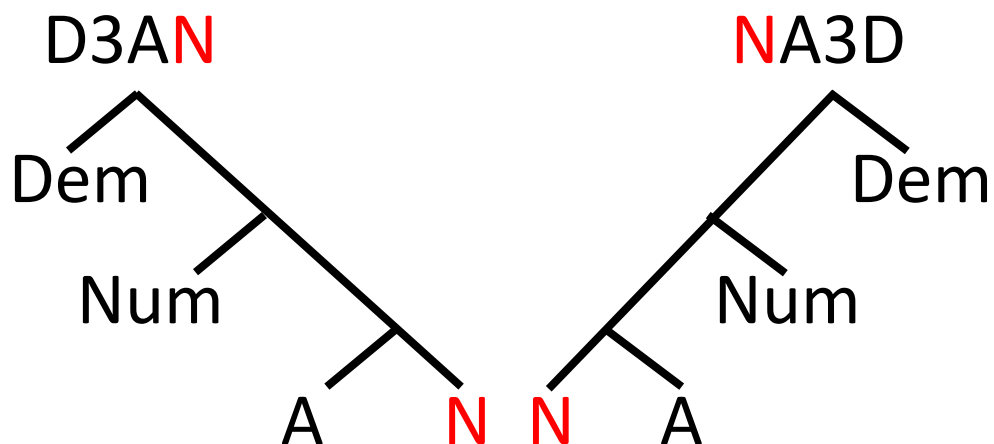


Derivace vzorců

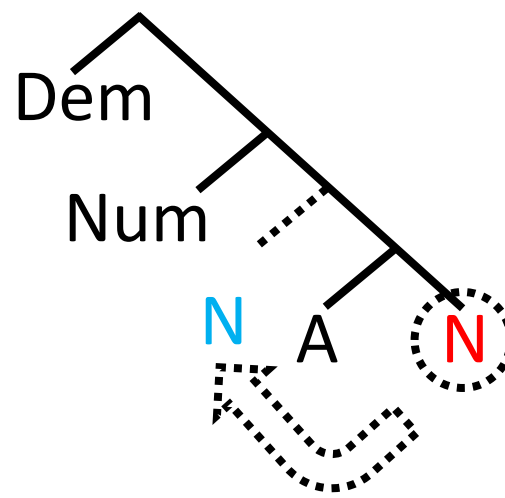
=> jednotlivé vzorce jsou derivovány ze základní univerzální struktury

=> 2 nejfrekventovanější vzorce: **D3AN** (čeština) a **NA3D** (thajština)

=> shoda = izomorfie (zachování hierarchie), rozdíl = lineární seřazení



D3AN -> **D3NA** (abcházština, perština)



Fonologie

=> 2 hlavní typy hlásek: samohlásky/**vokály** X souhlásky/**konsonanty**

ploutvička = **CCVVC****CVCCV**

=> různý způsob artikulace i různé akustické vlastnosti

	artikulace	akustika
C	striktura	šum
V	apertura	složený tón

Dvojice konsonantů na začátku slova: #CC

	čeština	angličtina		čeština	angličtina
pr	prak	price	rp	Rpety	*
br	brečet	broom	rb	*	*
kr	kráva	cream	rk	*	*
dr	dráha	dream	rd	rdousit	*
pl	plyn	plain	lp	lpět	*
bl	blázen	blast	lb	lbi (leb)	*
kl	klepat	clean	lk	lkát	*
dl	dluh	*	ld	*	*

náhodné mezery vs. **systematické** mezery (= gramaticky podmíněné)

#CC: typologie (Greenberg 1978)

[p b k d] = šumové konsonanty/obstruenty => T

[r l] = sonorní konsonanty => R

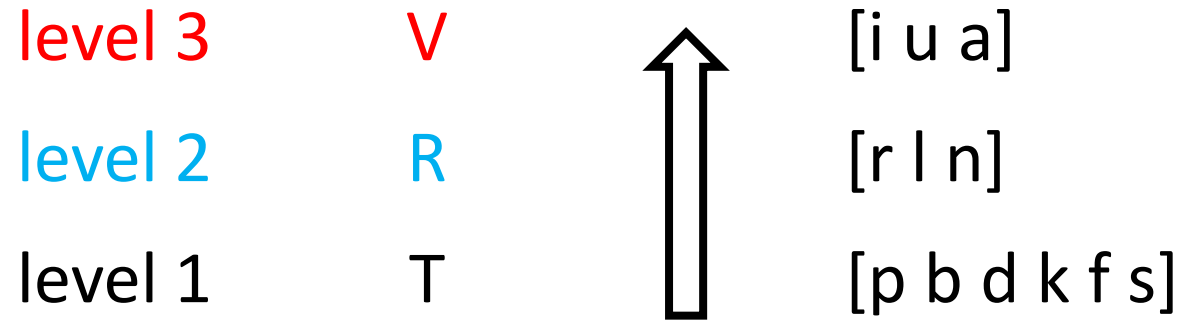
generalizace: 2 typy jazyků (pokud jde o fonotaktiku #CC)

	#RT	#TR	
typ 1	✓	✓	čeština, polština, gruzínština, paštunština
typ 2	✗	✓	angličtina, dánština, albánština
typ 3	✓	✗	*

Sonorita

=> T a R mají různou distribuci, protože mají různou sonoritu

=> sonorita = skalární vlastnost, koresponduje s akustikou a artikulací



#TR vs. #RT

#TR

stoupající sonorita

#RT

klesající sonorita

level 3

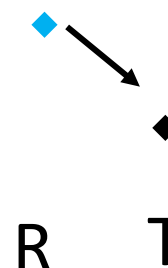
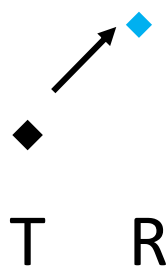
V

level 2

R

level 1

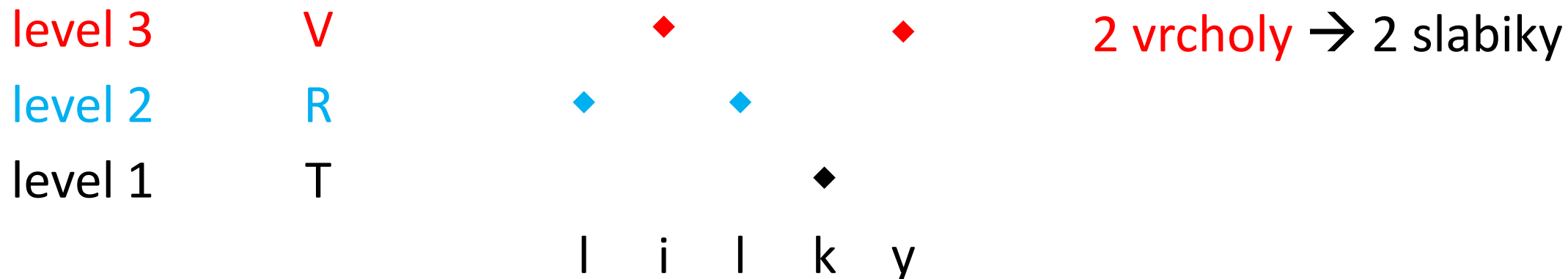
T



Slabičná struktura

=> začátek slova = začátek slabiky

=> počet slabik odpovídá počtu sonorních vrcholů ve slově



**li.lky* vs. *lil.ky*

=> sonorní vrchol = jádro slabiky

=> sonorita uvnitř slabiky směrem k jádru stoupá

= **Sonority Sequencing Principle** => univerzální struktura

level 3

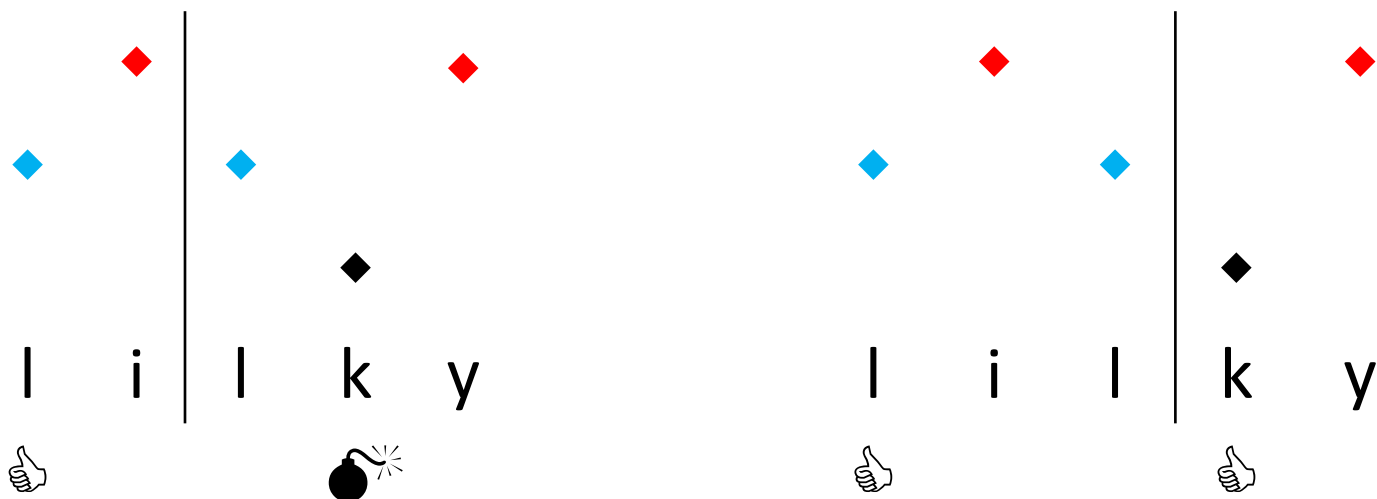
V

level 2

R

level 1

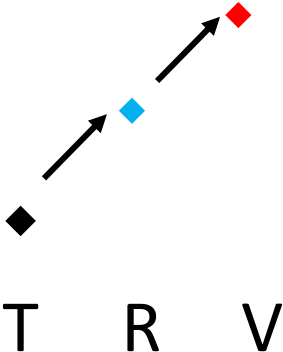
T



#TR vs. #RT

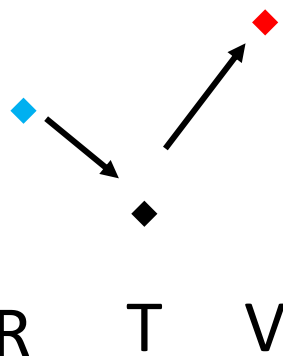
level 3 V
level 2 R
level 1 T

trus



TR = začátek slabiky

rtuť



R = extraslabičný C

Derivace typologie

	#RT	#TR	
typ 1	✓	✓	=> také extraslabičné konsonanty
typ 2	✗	✓	=> bez extraslabičných konsonantů
*typ 3	✓	✗	=> jenom extraslabičné konsonanty

=> implikace: specifické (extraslabičné #RT) implikuje defaultní (#TR)

Lhasa

	#RT	#TR	
typ 1	✓	✓	čeština: <Lhasa> = [l ^h asa]
typ 2	✗	✓	angličtina: <Lhasa> = [la:sə]