

IV.

Vzájemné vztahy formální morfologie a derivace z hlediska možností automatické slovtvorné analýzy

Automatická slovtvorná analýza se opírá o pravidelnosti formálních změn, k nimž dochází při derivaci. Derivaci lze formálně popsat jako záměnu/záměny části/částí slova základového, které mají za následek vznik slova odvozeného. Slova základová, od nichž potenciálně mohou být derivována slova odvozená, lze definovat na rovině gramatické abstrakce zakódované v morfologických značkách. Tyto vztahy jsou do různé hloubky popisovány v klasických popisech české slovtvorby. Předložený text si klade za cíl na základě pozorování dat strojového slovníku češtiny a jazykových korpusů vytvořit formální popis a otestovat a) možnosti a meze formalizace; b) doplnit stávající popisy o případy, které dosud nebyly zaznamenány.

Morfematická segmentace slovesných tvarů a deverbativ

Řešení otázky segmentace slovesných tvarů a deverbativ je pro popis tvarosloví sloves a tvoření deverbativ klíčové. Týká se i problematiky morfologických alternací (alomorfie), konkrétně nejrůznějších alternací, které doprovázejí tvoření slovesných tvarů (konjugace) i některých pravidelných derivací od slovesného kmene (adjektivum slovesné na *-ný/-tý* – VA, substantivum slovesné na *-ní/-tí* – VSB). Těm se budeme věnovat v následujících kapitolách. V této kapitole se chceme zmínit o problémech spjatých se segmentací (morfémovou analýzou) slova v oblasti strojové analýzy přirozeného jazyka.

Otázce pravidel segmentace slovního tvaru (morfémové analýzy) jsou věnovány úvodní studie řady morfematických a retrográdních slovníků (Worth – Kozak – Johnson 1970), (Slavičková 1974), (Tichonov 1985), (Šiška 1998), (Sokolová a kol. 1999).

Hloubka morfémové analýzy je mnohdy závislá na účelu příslušné analýzy. Při segmentaci slovesných tvarů a interpretaci segmentů se přidržíme systému použitého v Mluvnici češtiny 2 (Komárek a kol. 1986). Jak jsme uvedli v úvodu, jádrem naší práce bude analýza mezí a možností automatického zpracování derivace některých typů deverbativ. Za tímto účelem se v rámci této kapitoly budeme věnovat zásadám segmentace slovesného tvaru a deverbativ.

U sledovaných typů deverbativ se přidržíme zásad segmentace E. Slavičkové (Slavičková 1974), protože tyto zásady (aplikované na korpusy nesrovnatel-

ně menšího rozsahu, než jsou ty, které máme dnes k dispozici) se nám nejeví v žádném směru jako překonané.⁸

Termíny, které zavádíme v následujícím přehledu, odpovídají (až na výjimky – konekt) termínům použitým v Mluvnici češtiny 2 (Komárek a kol. 1986). Zavedení segmentu, který nazýváme konekt, vychází ze zkušenosti se segmentací některých okrajových typů derivací v partiích věnovaných popisu tvoření slov v českých mluvnicích i v některých výše uvedených (českých) pracích slovníkového charakteru.

U slovesného tvaru budeme tedy rozlišovat následující typy segmentů:

prefix (_{1-n})	tvarotvorný základ (kořen)	kmenotvorná přípona (vč. 0)	konekt	tvarotvorný formant	tvarová koncovka
<i>ne-u-po-</i>	<i>-třeb-</i>	<i>-i-</i>		<i>-l-</i>	<i>-a</i>
<i>za-</i>	<i>-kry-</i>	<i>-0-</i>		<i>-t-</i>	<i>-a</i>
<i>ne-u-po-</i>	<i>-třeb-</i>	<i>-í-</i>			<i>-me</i>
<i>ne-na-</i>	<i>-sáz-</i>	<i>-ej-</i>			<i>-te</i>
<i>u-</i>	<i>-kry-</i>		<i>-v-</i>	<i>-š-</i>	<i>-e</i>

U deverbativ odvozených od slovesného tvaru konverzí budeme rozlišovat následující typy segmentů:

prefix (_{1-n})	tvarotvorný základ (kořen)	kmenotvorná přípona (vč. 0)	konekt	tvarotvorný formant	tvarotvorný formant	tvarová koncovka
<i>u-za-</i>	<i>-vř-</i>	<i>-e-</i>	<i>-v-</i>	<i>-š-</i>	<i>-í-</i>	<i>-ho</i>
<i>o-</i>	<i>-pi-</i>	<i>-0-</i>		<i>-l-</i>	<i>-ý-</i>	<i>-m</i>
	<i>nes-</i>	<i>-0-</i>		<i>-ouc-</i>	<i>-í-</i>	<i>-mi</i>
<i>u-</i>	<i>-tř-</i>	<i>-e-</i>		<i>-n-</i>	<i>-ý-</i>	<i>-ch</i>
	<i>kut-</i>	<i>-i-</i>		<i>-l-</i>	<i>-0-</i>	<i>-em</i>
<i>vy-</i>	<i>klouz-</i>	<i>-0-</i>		<i>-0-</i>	<i>-0-</i>	<i>-ovi</i>

U deverbativ odvozených od slovesného tvaru sufixací budeme rozlišovat následující typy segmentů:

prefix (_{1-n})	tvarotvorný základ (kořen)	kmenotvorná přípona (vč. 0)	konekt	tvarotvorný formant	slovotvorný formant	tvarová koncovka
<i>roz-</i>	<i>-mázl-</i>	<i>-0-</i>		<i>-en-</i>	<i>-0c-</i>	<i>-i</i>
<i>o-</i>	<i>-žr-</i>	<i>-a-</i>		<i>-l-</i>	<i>-0c-</i>	<i>-em</i>
<i>o-</i>	<i>pi-</i>	<i>-0-</i>		<i>-l-</i>	<i>-ec-</i>	<i>-0</i>
<i>za-</i>	<i>-ry-</i>	<i>-0-</i>		<i>-t-</i>	<i>-ec⁹</i>	

8 Zásady segmentace slovního tvaru jsou bez jasného zdůvodnění masivně měněny vzhledem k dosavadní praxi české mluvnické i lexikografické tradice v kapitole věnované tvoření slov v Mluvnici současné češtiny (Cvrček a kol. 2010, s. 81–124).

9 Upozorňujeme na segmentaci substantiva *zarytec*, které chápeme jako odvozené od tvaru participia/adjektiva slovesného : *zarýt/zaryt/zarytj/zarytec*, na rozdíl od níže uvedeného substantiva *rytec*.

U deverbativ odvozených od slovesného kmene budeme rozlišovat následující typy segmentů:

prefix (_{1-n})	tvartvorný základ (kořen)	kmenotvorná přípona (vč. 0)	konekt	tvartoslovný formant (_{1-n})	tvartvorný formant	tvárová koncovka
<i>s-</i>	<i>-běr-</i>	<i>-a-</i>	<i>0</i>	<i>-tel-</i>	<i>-0-</i>	<i>-em</i>
	<i>hnět</i>	<i>0</i>	<i>-a-</i>	<i>-č-</i>	<i>-0-</i>	<i>-ů</i>
<i>ne-</i>	<i>-trp-</i>	<i>-ě-</i>	<i>-l-</i>	<i>-iv-</i>	<i>-ý-</i>	<i>-mi</i>
<i>nej-ne-</i>	<i>-pře-</i>	<i>-j-</i>		<i>-íc-n-ějš-</i>	<i>-í-</i>	<i>-mi</i>

U deverbativ odvozených od slovesného kořene budeme rozlišovat následující typy segmentů:

prefix (_{1-n})	kořen	konekt	tvartoslovný/tvartoslovné formant/formanty	tvartvorný formant	tvárová koncovka
	<i>soud-</i>		<i>-c-</i>		<i>-e</i>
	<i>ry-</i>	<i>-t-</i>	<i>-0c-</i>		<i>-em</i> ¹⁰
<i>vý-</i>	<i>-běr-</i>	<i>0</i>	<i>-č-</i>	<i>-l-</i>	<i>-m</i>
	<i>smír-</i>		<i>-č-</i>	<i>-l-</i>	<i>-ho</i>
<i>pří-</i>	<i>-sa-</i>	<i>-v-</i>	<i>-n-</i>	<i>-ý-</i>	<i>-ch</i>
<i>ú-</i>	<i>-spě-</i>	<i>-š-</i>	<i>-n-</i>	<i>-ě-</i>	<i>-m</i>

10 Upozorňujeme na segmentaci substantiva *rytec*, které chápeme jako odvozené od základu (kořene) *ryt/rytec*, na rozdíl od výše uvedeného substantiva *zarytec*.