

Morfém – kořeny vs. afixy, typy afixů

Úkol:

U různých druhů morfů, ať už jde o kořenové morfy/lexikální kořeny, afixální morfy (slovotvorné i tvarotvorné) se vydělení morfu děje na základě opakování a analogie. Rodilý mluvčí většinou nemá problém odlišit případy, kdy jde o morf a kdy jde o náhodnou kombinaci /řetězec hlásek/písmen. Tyhle problémy nám běžně chystají našeptavače. Opačně formulovaný dotaz ovšem může dělat mluvčímu problémy. A zde můžeme najít odpověď právě v korpusu.

Slova *kočka*, *kočičí*, *kočkovitý* mají stejný kořenový morf. Najděte slova, v nichž se a) vyskytuje řetězec *koč*, který b) není kořenový morf opakující se ve trojici výše uvedených slov. Vytvořte dotaz, pomocí kterého najdete doklady v korpusu.

Slovo *černozem* se skládá ze dvou kořenových morfů *čern* a *zem* a dále konektu *o*. Umíte si vzpomenout na další podobná slova. Zkuste vzít na pomoc korpus a sestavte dotaz, kterým byste si pomohli, pokud byste chtěli najít příklady stejně utvářených slov v korpusu.

Vzpomenete si na slova, která by obsahovala řetězec *ař*, *ář*, *íř*, *yř* a přitom by nešlo o derivační sufix? Jak byste hledali doklady v korpusu.

Vzpomenete si na slova, které by obsahovala řetězec *ky*, *aty*, *avy*, a přitom by nešlo o derivační sufix? Jak byste hledali doklady v korpusu.

Kombinatorika morfémů – kompozicionalita, produktivita, rozdíl mezi kořenem a kmenem, alomorfie (supletivní alomorfy a jejich distribuce, fonologicky podmíněné alomorfy, morfonologické alternace)

Jak se vypořádat s otázkou a jak nám může pomoci jazykový korpus při přípravě?

Rozdíl mezi morfem (znakovou jednotkou) a náhodným řetězcem znaků (grafických znaků - grafémů) je jedním z hlavních problémů, které řeší aplikace NLP. Na úrovni lingvistických analýz můžeme narazit na řadu problémů, které nám pomohou uvědomit si (noticing), že i pro lingvistickou analýzu může hrát roli ambiguita v oblasti morfologie.

Uvědomme si, že alomorfie je založena na HA. Které hlásky se často střídají? Vyskytuje se hláska v nějakém slově, které pojmenovává předmět v mém okolí. Umím najít další slova, v nichž se alternace opakuje? Umím na základě uplatnění principu analogie a opakování zobecnit pozorování a ukázat, jak funguje morfémová analýza?

ZÁLUDNÉ PŘÍKLADY

Mějme substantivum *tužtička*. Z kolika morfů se skládá?

Návod: vytvořte (napište) slovotvornou řadu: *tuha- tužka- tužtička*

Uzávorkujte () gramatické/flektivní morfy **tučně** lexikální morfy: **tuh(a)- tužk(a)- tužtičk(a)**

Vydělte pomlčkami slovotvorné afixy a počítejte (nestřílejte): **tuh(a)- tuž-k-(a)- tuž-t-ičk-(a)**

A teď záludná otázka: jak byste vysvětlili alternaci *k/t*? Jde o běžnou alternaci? V co všechno se *k* mění, s kterými hláskami se ve flexi a slovtvorbě střídá? Ve 100milionovém korpusu je skoro 19 milionů slovních tvarů, v nichž se *k* vyskytuje. 4 nejfrekventovanější synsemantika jsou *rok*, *člověk*, *velký*, *řekl*. Vytvoříme-li slovtvorné hnízdo, nikde se *k t* nedostaneme. Jak byste našli data, která by pomohla? V korpusu? Tak jdeme na to.

. *žt[ij]. *

	Filtr	lemma	Freq	i.p.m.
1	p / n	nebožtík	568	4,66
2	p / n	služtička	28	0,23
3	p / n	idžtihád	26	0,21
4	p / n	nůžtičky	24	0,2
5	p / n	nebožtíkův	20	0,16
6	p / n	nádržtička	10	0,08
7	p / n	držtička	6	0,05
8	p / n	ponožtička	6	0,05
9	p / n	tužtička	6	0,05
10	p / n	nebožtíky	1	0,01
11	p / n	využití	1	0,01

Ale jak je to s *nebožtíkem*?

ne-boh-ý/ne-bož-t-ík

Buď musíte vysvětlit, kde se vzal afix *-t-* nebo afix *-tik-*

Kolik substantiv je v češtině tvořeno sufixem *-tik*?

1 / 4 (celkem: 163 položek)

	Filtr	lemma	Freq	i.p.m.
1	p / n	nebožtík	568	4,66
2	p / n	puntík	450	3,69
3	p / n	dortík	215	1,77
4	p / n	pultík	202	1,66
5	p / n	tatík	200	1,64
6	p / n	prstík	194	1,59
7	p / n	žertík	172	1,41
8	p / n	tloušťík	170	1,4
9	p / n	Fojtík	169	1,39
10	p / n	Bertík	109	0,9

Co můžeme dále dělat?

Můžeme v korpusu najít doklady *. *hý* a podívat se, zda a jak bychom sufixem *-ík* tvořili jména osob. Uvidíme-li, že se názvy osob tohoto typu netvoří, pak hledejme dále.

Podívejme se, zda se netvoří jiným sufixem, který alternaci *h/ž* působí (*čtyřnožec*).

Tvoří se sufixem *ík* deriváty od substantiv na *ec*? Podívejte se do korpusu (*Němčik*, *Japončik*).

Zkusme najít alespoň deriváty na *. *žčik* (*nosorožčik*).

Proč tedy máme *nebožtíka* a nikoli *nebožčika*? Protože nemáme *nebožce*. Opravdu? Podívejte se do korpusu.

Existovalo-li kdysi substantivum *nebožec*, není možné, aby právě ono stálo za současným *nebožtíkem*? Podívejme se, zda někde v korpusu toho *nebožčíka* nenajdeme?

https://www.korpus.cz/kontext/view?viewmode=kwic&pagesize=40&attrs=word&attr_vmode=visible-kwic&base_viewattr=word&ctxattrs=word&refs=%3Ddoc.did&q=~oEE6AqKsoQsy&QSEnable=1

https://www.korpus.cz/kontext/view?viewmode=kwic&pagesize=40&attrs=word&attr_vmode=visible-kwic&base_viewattr=word&refs=%3Ddoc.id&q=~jK0g6KOU6wGW

<http://kott.ujc.cas.cz/index.php?vstup=Nebo%9E%E8%EDk&hledat=fulltext&heslo=checked&Zvyraznit=&Vyber=&Zobrazeni=&Strana=&stranaVys=1>

ÚKOL

Sloveso *pečí* je odvozovacím základem řady příbuzných slov. Při tvoření slov dochází v češtině k hláskovým alternacím. Lexikální morf mívá řadu alomorfů. Doplňte následující tabulku o všechny možné doklady příbuzných slov. Pokud si nebudete moci vzpomenout, zda je některá kombinace přípustná, vezměte si na pomoc korpus.

	P-K	P-C	P-Č
É	<i>pék</i>	<i>pec</i>	<i>peč</i>
Í	<i>pík</i>	<i>píc</i>	<i>pič</i>
E	<i>pek</i>	<i>pec</i>	<i>peč</i>