

PROJEKT: Inovace ve vzdělávání v oboru čeština jako druhý jazyk

KLÍČOVÁ AKTIVITA 06: Inovace didaktických metod 2 – tvorba korpusu se zapojením studentů

NÁVOD K ANOTACI CHYBOVÉHO KORPUSU

- o pracovní verze 1 (30.4.2011)

Obsah

NÁVOD K ANOTACI CHYBOVÉHO KORPUSU	1
SEZNAM IKON	2
TERMÍNY.....	3
1 ÚVOD.....	4
1.1 Emendace a anotace	4
1.2 Manuální a automatická anotace.....	4
1.3 Míra podrobnosti chybové anotace	4
2 FORMÁLNÍ PROSTŘEDKY PRO ANOTACI.....	5
2.1 Roviny anotace	5
2.2 Vztahy mezi sousedními rovinami.....	7
2.3 Odkazování.....	7
2.4 Kódy chyb	8
2.5 Automatická anotace	8
3 ZÁKLADNÍ PRINCIPY EMENDACE / ANOTACE.....	9
3.1 Pravidla pro emendaci (prostou opravu textu).....	9
3.2 Pravidla pro anotaci.....	10
3.3 Technická poznámka	11
4 CHYBY OPRAVOVANÉ NA ROVINĚ 1 (R1).....	12
4.1 Jak naložit s přepisovými kódy	12
4.2 Obecné zásady emendace na R1.....	12
4.2.1 Typy chyb opravovaných na R1 a jejich značkování.....	13
4.2.2 Zásada minimálního zásahu do tvaru slova.....	15
4.2.3 Zásada postupné emendace.....	16
4.2.4 Zásada nejkratší smysluplné trasy.....	17
4.2.5 Opravy typu <i>dal – dál</i>	17
4.2.6 Autorská univerbizace	18
4.3 Nesprávný tvar <i>incor</i>	18
4.3.1 Nesprávná báze <i>incorBase</i>	18
4.3.2 Nesprávná flexe <i>incorInfl</i>	19
4.3.3 Auto: Ostatní nesprávné tvary <i>incorOther</i>	20
4.4 Nově utvořené / cizí / neidentifikovatelné slovo <i>fw</i>	21
4.4.1 Nově utvořené („české“) slovo <i>fwFab</i>	21
4.4.2 Cizí / neidentifikovatelné slovo <i>fwNc</i>	21
4.4.3 Přítomnost flexe <i>flex</i>	21
4.5 Chybná hranice slov <i>wbd</i>	22
4.5.1 Prefix oddělený mezerou a předložka bez mezery <i>wbdPre</i>	22
4.5.2 Neoprávněně rozdělená či spojená kompozita <i>wbdComp</i>	23

4.5.3	Ostatní chybně rozdělená nebo spojená slova wbdOther	23
5	CHYBY OPRAVOVANÉ NA ROVINĚ 2 (R2)	24
5.1	Narušení shody agr	24
5.2	Chyba ve vyjádření syntaktické závislosti (jinak než shodou) dep	28
5.2.1	Způsoby odkazování u předložkových vazeb.....	29
5.2.2	Opravy vedlejších vět	31
5.3	Chyba v analytickém slovesném tvaru a složených predikátech vbx	34
5.4	Chyba ve zvrtném výrazu rflx.....	35
5.5	Chyba v negaci neg	37
5.6	Přebytečné / chybějící výraz odd / miss.....	37
5.7	Auto: Chybný slovosled wo.....	38
5.8	Chyba v lexiku a frazeologii lex	39
5.9	Chyba ve významovém užití gramatické kategorie use	40
5.9.1	Neuzuální užití singuláru či plurálu	40
5.9.2	Chybné použití času.....	40
5.9.3	Chybné užití vidu	40
5.9.4	Nevhodně užitý komparativ nebo superlativ	41
5.9.5	Vokalizace předložek.....	41
5.10	Chyba v zájmeném odkazu ref.....	41
5.11	Sekundární chyba sec.....	42
5.12	Stylově nevhodný výraz styl.....	42
5.12.1	Obecněčeský výraz stylColl	42
5.12.2	Jiný stylově nevhodný výraz stylOther.....	43
5.13	Rozvrácená konstrukce disr.....	43
5.14	Problémová chyba problem.....	44
6	Kódy chyb	45
	chyba ve významovém užití gramatické kategorie	45

SEZNAM IKON



pamatujte si / vysvětlení / výklad



nedělejte / ne / nepatří sem



příklad / ukázka



prostudujte si / pozor / výjimka

TERMÍNY

přepis	
emendace	prostá oprava chybných tvarů
anotace	přiřazení chybové značky
R1	rovina, na níž se opravují izolované tvary slov
R2	rovina, na níž se opravuje
pavouk	způsob propojení výrazů na jednotlivých rovinách, který má jeden uzel
pahýl	částečná hrana u výrazů, které doplňujeme nebo odstraňujeme
odkaz	šipka ve směru důvodu opravy
uzel	X, střed hrany, na kterém umísťujeme chybové kódy
hrana	spojnice mezi výrazy na jednotlivých rovinách
chybový kód (značka, tag)	značka, kterou přiřazujeme určité chybě v textu
chybová taxonomie	seznam chyb, které se při anotaci značkují
rovina	R1, R2
doména	např. chybová doména <code>incor</code>
typ	např. chybové typy <code>incorBase</code> , <code>incorInfl</code>
FEAT	anotační program

1 ÚVOD

1.1 Emendace a anotace



emendace, tj. samotná oprava textu
➤ viz i rektifikace, hypotéza

anotace, tj. popis a klasifikace chyb v textu
➤ viz značkování

Instrukce jsou společné (a propojené) pro emendaci i anotaci.

Typy chyb s kódy a základními údaji jsou uvedeny v tabulce 6.1 **Kódy chyb**.

1.2 Manuální a automatická anotace



manuální anotace, tj. ruční značkování textu, které provádí anotátor

automatická anotace, tj. využití programových nástrojů k automatickému značkování textu

Návod popisuje manuální emendaci a anotaci. Výsledky manuální emendace a anotace využívá automatická procedura. Ta určuje typy chyb, které lze odvodit na základě porovnání původních a opravených forem, nebo zpřesňuje již určené typy chyb.



Některé automaticky určené typy chyb jsou uvedeny pro přehlednost i v tomto textu a v hlavičce jsou označeny např. takto: **Auto: Chybný slovosled wo** (viz např. 5.8).

1.3 Míra podrobnosti chybové anotace

V návodu je popsána anotace, která umožňuje pozdější rozšíření o detailnější klasifikaci chyb, automatickou či manuální.

2 FORMÁLNÍ PROSTŘEDKY PRO ANOTACI

The screenshot shows the annotation software interface. At the top, there are three text levels (R0, R1, R2) for the text: "Bydlíme ve svém domě. Máme hezkých psa a kočku. Rása psa." The R0 level shows the original text with red lines indicating connections to the R1 level. The R1 level shows the text with green lines indicating connections to the R2 level. The R2 level shows the text with green lines indicating connections to the R1 level. The R1 level has a green line labeled "HRANA" connecting "ve" to "svém". The R2 level has a green line labeled "UZEL" connecting "ve" to "svém". The R1 level has a green line labeled "ODKAZ" connecting "hezkých" to "hezkého". The R2 level has a green line labeled "CHYBOVÝ KÓD" connecting "hezkého" to "hezkých". The R1 level has a green line labeled "PAHÝL" connecting "hezkých" to "hezkého". The R2 level has a green line labeled "PAHÝL" connecting "hezkého" to "hezkých". The R1 level has a green line labeled "ODKAZ" connecting "hezkých" to "hezkého". The R2 level has a green line labeled "CHYBOVÝ KÓD" connecting "hezkého" to "hezkých". The R1 level has a green line labeled "PAHÝL" connecting "hezkých" to "hezkého". The R2 level has a green line labeled "PAHÝL" connecting "hezkého" to "hezkých".

Moje rodina. **PŘEPIS**

Má rodina není velká. Bydlím s otcem a mátkou. Jaem jeden syn v rodině. Ještě mám 2 nevládné sestry. Jedna má manžela a dceru. Její dcera je má nětet. Studuje ve střední škole. Druhá sestra bydlí v nem městě. Me rodič jsou uředníké. Mají malý bu biznes. Bydlíme v domě ve svém domě. Máme hezkých psa a kočku. Rása psa je „dvůrterjér“. Tak jmenuje ho můj tatněk. Pes se jm(e)jmuje „Šarik“. Kočka se jmenuje „Bars(j)eK“. Mame ještě velkou zahrádu. Tam je malý proud – rybník. S zlatými rybami. Má Můj otec je S[is], a proto bydlíme v domě i nebydlíme v bytě. Má maminka vaří moc dobře. Chutné mi j jídlo.

2.1 Roviny anotace



Anotační schéma:

rovina 0 (R0), tj. zkonvertovaný přepis originálního textu

rovina 1 (R1), tj. posloupnost správných tvarů existujících spisovných českých slov

rovina 2 (R2), tj. gramaticky správný text (věta)

Anotátor pracuje se stejným textem v několika podobách, které jsou v anotačním programu přístupné současně. Tři z těchto podob (roviny 0, 1 a 2) jsou součástí anotačního schématu. Chyby se opravují (tj. emendují) podle typu na rovině 1 nebo 2, jejich typy se určují v **uzlech**, které leží na spojnicích (tj. **hranách**) mezi původními a opravenými formami.

Anotační program zahrnuje:

- Faksimile rukopisného **originálu**
- **Přepis** originálu se zachováním některých úprav (rektifikací) autora a učitele (pravidla přepisu viz *Manuál pro přepis psaných materiálů*)

- **Rovina 0**

- obsahuje text z přepisu, zkonvertovaný do formátu vhodného pro anotaci
- obsahuje výslednou podobu textu, tj. na této rovině nejsou zachyceny autorské rektifikace, škrty apod.

- **Rovina 1**

- výsledkem opravy na R1 je posloupnost správných tvarů existujících spisovných českých slov s výjimkou slov, která nelze emendovat (viz bod 4.4 **Nově utvořené / cizí / neidentifikovatelné slovo**)
- zde se provádějí opravy týkající se jednotlivých slov a slovních tvarů, např.



*kostích -- kostech, chlapca -- chlapce,
dědečka -- dědečka, kadeřníčka -- kadeřnice,
napisovat -- napsat, dobrej -- dobrý, mam -- mám*

- je třeba opravit všechny slovní tvary a formy slov (tj. chyby hláskové, slovtvorné, tvarotvorné), které nelze považovat za spisovně české
- navíc je třeba opravit všechny tvary, které sice formálně za české považovat lze, ale intence pisatele byla zřetelně jiná (viz bod 4.3 **Nesprávný tvar incor**), např.



*on mi ho dal -- dal, jeli jsme na jich -- jih, teď
mužů říct -- můžu*

- neopravují se tvary, které nejsou chybné samy o sobě, ale ve vztahu k jinému slovu (tj. chyby ve shodě, valenci aj.)



NE

*včera přšeli; máme hezkých psa;
Marta hezká;*

- **Rovina 2**

- výsledkem anotace na R2 je gramaticky správná věta
- zde se provádějí opravy týkající se více slov, anotují se chyby plynoucí ze vztahů mezi (izolovaně) správnými tvary na R1
- chyba může být pouze v jednom slově, ale jejím důvodem je narušení vztahu se slovem jiným (shoda, valence, slovosled, dlouhá forma osobního zájmena v příklonkové pozici ap.)
- zde se opravují lexikální a některé stylové chyby

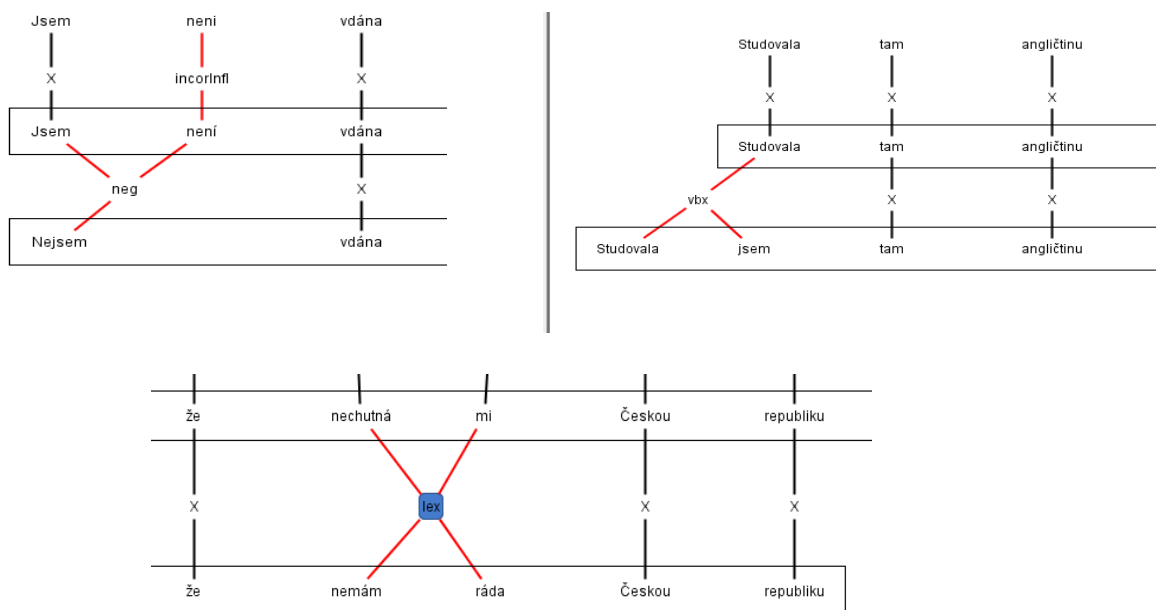


*pro maminka -- pro maminku; včera přšeli -- včera přšelo;
až mi to dal, půjdeme -- až mi to dá, půjdeme;
my by jsme rádi -- my bychom rádi; Petr bude dělá -- Petr
bude dělat atp.*

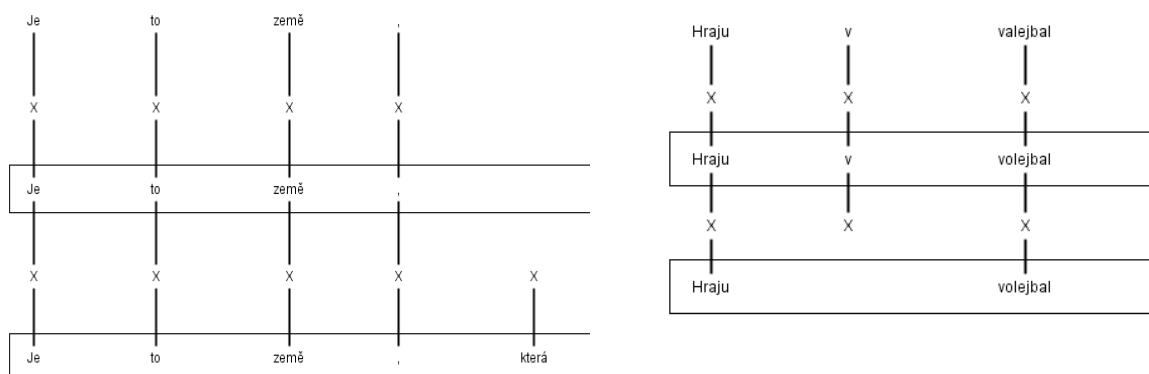
2.2 Vztahy mezi sousedními rovinami

Běžně platí vztah 1:1, tj. že jednomu slovu na jedné rovině odpovídá jedno slovo na rovině další. Ale při rozdělování, spojování slov, přidávání a vypouštění slov a také při opravách více slov najednou může vzniknout složitější vztah, tzv. **pavouk**:

- o jednomu slovu z jedné roviny může odpovídat více slov z roviny sousední,
- o více slovům na jedné rovině může odpovídat jedno slovo na rovině sousední,
- o více slovům na jedné rovině může odpovídat více slov na rovině sousední.



Na sousední rovině žádný ekvivalent být nemusí, v tom případě pak jedna nebo více hran spojující roviny končí v uzlu mezi rovinami. Takovým neúplným hranám říkáme "**pahýly**".



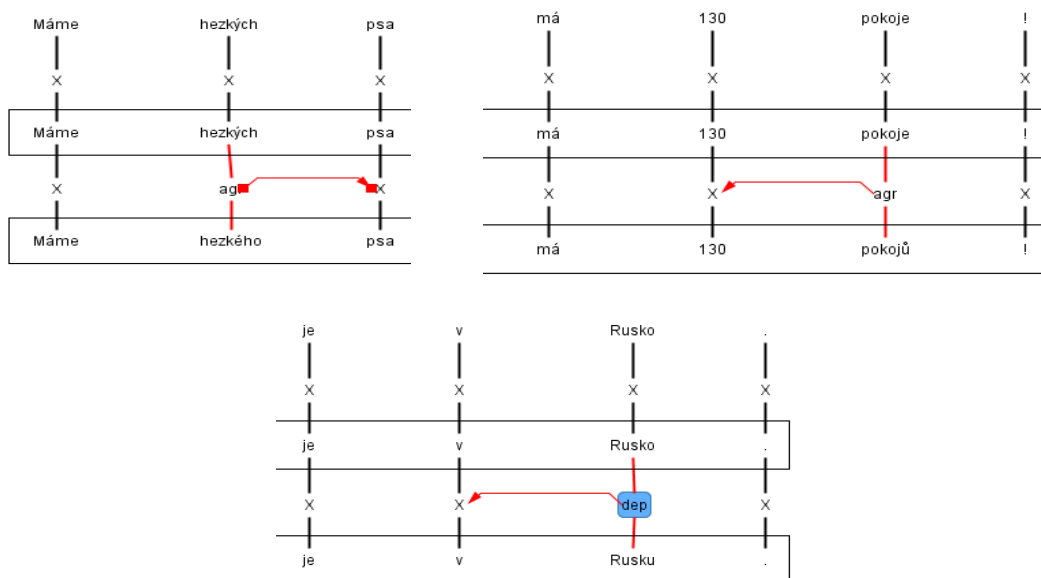
2.3 Odkazování

Odkazy slouží k objasnění důvodu opravy chyby, je-li jím vztah mezi opravovaným slovem a jiným slovem, kterým je správný tvar určen. Používají se u chyb z důvodů chybné shody, porušení valenčních požadavků, chybování v zájmeném odkazování, chyb ve složených přísudcích, u

analytických tvarů sloves, v negaci, u chyb v užití zvrtných výrazů, v odůvodněných případech i u chyb lexikálních, příp. dalších.

Odkazy vedeme vždy ve směru od opravovaného tvaru, a to i v případech jako *jen jedna holky zůstala doma* (narušení shody s odkazy od *holka* k *jedna* a *zůstala*).

Odkazovat lze pouze v případě, je-li na uzlu určen typ chyby.



hrana, tj. spojnice mezi formami/výrazy na jednotlivých rovinách, resp. mezi formou/výrazem a uzlem
uzel, tj. bod na hraně mezi rovinami, na kterém se vyznačuje kód chyby a ze kterého je veden odkaz
pavouk, tj. kombinované spojování/rozpojování hran se společným centrem v uzlu
odkaz, tj. objasnění důvodu opravy
pahýl, tj. hrana ukončená v uzlu (v případě odebrání či přidání výrazu)

2.4 Kódy chyb

Všechny kódy chyb, které jsou důležité pro ruční anotaci, obsahuje tabulka **Kódy chyb** v bodě 6.1.

2.5 Automatická anotace

Automatická anotace se spouští až po provedení anotace manuální.



Automaticky se určuje:

lemma, slovní druh a hodnoty morfologických kategorií slovních tvarů

- na R1 (mohou být víceznačné)
- na R2 (jednoznačné)

typ chyby nebo zpřesnění typu chyby na R1

- porovnáním chybného a opraveného řetězce

Perspektivně se počítá s klasifikací / specifikací chyb

- na obou rovinách
 - formální popis chyby, tj. výraz chybí/přebývá/je chybný/chybný sled;
- na rovině 1
 - chyby v diakritice, flektivních afixech, grafémice, palatalizaci, znělosti, metatezích apod.;
- na rovině 2
 - podrobnější určení typu morfosyntaktické chyby (chyba v pádu, rodě, čísle) související s chybou typu agr nebo dep na rovině 2, které lze provést porovnáním morfosyntaktických značek na rovině 1 a 2;
 - podrobnější určení typu chyby vbx (chyba v analytickém slovesném tvaru a složených predikátech).

3 ZÁKLADNÍ PRINCIPY EMENDACE / ANOTACE

3.1 Pravidla pro emendaci (prostou opravu textu)

- Opravu je třeba provést s ohledem na kontext a pravděpodobnou intenci autora.
- Opravu je třeba provést tak, aby byla co nejbližší originálu, tj. co nejméně měnit tvar slova, slovosled, počet slov ve větě, slovník ap. Konečným výsledkem opravy (tj. zápis věty na rovině R2) je gramaticky správná, nikoli stylisticky vybroušená věta.
- Je-li možné si vybrat z více možností emendace jednoho tvaru, volíme formu, která je nejbližší jeho původní podobě. Podrobněji viz část 4.1 **Obecné zásady**.
- Často je vhodné chybný tvar opravovat ve dvou fázích. Opravy na R1 by měly vést k formám, které jsou správné, ale co možná nejbližší původním chybným tvarům. Konečný tvar, správný i z hlediska kontextu, je pak uveden až na R2. Podrobněji viz také část 4.1 **Obecné zásady** (princip postupné opravy forem).



R0 – R1 – R2

*dvá další musea -- dva další musea -- dvě další musea
nekde jsem nosil džíny -- někde jsem nosil džíny -- někdy jsem nosil džíny*

R1 – R2

*Vladimír je lékařka -- Vladimír je lékař
díval se televize -- díval se na televizi
uběhlo už tři měsíc -- uběhly už tři měsíce
on je ne velký – on není velký*

*líbí se mým rodičům Finsko -- mým rodičům se líbí Finsko
a jsem šla pryč -- a šla jsem pryč
se učím česky dvě hodiny -- učím se česky dvě hodiny*

- V případě, že je konstrukce jazykově přijatelná (resp. gramaticky správná), neopravujeme ji.



NE

bydlíme ve svém domě

- neopravujeme na *bydlíme ve vlastním domě* apod.

celé léto jsem odpočívala s babičkou na chatě

- neopravujeme na *celé léto jsem trávila /se rekreovala s babičkou na chatě* apod.

používám češtinu v restauraci

- neopravujeme na *v restauraci mluvím česky*

Na skříni je váza. Rádio je taky na skříni.

*Česká republika je krásné místo ... Také v Čechách máte hezkou přírodu
líbí se mi moc Pražský hrad*

- neopravujeme slovosled



- Pokud nelze rozhodnout podle kontextu ani podle zásady minimálního zásahu, platí pravidlo zachovat podobu řídicího slova. Např.:

*má **hezkou** tužky -- má **hezké** tužky*

3.2 Pravidla pro anotaci

Pokud nelze přiřadit kód chyby jednoznačně, dáváme přednost „náročnější“ chybě, protože chybu „triviálnější“ lze odvodit automaticky, např. ve větě *kočka chytla myši* značujeme chybu ve valenci, ne ve flexi nebo ortografii.



Pro emendaci / anotaci na rovině 1 i 2 platí

- zásada gramatické korektnosti
- zásada minimální intervence
 - tj. minimální zásah do tvaru slova
- zásada respektování kontextu
- zásada zachování podoby řídicího slova
 - platí pouze v problémových případech, kdy nelze uplatnit výše uvedené principy

Rovina 1

- princip nejkratší smysluplné trasy (viz bod 4.2.4)
- princip postupné opravy forem

Rovina 2

- princip postupného odkazování
 - u předložkových vazeb
- principy opravy slovosledu
 - co nejméně zasahovat do původního slovosledu
 - co nejméně měnit počet slov ve větě
 - provedeme takový přesun, který je nejméně pracný

3.3 Technická poznámka

U většiny chybových kódů uvádíme dále v manuálu



oddíl **PŘÍKLADY** prezentující typy chyb, které řadíme k dané kategorii;

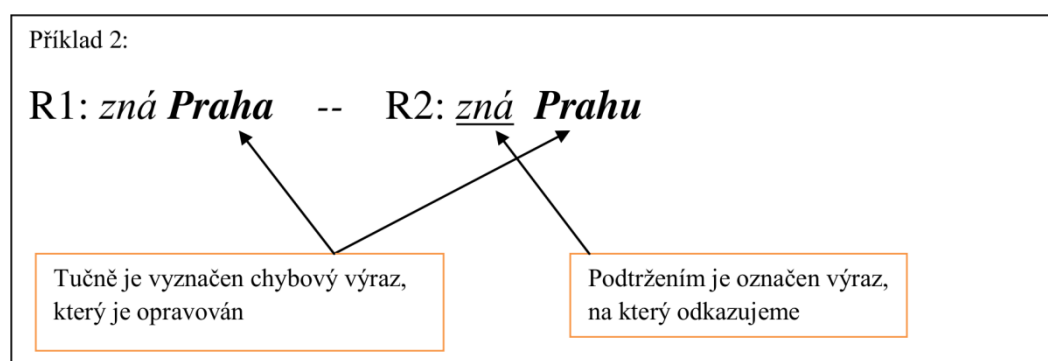
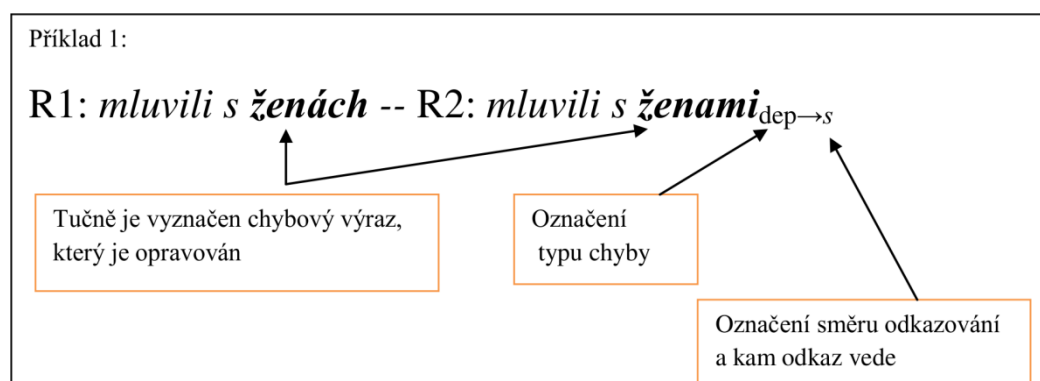


oddíl **POZOR!** zahrnující specifické a problémové jevy, jež lze řadit k danému chybovému typu;



oddíl **Pod tento typ chyby NEPATŘÍ**, tj. negativní příklady, jevy, které k danému typu chyby nezařazujeme.

Příklady uvádíme v těchto podobách:



4 CHYBY OPRAVOVANÉ NA ROVINĚ 1 (R1)

4.1 Jak naložit s přepisovými kódy

- vsuvky, přesuny, rozdělení, škrty
 - na R0 je uvedena už správná/přesunutá/opravená varianta

Některé z přepisových kódů zůstávají zachovány i v konvertovaném textu na R0. Podléhají specifickým pravidlům pro emendaci.

- varianty a autorské varianty (v přepisu kód <st>)
 - zůstávají na R0 viditelné pro anotátora, např. {Č/č}esko
 - emendují se na R1, tj. zvolí se správná varianta: Česko
- nečitelné řetězce bez hypotézy (v přepisu značka XXX)
 - značka XXX se chápe jako jeden celek, který se užívá pro písmeno, skupinu písmen i pro slovo; samostatné X nebo XX se v přepisu nepoužívá
 - značka XXX zůstává na všech rovinách bez emendace, bez hypotézy, bez kódu chyby
- nečitelné řetězce s hypotézou (v přepisu kód <ni>)
 - jde o nečitelné řetězce, kdy je ale přepisovač schopen na základě kontextu nebo jiných skutečností interpretovat nějakou možnost (*Nemůžu přijít, mám málo času*)<ni>
 - s hypotézou se pracuje stejně jako s běžným výrazem; údaj o tom, že jde o hypotézu, se projeví na všech rovinách automaticky
- předepsané, tj. neautorské části textu (v přepisu kód <dt>)
 - zůstávají zachovány a jsou viditelné na R0, např. {*Na obrázku je hezký pokoj.*}<dt>
 - neemendují se, na R1 i R2 zůstávají ve stejné podobě

4.2 Obecné zásady emendace na R1

Zde se provádějí opravy týkající se jednotlivých slov a slovních tvarů, např.



*kostích -- kostech, chlapca -- chlapce,
dědečka -- dědečka, kadeřníčka -- kadeřnice, napisovat --
napsat, dobrej -- dobrý, mam – mám atp.*

- Je třeba opravit všechny nesprávně utvořené a zapsané slovní tvary a formy slov (tj. chyby grafické, hláskové, slovotvorné, tvarotvorné). Opravujeme tedy slovní tvary, které nelze považovat za české.
- Navíc je třeba opravit všechny tvary, které sice formálně za české považovat lze, ale intence pisatele byla zřetelně jiná (viz bod 4.2.1 *Opravy typu dal – dál*), např.



on mi ho dál -- dal, jeli jsme na jich -- jih, teď mužů říct -- můžu, podlé mě -- podle

- neopravují se (nevypouštějí, neřadí jinak) tvary, které nejsou chybné samy o sobě, ale ve vztahu k jinému slovu (tj. chyby ve shodě, valenci aj.)



NE

včera přšeli; máme malých psa; Marta hezká; mají hezky velmi dům

Výsledkem opravy na R1 je posloupnost korektních českých slovních tvarů, resp. správných tvarů existujících českých slov. Výjimkou jsou opravy typu *dal – dál* (viz 4.2.1 *Opravy typu dal – dál*), a tzv. nově utvořená slova, resp. cizí a neidentifikovatelné řetězce, které neemendujeme (viz bod 4.4 *Nově utvořená slova fw*)



- Odlišujeme morfologii a morfosyntax. Morfosyntax se řeší na rovině 2 spolu s chybami ve shodě a valenci.
- Chyby ve složených tvarech slovesných se týkají více slov a řeší se na rovině 2 (*budu napsat*).
- Správná česká slova, která však nevyhovují kvůli významu, stylu nebo užití, se opravují na rovině 2.
- Interpunkce se opravuje na rovině 2.

4.2.1 Typy chyb opravovaných na R1 a jejich značkování

- pravopis
 - velká/malá písmena, y/i, ú/ů ap.
 - např. *analizovat, odpovídají, viprat, ůžech, dělá, říc, lef, jich, jih, o~~t~~počívám, besplatně, besplatnie, bezplatnye, czeský, tscheský;*
- inverze/metateze hlásek
 - např. *kraští/kratší, cerel/celer;*
- chyby v diakritice
 - chybějící (*česka*), nadbytečná (*manžéla*);

- chybná (neprovedená) palatalizace
 - např. *ledničke, článkech*;
- pohyblivé -e-
 - např. *domečkeku, Havela, tatíneka*;
- chyba ve flexi, resp. použití nekompatibilního afixu
 - chybná existující koncovka, resp. jiné deklinační či konjugační paradigma téhož slovního druhu (*čaju, spám*);
 - chybné použití zcela nekompatibilní koncovky (*rád pivuju*);
 - chybná neexistující koncovka (*zlatymi rybami*);
 - obecněčeská koncovka (*dobrej*);
 - chybná tvarotvorná předpona (*půjedu*); atd.
- chyby v hranicích slov
 - rozdělené slovo (přebývajících mezer) (*jsem*);
 - spojené slovo (chybějících mezer) (*byses*).



Pravopisné chyby NEZNAČKUJEME MANUÁLNĚ,

tj. nepřirazujeme jim chybový kód *incor*. Budou značkovány a blíže specifikovány automaticky.

Pro potřeby této anotace řadíme k pravopisným chybám následující:

- chyby v použití velkých/malých písmen (*praha*)
- použití *i/y* v kořenných morfémech domácích slov, výjimečně i v příponě (tj. problematika vyjmenovaných slov, např. *mamynka, obivatel*)
- záměna *ú/ů* (*dúm*)
- asimilace a neutralizace znělosti (např. *lef, odpověď, spívat, chutší*)
- záměna *jeXě* (*pjet*); *něXě* (*soukromně*); *de/ňe/te* X *dě/tě/ně* (*dědeček*)
- chyby v použití předpon *s-/z-* (*sjišťovat, sničit, ztýskat ...*)
- chyby v použití zdvojených hlásek
 - *-n/nn-* (*kamený, pana, denník*)
 - *-c/cc-* (*lecos*)
 - *-č/čč-* (*babičín*) atd.



Chyby v použití velkých písmen:

- na R1 opravujeme pouze chybné použití velkých / malých písmen u vlastních jmen, příp. vyjádření zvláštního vztahu k osobě nebo věci
praha -- Praha, v celené ulici -- v Celetné ulici
- chyby v neužití velkého písmena na začátku větného celku opravujeme až na R2 spolu s interpunkcí
šel ven. ona zůstala doma -- Šel ven. Ona zůstala doma.



Všechny ostatní chyby opravované na R1 z důvodu nesprávně utvořeného a zapsaného slovního tvaru **ZNAČKUJEME MANUÁLNĚ jako incor**.

Včetně

- interferenčních chyb (jako např. *czeský, tscheský, rólik, ko*)
- zjednodušování souhláskových skupin (např. *cera, kerý, řebík, děcký*)
- chyb v palatalizaci (např. *v ledničce, na soche*)
- chyb v metatezi hlásek (např. *kraští cesta*)

4.2.2 Zásada minimálního zásahu do tvaru slova

Někdy je možné nesprávné tvary opravit různými způsoby; v tom případě dáváme přednost takové opravě, která znamená nejmenší zásah do tvaru slova, tj. platí **zásada minimálního zásahu do tvaru slova**.



Např. ve větě *vidím úředníky* lze slovo *úředníky* opravit na *úředníci* nebo *úředníky*.
➤ preferujeme opravu na *úředníci*, a to kvůli stejné výslovnosti jako v původním tvaru.

Postupujeme takto:

1. Za vůbec nejmenší zásah považujeme opravu, která **nemění zvukovou podobu tvaru**, tedy opravu čistě pravopisnou: velká/malá písmena, y/i, ú/ů, palatalizace, znělost apod. Např. *mám rád praha -- Praha; znám nekdo -- někdo; potřebujou úředníky -- úředníci; odpovídají na otázky -- odpovídají; to je český film -- český; to je špatný odpověť -- odpověď* atd. Všechny takové chyby opravujeme rovnou na R1.



Případy typu *vidí mně // přišel ke mě* nejsou pravopisnou záležitostí, takové chyby se opravují až na rovině 2 jako chyby dep.

2. Pokud lze k existujícímu tvaru dojít pouhou **opravou diakritického znaménka**, či **záměny i za y** a naopak, opravíme slovo tímto způsobem na R1, i když se zvuková podoba slova změní (např. *krasný, deti, utery, menšy*), a případnou definitivní opravu provedeme na R2 s přiřazením příslušného typu chyby a event. odkazu (viz 4.2.3 **Zásada postupné emendace**).



RO: *me rodiče* -- R1: *mé rodiče* -- R2: *mí_{agr→rodiče} rodiče*

- nikoli hned na R1 na *mí*

3. Teprve jako další nastupují opravy, kterými se **mění jeden nebo více znaků**. Odebrání či doplnění znaku je větší zásah, než změna znaku.

4.2.3 Zásada postupné emendace

Opravou na R1 může vzniknout tvar, který je správný lokálně, nikoli však kontextově, a bude nutné jej dále opravit i na R2. Platí zde **zásada postupné emendace**.

Tento případ může nastat pouze v situaci, kdy emendujeme v souladu s body 1 a 2 výše, tj. je-li chyba způsobena nedostatkem pravopisným, záměnou *i/y*, či ne/užitím diakritického znaménka (viz 4.2.2 **Zásada minimálního zásahu do tvaru slova**).



RO: *dvá další musea* -- R1: *dva další musea* -- R2: *dvě další musea*

- nikoli hned na R1 na *dvě*

RO: *když jsem byl malý, někde jsem nosil džíny* -- R1: *někde jsem nosil džíny* -- R2: *někdy_{lex} jsem nosil džíny*

RO: *mluvili s ženach* -- R1: *mluvili s ženách* -- R2: *mluvili s ženami_{dep→s}*

- nikoli hned na R1 na *ženami*

RO: *vidí měste* -- R1: *vidí městě* -- R2: *vidí město_{dep→vidí}*

RO: *me dědečki neziví* -- R1: *mé dědečky neziví* -- R2: *mí_{agr→dědečki} dědečki_{agr→nežijí} nežijí_{lex}*



Nelze-li provést opravu dle zásady postupné emendace (viz výše), **uvedeme definitivní tvar, správný kontextuálně i formálně, už na R1**. V těchto případech nelze aplikovat postupnou opravu, protože nelze jednoznačně určit odpovídající formu na R1, aniž bychom výrazně neinterpretovali.



RO: *šel jsem na kavárn* -- R1: *šel jsem na kavárny* -- R2: *šel jsem do_{dep→šel} kavárny*

RO: *vidím mužů* -- R1: *vidím muže* -- R2: *vidím muže*

RO: *můj příšte jídlo je oběd* -- R1: *můj příští jídlo je oběd* -- R2: *moje_{agr→jídlo} příští jídlo je oběd*

RO: *šla jsem do museah* -- R1: *šla jsem do museí* -- R2: *šla jsem do museí*

RO: *v této loteryj* -- R1: *v této loterii* -- R2: *v této loterii*

RO: *hodne pribuzen jsou v německu* -- R1: *hodně příbuzných jsou v Německu* -- R2: *hodně příbuzných je_{agr→hodně} v Německu*

RO: *rodiči jsou uředniké* -- R1: *rodiči jsou úředníci* -- R2: *rodiče_{agr→jsou} jsou úředníci*

Takto, tedy uvedením definitivního tvaru na R1 a bez odkazu na R2, řešíme i případy, kdy lze předpokládat znalost funkce gramatického jevu (jde nejčastěji o chybu v kompatibilitě morfémů, tj. aplikaci nevhodného flektivního vzoru).

Problematiku tohoto typu ponecháváme k následné analýze.



RO: *byla jsem v Finske* -- R1: *byla jsem v Finsku* -- R2: *byla jsem ve_{lex} Finsku*

RO: *cestovala s rodičami* -- R1: *cestovala s rodiči* -- R2: *cestovala s rodiči*

RO: *byla jsem v museah* -- R1: *byla jsem v museích* -- R2: *byla jsem v museích*

RO: *jedu do domů a vařu večeri* -- R1: *jedu do domů a vařím večeri* -- R2: *jedu domů a vařím večeri*

RO: *nebo pracovají ve továrně* -- R1: *nebo pracují ve továrně* -- R2: *nebo pracují v továrně*

4.2.4 Zásada nejkratší smysluplné trasy

Máme-li více možností, jak na R1 dospět k existujícímu tvaru, zvolíme tu, která se více blíží správné formě. Platí zde **zásada nejkratší smysluplné trasy**.



R0: *to je skvělě nápad*

-- R1: *to je skvělé nápad*

- opravou na *to je skvěle nápad* bychom se vzdalovali od konečného znění

-- R2: *to je skvělý_{agr→nápad} nápad*

4.2.5 Opravy typu *dal* – *dál*

Tento typ chyby hodnotíme jako „překlep/přepis“ a opravujeme jej už na R1. Tzn. tuto chybu neklasifikujeme jako záměnu lexému.



Tuto chybu **MANUÁLNĚ NEZNAČKUJEME!**



Pravidla pro tento typ opravy

- mezi původním a opraveným tvarem není slovtvorný vztah, srov. *dal* - *dál*, *mužů* – *můžu* apod.
- jde primárně jen o záměnu jednoho grafému nebo písmenné spřežky, tedy včetně záměn *ou* – *u*, *ou* – *o*, *ch* – *h*, *cc* – *c*, *šš* - *š* apod.
- do této skupiny chyb patří i záměna *i/y* v kořenných morfémech domácích slov, výjimečně i v příponě (např. u slov jako *brzy*, *slepýš*)



PŘÍKLADY:

R0: *on mi to dál* -- R1: *on mi to dal*

R0: *jeli jsme na jich* -- R1: *jeli jsme na jih*

R0: *ted' mužů říct* -- R1: *ted' můžu říct*

R0: *podlé mě* – R1: *podle mě*

R0: *cedí na židli* -- R1: *sedí na židli*

R0: *bydlel jsem tam lok* -- R1: *bydlel jsem tam rok*

R0: *viděla jsem velký hlad* -- R1: *viděla jsem velký hrad*

R0: *puk je na podzim krásný* -- R1: *buk je na podzim krásný*

R0: *to je muka* -- R1: *to je mouka*

R0: *Radka mihá kávu* -- R1: *Radka míchá kávu*

R0: *Petr plašší holuby.* -- R1: *Petr plaší holuby.*

R0: *je to křečí domeček* -- R1: *je to křeččí domeček*

R0: *chce bít profesorem* -- R1: *chce být profesorem*

R0: *včely opilují květiny* -- R1: *včely opylují květiny*



Pod tento typ chyby NEPATŘÍ:

R0: ženy **přišli** -- R1: *přišli* -- R2: *přišly*

R0: vidím **páni** -- R1: *páni* -- R2: *pány*

R0: posílám **srdeční** pozdrav -- R1: *srdeční* -- R2: *srdečný*

R0: nelíbily se mi **německy** kluby -- R1: *německy* -- R2: *německé*

R0: tahle země má **malý** co nabídnout -- R1: *malý* -- R2: *málo*

R0: jede **rychlý** -- R1: *rychlý* -- R2: *rychle*

R0: dodržují **jejích** tradici -- R1: *jejích* -- R2: *jejich*

- jestliže je mezi výrazy slovotvorný vztah, chybu klasifikujeme jako chybu `agr/dep/lex` a opravujeme ji až na R2

4.2.6 Autorská univerbizace

V případě tzv. autorské univerbizace, tj. kdy z víceslovného vyjádření byla utvořena jednotka jednoslovná, česky nekorektní, opravujeme na R1 pomocí pavouka a tuto chybu **MANUÁLNĚ NEZNAČKUJEME**. Bude specifikována při automatické proceduře (jako specifický typ `incorOther`).



R0: kluk **drzoval** -- R1: *kluk byl drzý*

R0: rád **pivuju** -- R1: *rád piju pivo*

4.3 Nesprávný tvar `incor`

- tj. nesprávně utvořený výraz, který lze opravit

4.3.1 Nesprávná báze `incorBase`

- chyba ve slovotvorné struktuře a/nebo v derivační morfologii, tj. v kořeni a slovotvorných afixech. Jde o nekorektní formu lexému.



PŘÍKLADY:

R0: bydlím s **mátkou** -- R1: *bydlím s matkou*

R0: lidé jsou moc **měrný** -- R1: *lidé jsou moc mírný*

R0: kočka se **jmemuje** -- R1: *kočka se jmenuje*

R0: moje sestra je **kadeřníčka** -- R1: *moje sestra je kadeřnice*

R0: musíš to **posvětlit** -- R1: *musíš to vysvětlit*

R0: **libili** se mi -- R1: *líbili se mi*

R0: byla velká **tema** -- R1: *byla velká tma*

R0: mluvím **česka** -- R1: *mluvím česky*

R0: přišel **brze** -- R1: *přišel brzo*

- o do této kategorie řadíme také chyby v



- postfixech *kdokol -- kdokoli*
- infixech *malelinký -- malilinký*
- příp. interfixech *českoněmecký -- česko-německý*



Pod tento typ chyby NEPATŘÍ:

bojal se -- bál se
rád psu -- rád píšu
neje -- není

- příklady považujeme za chybu flexe (existují kořenné morfémy *boj-, ps-, atd.*).

Pokud vhodný, existující lexém nelze najít, ponechá se původní tvar s údajem, že jde o „vymyšlené“, obtížně emendovatelné slovo (viz bod 4.4.1 *Nově utvořené („české“) slovo fwFab*). V tomto případě tedy opravu neprovedeme.

Stejně tak při emendaci neopravujeme ani cizí slovo, příp. neidentifikovatelný řetězec, a ponecháme je na všech rovinách nezměněné; k těmto typům chyb uvádíme chybový kód *f_wNc* (viz bod 4.4.2 *Cizí / neidentifikovatelné slovo fwNc*).

4.3.2 Nesprávná flexe *incorInfl*

- o jde o opravu nekorektního tvaru slova, tj. chybu ve flexému.
- o chyba v tvarotvorných afixech
 - v tvarotvorném prefixu *půjedu*
 - v koncovce *spám*
 - v kmenotvorné příponě *dělel, kuřem*
 - příp. jiných tvarotvorných sufixech *nesna – nesena*
- o do této kategorie řadíme i chyby způsobené nevhodným použitím existujícího alomorfu daného lemmatu, např.



dír -- děr, stůlu -- stolu, Havela -- Havla, hoš -- hoch



PŘÍKLADY:

RO: *byla jsem v museah* -- R1: *byla jsem v museích*

RO: *v šestě třídě* -- R1: *v šesté třídě*

RO: *nosit sukné* -- R1: *nosit sukně*

RO: *už to stači* -- R1: *už to stačí*

RO: *dlouho jsem nepsála* -- R1: *dlouho jsem nepsala*

RO: *to jsou drahý prsteni* -- R1: *to jsou drahý prsteny*

RO: *se svou rodinou a svými dětaty* -- R1: *se svou rodinou a svými dětmi*

RO: *bojal se* -- R1: *bál se*

RO: *rád psu* -- R1: *rád píšu*

RO: *jetu na kole* -- R1: *jedu na kole*



Pod tento typ chyby NEPATŘÍ:

tema -- *tma*; *kadeřníka* -- *kadeřnice*; *buřka* -- *bouřka*; *měseci* -- *měsíce*
apod.

- příklady považujeme za chybu báze (užité báze formy se v paradigmatu daného lexému nevyskytují)

4.3.3 Auto: Ostatní nesprávné tvary incorOther



Jeden výraz může být chybný slovotvorně (*incorBase*) i tvarotvorně (*incorInfl*), tzn. mohou mu být přiřazeny oba chybové kódy, např.

RO: *tři měseci* -- R1: *tři měsíce*_{incorBase+incorInfl}

RO: *sestra nemá manželu* -- R1: *sestra nemá manželá*_{incorBase+incorInfl}

RO: *nebu nosit to triko* -- R1: *nebudu*_{incorBase+incorInfl} *nosit to triko*

V případě, že je u jednoho výrazu více chyb stejného typu, přiřazujeme příslušnou značku jen jednou, např.

RO: *půjedíme* -- R1: *pojedeme*_{incorInfl}

4.4 Nově utvořené / cizí / neidentifikovatelné slovo fw

Značkování této chybové kategorie reflektuje slovo tvorné strategie, resp. práci se slovo tvorným subsystémem jazyka. Nově utvořený výraz lze obvykle jen obtížně emendovat, příp. jej opravit nelze.

4.4.1 Nově utvořené („české“) slovo fwFab

- „autorský neologismus“
- obsahuje českou slovo tvornou strukturu
- student se pokusil vytvořit nové slovo/kmen/kořen ...



PŘÍKLADY:

R0: je tam hodne **jinaků** -- R1: je tam hodně **jinaků** -- R2: je tam hodně **jinaků**

R0: pokud nechces slyšet **smášky** -- R1: pokud nechceš slyšet **smášky** -- R2: pokud nechceš slyšet **smášky**

R0: anglicky jídlo je strašny všekno je **chlebasa** -- R1: anglický jídlo je strašný všechno je **chlebasa** -- R2: anglické jídlo je strašné, všechno je **chlebasa**



PŘÍKLADY:

R0: jím rad **eggs** -- R1: jím rád **eggs** -- R2: jím rád **eggs**

R0: jdu do **shopu** -- R1: jdu do **shopu** -- R2: jdu do **shopu**

R0: váza je na **Tische** -- R1: váza je na **Tische** -- R2: váza je na **Tische**

R0: byla v **hangu** -- R1: byla v **hangu** -- R2: byla v **hangu**

R0: jsem v **truong** -- R1: jsem v **truong** -- R2: jsem v **truong**

R0: **slichtování** názory o moderní obrázcích -- R1: názory o moderní obrázcích -- R2: **slichtování** názory o moderních obrázcích

R0: já **gymato** že -- R1: já **gymato** že -- R2: já **gymato**, že

4.4.3 Přítomnost flexe flex

- Tento kód uvádíme pouze jako doplňkovou klasifikaci pro chyby typu fwFab a fwNc. Poskytuje informaci, zda je v nově utvořeném, příp. cizím výrazu přítomna nějaká tvaro tvorná struktura. Zároveň neřešíme, zda je daný flexém „korektní“, či nikoli.
- Pokud nelze rozhodnout, zda se jedná o český flexém, či nikoli, je rozhodnutí na anotátorovi. Např. u výrazu v *truonge* není jasné, zda jde o koncovku Lsg., příp. jinou, nebo zda původní tvar slova je *truonge*, tudíž lze označit pouze jako fwNc, tj. výraz bez flexe .



PŘÍKLADY ZNAČKOVÁNÍ:

RO: *jím rad eggs fwNc*
 RO: *jím rad eggy fwNc+Flex*
 RO: *jím rad eggové fwNc+Flex*
 RO: *jdu do shopu fwNc+Flex*
 RO: *jdu do shopa fwNc+Flex*
 RO: *jdu do shop fwNc*
 RO: *musím to dopaintovat fwNc+Flex*

RO: *jdu do Geschäftu fwNc+Flex*
 RO: *jdu do Geschäft fwNc*
 RO: *váza je na Tischi fwNc+Flex*
 RO: *váza ie na Tische fwNc+Flex*

RO: *byla v hangu fwNc+Flex*
 RO: *byla v hang fwNc*
 RO: *jsem v truongu fwNc+Flex*
 RO: *jsem v truonge fwNc+Flex*

RO: *ja gymato fwNc*

RO: *je tam hodně jinaků fwFab+Flex*
 RO: *je tam hodně jinakové fwFab+Flex*
 RO: *pokud nechces slyšet smášky fwFab+Flex*

4.5 Chybná hranice slov *wbd*

Případy chybně odděleného prefixu, chybně rozděleného kompozita, příp. chyby v neoddělení předložky opravíme na R1 a uvedeme chybový kód *pre* (příp. *comp*) .

4.5.1 Prefix oddělený mezerou a předložka bez mezery *wbdPre*

- Tyto dva chybové typy (tj. prefix oddělený mezerou a předložka bez mezery) se vzájemně oddělí automaticky podle počtu hran nahoru nebo dolů.
- U oddělených prefixů řešíme na R1 primárně případy, kdy prefix a slovo stojí bezprostředně vedle sebe (prefix před nebo za slovem, ke kterému patří).
- Řadíme sem ale i řídké případy, kdy je prefix od slova oddělen nějakým dalším výrazem.



PŘÍKLADY:

RO: *při pravit* -- R1: *připravít*
 RO: *ne dělá* -- R1: *nedělá*
 RO: *Petr při jde* -- R1: *Petr přijde*
 RO: *já jsem to roz rychle vázal* -- R1: *já jsem to rychle rozvázal*
 RO: *Petr už jel od* -- R1: *Petr už odjel*
 RO: *potřebuju si to číst pře* -- R1: *potřebuju si to přečíst*
 RO: *on je ne velký* -- R1: *on je nevelký*
 - jde-li dle kontextu opravdu o lexém *nevelký*, nikoli o chybu *není velký*)

4.5.2 Neoprávněně rozdělená či spojená kompozita wbdComp

- takto označujeme
 - chybně oddělené lexikální morfémy
 - chyby ve spojení kompozit se spojovníkem
- opravujeme pomocí pavouka



PŘÍKLADY:

R0: *sádro karton* -- R1: *sádrokarton*

R0: *šesti násobný* -- R1: *šestinásobný*

R0: *anglo-americká literatura* -- R1: *angloamerická literatura*

R0: *českoněmecký slovník* -- R1: *česko-německý slovník*

4.5.3 Ostatní chybně rozdělená nebo spojená slova wbdOther

- chybně spojená, či rozdělená slova



PŘÍKLADY:

R0: *socialis mus* -- R1: *socialismus*

R0: *mochezký* -- R1: *moc hezký*

R0: *neměl čas, atak jsem šel* – R1: *neměl čas, a tak jsem šel*

R0: *přišel, kdykoli měl čas* – R1: *přišel, kdykoli měl čas*

5 CHYBY OPRAVOVANÉ NA ROVINĚ 2 (R2)

5.1 Narušení shody agr

- takto označujeme
 - chyby ve tvaru shodného přívlastku
 - chyby v rodě a čísle u finitního slovesného tvaru
 - chyby v rodě a čísle u jmenné části verbonominálního přísudku
 - chyby ve jmenných kategoriích u doplňku
- odkazování
 - chyby typu agr vyžadují odkaz na slovo, s nímž se má chybný tvar (příp. chybné tvary) shodovat
 - obvykle se odkazuje k podmětu
 - ✓ od chybného slovesa
 - ✓ od chybné jmenné části přísudku
 - nebo k řídicímu podstatnému jménu
 - ✓ od chybného shodného přívlastku
 - ✓ od chybného doplňku
- několikanásobný subjekt
 - pokud je subjekt několikanásobný (u souřadného spojení), odkazujeme na první spojku, na kterou v tomto spojení při čtení zleva doprava narazíme. Není-li v koordinační vazbě žádná spojka, odkazujeme na první čárku.

R1: Petr, Martin a Eva **přišel**. -- R2: Petr, Martin a Eva přišli.

- v případech typu otec s matkou odkazujeme na předložku s

R1: Adam s Evou **přišly**. -- R2: Adam s Evou **přišli**.

- v případech, kdy stojí podmět za přísudkem, a přísudek opravujeme na formu singuláru, odkazujeme na první člen koordinační vazby, např.:

R1: **Přišla** Timur a jeho parta -- R2: **Přišel** Timur a jeho parta.

- v těchto případech je možná i oprava na formu plurálu

R1: **Přišla** Timur a jeho parta -- R2: **Přišli** Timur a jeho parta..

- základní předpoklad: správně je člen, který určuje morfologické kategorie spojení
 - tj. podmět, který určuje tvar slovesa v přísudku;
 - nebo podstatné jméno, které určuje tvar adjektiva v přívlastku.
- pro problémové / nejasné / komplikovanější struktury (např. všechny studentka odešly) platí tato kritéria:

- a. princip minimální opravy (oprava by měla být co nejméně náročná)
- b. respektování kontextu
 - V návaznosti na tato kritéria emendace je v takových strukturách možné opravit i tvar řídicího substantiva podle tvaru shodného přívlastku nebo tvaru podmětu podle tvaru přísudku.
 - Do této kategorie patří tedy i chyby v pádu podmětu, protože při pádové synkrezí často nelze odlišit od narušení shody v čísle, např. *Hanky přišla (celá rozjařená)*.



PŘÍKLADY 1:

R1: *děvčata spaly* -- R2: *děvčata spala*

R1: *Petr vařím oběd.* -- R2: *Petr vaří oběd.*

R1: *Michal budu dělá* -- R2: *Michal bude_{agr→Michal} dělat_{vbx→bude}*

R1: *uběhlo už tři měsíce* -- R2: *uběhly už tři měsíce*

R1: *přišli pět studentů* -- R2: *přišlo pět studentů*

R1: *přišli pět studenti* -- R2: *přišlo_{agr→pět} pět studentů_{dep→pět}*

R1: *zítra budou pršet* -- R2: *zítra bude_{agr} pršet*

- uvedeme typ chyby bez odkazu, protože sloveso *pršet* je bezpodmětové

R1: *byli jsem překvapení* -- R2: *byli jsme překvapení*

- v případě dvoučlenných vět bez explicitně vyjádřeného podmětu: najdeme-li ve větě jiný tvar, s nímž se opravovaný má shodovat, odkážeme na něj

R1: *člověk objet celý svět* -- R2: *člověk objel_{agr→člověk} celý svět*

R1: *nerozumět tomu* -- R2: *nerozumím_{agr} tomu*

R1: *já jsme přišel* -- R2: *já jsem přišel*

R1: *přišel jsme* -- R2: *přišel jsem*

R1: *přišel jsme* -- R2: *přišli jsme*

- chyby shody ve složených slovesných tvarech opravujeme podle kontextu

R1: *hezké chlapci* -- R2: *hezcí chlapci*

R1: *máme hezkých psa* -- R2: *máme hezkého psa*

R1: *té věci patřili jim* -- R2: *ty_{agr→věci} věci patřily_{agr→věci} jim*

R1: *dům je velká* -- R2: *dům je velký*

R1: *Petr a Lucie jsou kamarád.* -- R2: *Petr a Lucie jsou kamarádi.*

R1: *Petra jsem nikdy neviděl opilý.* -- R2: *Petra jsem nikdy neviděl opilého.*

R1: *Čas jsou peníz.* -- R2: *Čas jsou peníze.*

- značíme jako chybu agr, protože jde o chybu v čísle, nikoli v pádu, a odkazujeme na sloveso jako zdroj shody)

R1: *Dítě je starosti.* -- R2: *Dítě je starost.*

- dtto



PŘÍKLADY 2:

- několikanásobný podmět

R1: Petr, Martin a Eva **přišel**. -- R2: Petr, Martin a Eva **přišli**.

R1: Otec a matka a děti **šla** do kina. -- R2: Otec a matka a děti **šli** do kina.

R1: Petr, Martin, Eva **přišel**. -- R2: Petr, Martin, Eva **přišli**.

R1: Adam s Evou **přišli**. -- R2: Adam s Evou **přišli**.

- několikanásobný větný člen (opravujeme a odkazujeme podle smyslu)

R1: Mám **krásná** růži a orchidej. -- R2: Mám **krásnou** růži a orchidej.

R1: Mám **krásná** růže a orchideje. -- R2: Mám **krásné** růže a orchideje.

R1: Mám **pichlavá** růže a orchideje. -- R2: Mám **pichlavé** růže a orchideje.

- chyby v doplňku

R1: pět chlapců **přiběhli určité** -- R2: pět chlapců **přiběhlo**_{agr→pět} **určicých**_{agr→chlapců}



- jako chyba shody se chápe i chyba ve shodném přívlastku vyjádřeném podstatným jménem, např.

R1: přišel s **chudák** maminkou -- R2: přišel s **chudákem** maminkou

R1: neviděli jste **slečna** Moniku? -- R2: neviděli jste **slečnu** Moniku?

R1: neviděli jste slečnu **Monika**? -- R2: neviděli jste **slečnu** Moniku?

R1: mluvil s paní **Nováková** -- R2: mluvil s paní **Novákovou**_{agr→paní}

- k agr řadíme také širše pojaté chyby v přechylování

R1: přišel pan **Nováková** -- R2: přišel pan Novák

R1: mluvil s panem **Nováková** -- R2: mluvil s panem **Novákem**

R1: mluvil s **pan Nováková** -- R2: mluvil s **panem**_{dep→s} **Novákem**_{agr→panem}



R1: věřím Petra Nováka -- R2: věřím **Petru**_{agr→Novák} **Novákovi**_{dep→věřím}

- u komplexních větných členů považujeme v případě osob za řídicí poslední člen skupiny

R1: viděl to v řece Vltavu -- R2: viděl to v řece Vltavě_{agr→řece}

- u neživotných komplexních větných členů považujeme za řídicí první člen

R1: k hoře Říp // do kavárny Slavia -- R2: k hoře Říp // do kavárny Slavia

- je-li v originálu použit nominativ jmenovací, ponecháváme a neopravujeme



R1: *dal si jeden čaje* -- R2: *dal si jeden čaj*_{agr→jeden}

- uvedený typ řadíme k chybě *agr* (chyba v určujícím členu - dle principu minimální opravy a respektování kontextu)
- nejedná se tedy o chybu *dep* (viz i oddíl 5.2, *vzal si dobrou ženy* apod.)



R1: *byla jsem v hezkého paláce* -- R2: *byla jsem v hezkém*_{sec+agr→paláci} **paláci**_{dep→v}

- jde o sekundární chybu (*sec*) související s chybou valenční (*dep*) u výrazu *paláce*, musíme označkovat jako *sec+agr*



Pod tento typ chyby NEPATŘÍ:

R1: *Eva je studentce.* -- R2: *Eva je studentka.*

R1: *Zločinec byl malá postava.* -- R2: *Zločinec byl malý*_{sec+agr→postavy} **postavy**_{dep→byl}

- je-li chyba ve jmenné části přísudku a jde jednoznačně o substantivum v chybném pádu, označujeme chybu jako *dep*

R1: *do polévky dám červené čočky* -- R2: *do polévky dám červenou*_{sec+agr→čočku} **čočku**_{use}

- *čočky* opravíme na *čočku* jako chybu *use* (viz oddíl 5.10), zde neuzuální užití plurálu
- oprava adjektiva je vyvolána změnou tvaru řídicího substantiva, proto se přidává kód pro sekundární

R1: *do polévky dám červená čoček* -- R2: *do polévky dám červenou*_{agr→čočku} **čočku**_{dep→dám}

- *čoček* opravíme na *čočku* jako chybu *dep*, opravu chybného tvaru adjektiva značujeme standardně jako *agr*

Srov. i

R1: *mají tam dobrou obsluha* – R2: *mají tam dobrou obsluhu*_{dep→mají}

R1: *mají tam dobrá obsluha* – R2: *mají tam dobrou*_{sec+agr→obsluhu} **obsluhu**_{dep→mají}

5.2 Chyba ve vyjádření syntaktické závislosti (jinak než shodou) dep

- Všechny chyby ve vyjádření syntaktické závislosti ve smyslu větné struktury, včetně chyb v nevalenčních doplňcích. Nepatří sem chyby shody, které řešíme výše v části 5.1.
- takto označujeme
 - chyby ve slovesné, substantivní, adjektivní, číslovkové, adverbialní rekcí
 - chyby ve vyjádření subjektu (např. záměna slovního druhu – v tomto případě nejde o chybu shody)
 - chyby v předložkové frázi
 - chyby v užití podřadicích spojek a ve tvaru slovesa po podřadicí spojce
 - chybné užití slovního druhu v příslovečném určení (např. záměna adverbia za adjektivum)
- odkazování
 - od opravovaného tvaru ke tvaru, který objasňuje důvod opravy, většinou od tvaru řízeného k tvaru řídícímu
 - pokud odkazujeme na sloveso, které je ve složeném tvaru, odkazujeme vždy na jeho plnovýznamovou část, např.

R1: bude se **ho** líbit **tu knihu** – R2: bude se **mu**_{dep→líbit} líbit **ta**_{sec+agr→kniha} **kniha**_{dep→líbit}

- potřebujeme-li odkázat na přísudek jmenný se sponou, odkazujeme na sponu, např.

R1: **kanceláře** je prázdná -- R2: **kancelář** je prázdná



PŘÍKLADY 1:

R1: věřím **učitelku** -- R2: věřím **učitelce**

R1: bude se líbit **tu knihu** -- R2: bude se líbit **ta kniha**

R1: káva bez **mléko** -- R2: káva bez **mléka**

R1: otázka **čas** -- R2: otázka **času**

R1: Petr čte **špatný** -- R2: Petr čte **špatně**

R1: dal obraz **dole** -- R2: dal obraz **dolů**

R1: bojím se jí **zavolám** -- R2: bojím se jí **zavolat**

R1: **studenta** odešel -- R2: **student** odešel

R1: **kanceláře** je prázdná -- R2: **kancelář** je prázdná

R1: doufali ve **vítězí** -- R2: doufali ve **vítězství**



R1: **Ondřeje** odejít v 5 hodin. -- R2: **Ondřej**_{dep→odešel} **odešel**_{agr→Ondřej} v 5 hodin.

- chyba ve tvaru subjektu i predikátu

R1: má schopnost **vytváří** -- R2: má schopnost **vytvářet**

• zde odkazujeme od chybného tvaru k řídícímu substantivu podobně:

R1: mám plán **pracuju** -- R2: mám plán **pracovat**_{dep→plán}



R1: Eva je **studentce**. -- R2: Eva je **studentka**.

R1: Zločinec byl **malá postava**. -- R2: Zločinec byl **malé**_{sec+agr→postavy}
postavy_{dep→byl}

- Je-li chyba ve jmenné části přísudku, považujeme ji za chybu ve shodě, kromě případů, kdy jde jednoznačně o substantivum v chybném pádu.

R1: maminka koupí -- R2: maminka **to**_{dep→koupí} koupí

- pokud chybí vyjádření předmětového komplementu, značujeme jako dep s odkazem na příslušný řídicí výraz

5.2.1 Způsoby odkazování u předložkových vazeb

5.2.1.1 Předložka správně, pád špatně (pád špatně i v kontextu s užitou předložkou)

- jeden odkaz – od substantiva v předložkovém pádu k předložce
- příklady: otázka pro maminka, údaj od pojišťovna, myslím na maminka, dal obraz na stěna, dal obraz ke stůl, dívá se na páni, dívá se na pán, pověsil to nad stole

- anotujeme takto:

R1: otázka pro **maminka** -- R2: otázka pro **maminku**_{dep→pro}

5.2.1.2 Předložka i pád špatně

- dva odkazy – od substantiva v předložkovém pádu k předložce (nikoli jako sec) a od předložky k řídicímu slovesu/substantivu/adjektivu
- příklady: myslím o maminka, díval se v televize, díval se od oknu, přišli v hřištěm, jdu do kamarád, email z kamarád

- anotujeme takto:

R1: myslím **o maminka** -- R2: myslím **na**_{dep→myslím} **maminku**_{dep→na}

5.2.1.3 Pád dobře v kontextu s předložkou, předložka špatně s ohledem na valenci řídicího slova

- týká se vícevazebných předložek, např. na, o, nad, pod, za ...
- dva odkazy – od substantiva v předložkovém pádu k prepozici (současně jako sec), a od předložky k řídicímu slovesu/substantivu/adjektivu
- příklady: otázka na mamince, přišli na hřišti, Díval se na tobě. Pavel sedí na kolo. Stůl je před křesla. Dívá se na pánech. Pověsil to nad stolem.

- anotujeme takto:

R1: přišli na **hřišti** -- R2: přišli **na**_{dep→přišli} **hřiště**_{dep+sec→na}

R1: stará se o **našich holčičkách** -- R2: stará se **o**_{dep→stará} **naše**_{sec+agr→holčičky} **holčičky**_{sec+dep→o}

- **Pravidlo postupného odkazování**

- problém s odkazy u chyb předložkových vazeb typu *myslím na tobě* řešíme vždy jako chybu ve volbě předložky, chybný pád podstatného jména je pak chybou sekundární, následnou. V tomto případě odkazujeme dvojmo – vedeme odkaz od "správné" předložky ke slovesu i od "chybné" podoby substantiva k předložce.

R1: *leží na postelemi* – R2: *leží na posteli*_{dep→na}

R1: *leží na postele* – R2: *leží na*_{dep→leží} *posteli*_{sec+dep→na}

R1: *dívám se na učitelce* – R2: *dívám se na*_{dep→dívám} *učitelku*_{sec+dep→na}

5.2.1.4 Předložka chybí, pád dobře

- pahýl pro předložku jako chyba *dep*, odkaz na řídicí slovo
- příklady: *díval se televizi*

- anotujeme takto:

R1: *díval se televizi* -- R2: *díval se na televizi*

5.2.1.5 Předložka chybí, pád špatně

- dva odkazy – od pahýlu pro předložku (*dep*) odkaz na řídicí slovo, od substantiva v předložkovém pádu odkaz na předložku
- příklady: *díval se televize, stará se pes*

- anotujeme takto:

R1: *díval se televize* -- R2: *díval se na*_{dep→díval} *televizi*_{dep→na}

5.2.1.6 Nadbytečně užitá předložka

- odstraníme předložku, odkážeme od pahýlu (*dep*) na řídicí slovo

A. tvar (pád) dobře

- anotujeme takto:

R1: *bojím se od psa* -- R2: *bojím se psa*

- odstraníme *od* a na pahýlu značujeme chybu *dep* s odkazem na *bojím*
- stejně jako u

R1: *bojím se před psa* -- R2: *bojím se [před_{dep→bojím}] psa*

B. tvar (pád) chybně i v kontextu s předložkou

- anotujeme takto:

R1: *bojím se před psovi* -- R2: *bojím se [před_{dep→bojím}] psa_{dep→bojím}*

- odstraníme *od* a na pahýlu značujeme chybu *dep* s odkazem na *bojím*
- opravíme *psa* a značujeme *dep* s odkazem na *bojím*

C. tvar (pád) správně v kontextu s předložkou, chybně ve vztahu k řídicímu výrazu

- anotujeme takto:

R1: *bojím se před psem* -- R2: *bojím se [před_{dep→bojím}] psa_{dep+sec→ bojím}*

- odstraníme *od* a na pahýlu značujeme chybu *dep* s odkazem na *bojím*
- opravíme *psa* a značujeme sekundární chybu (*sec*) *dep* s odkazem na *bojím*

5.2.2 Opravy vedlejších vět

- případy, kdy
 - chybí spojka
rozhodl se jdeme do kina
 - chybí spojka a sloveso je chybně (infinitiv, chybný určitý tvar)
řekl mu nedělat to
slyšel Petr přišla pozdě
 - spojka je správně, ale sloveso je chybném tvaru (infinitiv, chybný určitý tvar)
řekl mi, že napsat dopis
Chtěl, abych přijdu?
 - je chybná spojka
půjdu tam, proto chci ho vidět
 - je chybná spojka a chybný tvar slovesa
chodím tam, kdy mít čas
- značujeme
 1. chybějící a chybně užitá spojky jako *dep* s odkazem na predikát v hlavní větě
 2. chybné tvary sloves standardně podle typu chyby jako *agr*, *vb_x* nebo *use*



PŘÍKLADY:

R1: *rozhodl se jdeme do kina* -- R2: *rozhodl se, že jdeme do kina*

R1: *řekl mu Petr to neznat* – R2: *řekl mu, že_{dep→řekl} Petr to nezná_{agr→Petr}*

R1: *slyšel Petr přišla pozdě* – R2: *slyšel, že_{dep→slyšel} Petr přišel_{agr→Petr} pozdě*

R1: *řekl mi, že napsat dopis* -- R2: *řekl mi, že napíše_{agr} dopis*

- není-li kam jednoznačně odkázat, pouze přiřadíme značku a neodkazujeme

R1: *půjdu tam, proto chce ho vidět* -- R2: *půjdu tam, protože_{dep→půjdu} ho chci_{agr→chci} vidět*

- v tomto případě lze odkázat i na shodový slovesný určitý tvar (tj. na vše, co vykazuje shodové kategorie)

R1: *chodím tam, kdy mít čas* – R2: *chodím tam, když_{dep→chodím} mám_{agr→chodím} čas*

- dtto

R1: *přišel jsem Petr odcházíš* – R2: *přišel jsem, když_{dep→přišel} Petr odešel_{agr→Petr + use}*



KONDICIONÁLOVÉ SPOJKY *aby* / *kdyby*

R1: *řekl mu nedělat to* – R2: *řekl mu, aby_{dep→řekl} to nedělal_{vbx→aby}*

- chybí spojka, sloveso je infinitiv
- neopravujeme pomocí pavouka

R1: *řekl Petr přišla včas* – R2: *řekl, aby_{dep→řekl} Petr přišel_{agr→Petr} včas*

- chybí spojka, sloveso v chybném určitém tvaru (chyba v rodu)

R1: *Chtěl, abych přijdu?* – R2: *chtěl, abych přišel_{vbx→abych}*

- sloveso v chybném určitém tvaru (ale ve správné osobě a čísle, nejedná se tedy o chybu *agr*!)

R1: *Chtěl, abych přijde?* – R2: *chtěl, abych přišel_{vbx→abych}*

- řešíme stejně jako předcházející příklad!
- sloveso v chybném určitém tvaru (chybná osoba a číslo, přesto neznačujeme jako chybu *agr*)

R1: *řekl mi aby to nedělat* – R2: *řekl mi, abych_{agr} to nedělal_{vbx→abych}*

- spojka chybně, sloveso infinitiv
- u kondicionálové spojky odkazujeme na odpovídající subjekt, příp. na výraz, který nese shodové prvky, pokud to nelze, ponecháváme bez odkazu

R1: *řekl mu aby to neděláš* – R2: *řekl mu, abys_{agr} to nedělal_{vbx→abys}*

- řešíme stejně jako předcházející příklad!



VZTAŽNÉ VĚTY VEDLEJŠÍ

A. chybí vztažný výraz (vztažné zájmeno / vztažné adverbium)

R1: *přivedla muže Jana neznala* -- R2: *přivedla muže, kterého_{dep→neznala} Jana neznala*

- značujeme chybějící větný člen (*dep*)

R1: *pamatuješ si tu noc Petr se vyboural?* -- R2: *pamatuješ si tu noc, kdy se Petr vyboural?*

- značujeme chybějící větný člen (*dep*)

B. chybí vztažný výraz a slovesný tvar je chybný

R1: *vzal kabát ležet na židli* -- R2: *vzal kabát, který_{dep→ležet} ležel_{agr→který} na židli*

C. vztažný tvar je chybný

R1: *vzal kabát, kterou ležel na židli* -- R2: *vzal kabát, který_{dep→ležet} ležel na židli*

R1: *přivedla muže, kterému Jana neznala* -- R2: *přivedla muže, kterého_{dep→neznala} Jana neznala*

R1: *to je kniha, o kterou jsme mluvili* -- R2: *to je kniha, o které_{dep→o} jsme mluvili*

R1: *to je kniha, pro kterou jsme mluvili* -- R2: *to je kniha, o které_{dep→mluvili} jsme mluvili*



R1: manželka je v domácnosti starat se o děti

-- R2: manželka je v domácnosti a stará_{agr→manželka} se o děti_{dep→o}

-- R2: manželka je v domácnosti, stará_{agr→manželka} se o děti_{dep→o}

-- R2: manželka je v domácnosti, aby_{dep→je} se starala_{vbx→aby} o děti_{dep→o}

- opravujeme podle kontextu, volíme nejbližší a nejpřiléhavější opravu



Pod tento typ chyby NEPATŘÍ:

- ✓ R1: stará se o **malých** holčičky – R2: stará se o **malé**_{agr→holčičky} holčičky
 - nejde o chybu dep, ale je to chyba agr
- ✓ viděl jsem velkou **postele**, bojí se cizích **muže**, dal si jeden **čaje**, vzal si dobrou **ženy**
 - je-li v nesourodé jmenné frázi tvar atributu z kontextuálních důvodů signifikantní a u tvaru řídicího substantiva je z několika možných interpretací (např. postele – Gsg, Lsg, Npl, Apl) alespoň jedna správná s ohledem na vazbu/valenci řídicího členu (zde Apl), tak ji za ni a priori považujeme a chybu substantiva klasifikujeme jako chybu ve shodě agr. Např.

R1: viděl jsem velkou **postele** – R2: viděl jsem velkou **postel**_{agr→velkou}

- ✓ R1: knihu čte už **dlouze** – R2: knihu čte už **dlouho**_{lex}
 - opravujeme jako lex, protože nejde o narušení syntaktického vztahu
- ✓ R1: nemá rád **červenou třešni** -- R2: nemá rád **červené třešně**_{use}
 - opravujeme jako chybu úzu (use), tvar adjektiva opravujeme jako sekundární chybu
 - čocky opravíme na čocku jako use a červený opravíme na červenou jako agr a sec. Jde-li o neuzuální užití singuláru či plurálu, značíme tento typ chyby jako use.
- ✓ R1: až mi to **dal**, budu spokojen -- R2: až mi to **dá**_{use}, budu spokojen
- ✓ R1: **až** jsem udělal zkoušku, šel jsem na pivo -- R2: **když**_{lex} jsem udělal zkoušku, šel jsem na pivo

5.3 Chyba v analytickém slovesném tvaru a složených predikátech vbx

- značujeme ručně jen jako obecnou kategorii, bez dalšího manuálního rozlišení. Automaticky se následně tento typ chyby rozdělí dle jednotlivých kategorií.
- patří sem
 - chyby v analytické morfologii sloves (tj. chyb v préteritu, opisném futuru, kondicionálu, reflexivních slovesech, pasívu)
 - chyby v konstrukcích s modálními a fázovými slovesy
 - chyby ve verbonominálních predikátech
 - chyby v konstrukcích rezultativních
- nepatří sem
 - chyby ve shodě finitního tvaru nebo příčestí s podmětem, např. *Petr budu dělat; Petr přišla sám; kde jste byli, děvčata?*
(na rozdíl od chyb, kdy je třeba místo infinitivu příčestí, kdy je pomocné sloveso navíc, kdy chybí nebo je jiné, než má být, např. *udělat jsem úkol*)



PŘÍKLADY:

R1: *budu napsat* -- R2: *napišu*

R1: *budu napsat* -- R2: *budu psát*

R1: *musíš přijdeš* -- R2: *musíš přijít*

R1: *začal pracuje* -- R2: *začal pracovat*

R1: *Petr bude dělá* -- R2: *Petr bude dělat*

R1: *Petr budu dělá* -- R2: *Petr bude_{agr→Petr} dělat_{vbx→bude}*

R1: *učit jsem se* -- R2: *učil_{vbx→jsem} jsem se*

R1: *učím jsem si* -- R2: *učil_{vbx→jsem} jsem se_{reflx→učit}*

R1: *kluci jsou běhali* -- R2: *kluci běhali*

- chyba vbx na pahýlu *jsou*, s odkazem na *běhali* (automaticky se doplní chyba odd)

R1: *Petr je dělal večeri* -- R2: *Petr dělal večeri*

- chyba vbx na pahýlu *je*, s odkazem na *dělal* (automaticky se doplní chyba odd)

R1: *guláš bylo chutná mi* -- R2: *guláš mi chutnal_{vbx}*

- opravujeme pomocí pavouka



V případě, že chybí spona (nebo jiná část složeného predikátu, tj. auxiliár), značíme jako `vbx`, pomocí pahýlu a odkazu (nikoli pavouk). Automaticky se následně doplní klasifikace `miss` a specifikace `cvf` nebo `vnp` ap.



Pod tento typ chyby NEPATŘÍ:

- ✓ R1: Michal **budu dělat** -- R2: Michal **bude**_{agr→Michal} *dělat*
R1: zítra **budou pršet** -- R2: zítra **bude** *pršet*
- R1: Petr **si budu učit** -- R2: Petr **se**_{reflx→učit} **bude**_{agr→Petr} *učit*
 - k `vbx` nepatří chyby ve shodě finitního tvaru nebo přičestí s podmětem
 - příklady chybné shody s podmětem, které řadíme k `agr`
- ✓ R1: **jseš přišel** – R2: **jsi přišel**
 - značujeme na R1 jako `incorInfl+stylColl`
- ✓ R1: **udělalo se hezké** -- R2: **udělalo se hezky**
 - složené přísudky výše uvedeného typu se značují jako volné syntagma (`dep`)

5.4 Chyba ve zvrtném výrazu `rflx`

- při opravě odkazujeme obvykle na sloveso, k němuž zvrtný výraz patří. Kvůli odkazu označujeme tuto chybu manuálně jako `rflx`
- patří sem
 - chyby v reflexivech tantum
 - odkazujeme na sloveso
R1: *smála si* -- R2: *smála se*_{rflx}
 - chyby v reflexivním pasivu, deagentivu
 - odkazujeme na sloveso
R1: *dnes si neučí* -- R2: *dnes se*_{rflx} *neučí*



PŘÍKLADY:

R1: *troufl se přijít* -- R2: *troufl si přijít*

R1: *narodila jsem v Petrohradu* -- R2: *narodila jsem se v Petrohradu*

R1: *dívá na televizi* -- R2: *dívá se na televizi*

R1: *Pavel si raduje* -- R2: *Pavel se raduje*

R1: *tancovali jsme se celý večer* -- R2: *tancovali [se_{rflx}] jsme celý večer*

R1: *auta si vyrábějí* -- R2: *auta se vyrábějí*



Jako *rflx* **neoznačujeme manuálně** případy, kdy je chybu ve zvrtném výrazu možné hodnotit jako chybu jiného typu, tedy *ref*, *dep* nebo *agr*. Označení *rflx* se pak doplní automaticky podle toho, že se v chybné nebo opravené verzi objevuje zvrtný výraz.

Jde především o

- chyby ve valenci (*dep*)
 - tyto chyby odlišíme od ostatních případů chybného (ne)užití výrazů *se* a *si* náhradou za zájmena *sebe* nebo *sobě*, vznikne-li náhradou gramaticky správná věta, jde o tuto chybu ve valenci
R1: *umyla si celá* -- R2: *umyla se_{dep→umyla} celá*
R1: *mýt se ruce* -- R2: *mýt si_{dep→mýt} ruce*
- také valenční chyby v členu navíc nebo v chybějícím členu:
R1: *hned po obědě maminka myje se nádobí* -- R2: *hned po obědě maminka myje [se_{dep→myje}] nádobí*
R1: *ráno čistím zuby a umyju* -- R2: *ráno si_{dep→čistím} čistím zuby a umyju se_{dep→umyju}*
- chyby v referenci zájmen
R1: *Tomáš stojí před jeho dům* -- R2: *Tomáš stojí před svým_{ref→Tomáš} domem_{dep→před}*
(z kontextu vyplývá, že dům je Tomášův)
- chyby ve shodě
R1: *Petr stojí před její dům* -- R2: *Petr stojí před svým_{ref→Petr + agr→dům} domem_{dep→před}*
(z kontextu vyplývá, že dům je Petrův)



Chybu v užití zájmena 1. nebo 2. osoby místo zájmena zvrtného označujeme jako chybu *rflx*, užití zájmena 3. osoby místo zvrtného však označujeme jako chybu *ref*.

R1: *dívám se na moje auto* -- R2: *dívám se na svoje_{rflx} auto*

R1: *dívám se na mě* -- R2: *dívám se na sebe_{rflx}*

R1: *dívá se na jeho dům* -- R2: *dívá se na svůj_{ref} dům*

R1: *konečně si rozhodl oholit* -- R2: *konečně se rozhodl oholit*

- haplogonii značíme jako chybu s dvojím odkazem

5.5 Chyba v negaci neg

- včetně záporové shody
- chyby tohoto typu mohou obsahovat odkaz na slovo indikující záporovou shodu



PŘÍKLADY:

R0: *půjdu neráno* -- R1: *půjdu ne ráno* -- R2: *nepůjdu ráno*

- z R0 na R1 i z R1 na R2 oprava pomocí pavouka

R0: *on ne velký* -- R1: *on ne velký* -- R2: *on není velký*

R0: *on je ne velký* -- R1: *on je ne velký* -- R2: *on není velký*

R1: *půjdu ne do školy* -- R2: *nepůjdu do školy*

R1: *mám žádný čas* -- R2: *nemám_{neg→žádný} žádný čas*

R1: *žádný to ví* -- R2: *žádný to neví*

R1: *máma ani táta kouří* -- R2: *máma ani táta nekouří_{neg→ani}*

5.6 Přebytečné / chybějící výraz odd / miss

- vždy emendujeme, ale značky odd a miss se přiřazují automaticky
- interpunkční znaménka neznačujeme, tj. jsou bez pahýlu
- přebývající či chybějící výrazy značujeme podle kontextu jako
 - dep (u subjektů, objektů, atributů a adverbialní),
 - vbx (u částí slovesného tvaru)
 - lex apod.
- přidáváme příslušný odkaz, je-li to možné
 - nepřičítáme-li příslušnou značku, není třeba dělat pahýl ani odkaz



PŘÍKLADY:

R1: *Petr vstává v 6 hodin. On nesnídá, protože on nemá hlad.*

-- R2: *Petr vstává v 6 hodin. [on_{dep→nesnídá}] Nesnídá, protože [on_{dep→nemá}] nemá hlad.*

R1: *Petra já ho viděl* -- R2: *Petra [já_{dep→jsem}] [ho_{dep→nesnídá}] jsem_{vbx→viděl} viděl*

R1: *To je moje kamarádka Iva. Kamarádce koupil jsem jí zmrzlinu.*

-- R2: *To je moje kamarádka Iva. [kamarádce_{dep→koupil}] Koupil jsem jí zmrzlinu.*

R1: *já jsem je viděl je* -- R2: *já jsem je viděl*

R1: *Znal jsem.* -- R2: *Znal jsem ji_{dep→znal}*

R1: *Iva hezká* -- R2: *Iva je_{vbx→hezká} hezká*

R1: *chlapci dělali spali* -- R2: *chlapci spali*

- zde bez pahýlu a bez značky

5.7 Auto: Chybný slovosled wo

- slovosled je vhodné opravovat až po všech ostatních opravách
- chyby ve slovosledu emendujeme (= opravujeme), ale nepřihazujeme jim žádnou značku, tyto opravy budou označeny automaticky
- slovosledné opravy se zaměřují primárně na gramaticky správný slovosled, tj. na opravu chybného postavení příklonek, předložek, kvantifikátorů, adverbíí, shodných a neshodných přívlastků, doplňků
 - pořadí klitik:
 1. spojka *-li*;
 2. pomocné sloveso v minulém čase (*jsem, jsi (-s), jsme, jste..*) a v podmiňovacím způsobu (*bych, bys, by, bychom, byste..*);
 3. krátké tvary zvrátných osobních zájmen (*si, se*);
 4. krátké tvary osobních zájmen v dativu (*mi, ti, mu...*);
 5. krátké tvary osobních zájmen v akuzativu (*mě, tě, ho; to, tu ...*).

Učil jsem se. Zítra bych se učil. Snažil jsem se mu to vysvětlit. Uvařil jsem si to. ...

- Opravu slovosledu provedeme tak, aby byla vždy co nejbližší originálu. Tj. řídíme se principy:
 - co nejméně zasahovat do původního slovosledu
 - co nejméně měnit počet slov ve větě
- Pokud slovosled vykazuje rysy kontextového aktualizovaného užití, opravu ve slovosledu neprovádíme. V následujícím dialogu např. nepřesouváme *bydlím*:

*Ty pracuješ v Praze? To každý den dojíždíš? – Já v Praze už pět let **bydlím!***



PŘÍKLADY:

R1: *líbím se moc Pražský hrad ...* -- R2: *líbí se **mi** moc Pražský hrad*

- slovosled tedy neopravujeme na: *moc se mi líbí Pražský hrad*

R1: *mají hezký velmi dům* -- R2: *mají **velmi hezký** dům.*

- podobně: *hezký tento dům // **malí dva** kluci*

R1: *je to rychlé auto **hodně*** -- R2: *je to **hodně** rychlé auto*

R1: *koupil jsem **auto nové*** -- R2: *koupil jsem **nové auto***

R1: *bál jsem se, že **jí se** nebude líbit Praha* -- R2: *bál jsem se, že **se jí** nebude líbit Praha*

R1: *to by velmi vadilo **mi*** -- R2: *to by **mi** velmi vadilo*



V anotačním programu je možné přesouvat celé skupiny slov.
Primárně přesouváme závislé členy větné struktury!

5.8 Chyba v lexiku a frazeologii lex

- chyby ve významu, stylu, kolokabilitě, užití slova nebo konstrukcí
- neexistující nebo chybně psaná jednotlivá slova se opravují na rovině 1
- chybu v užití frazému identifikujeme automaticky
- je-li však ve frazému jiná chyba, např. v pravopise nebo ve shodě, opravujeme ji jako v ostatních volných spojeních



Chyby typu *lex* opravujeme umírněně, minimalizujeme stylové opravy!
Např. následující příklady není třeba opravovat:

tady nemyslím - není třeba opravit na *tím nemyslím*
mléčné potraviny - není třeba opravit na *mléčné výrobky*



PŘÍKLADY:

R1: ..., **proto** to by mi velmi vadilo. -- R2: ..., **protože**_{lex} to by mi velmi vadilo

R1: bydlíme v domě **i** nebydlíme v bytě -- R2: bydlíme v domě **a**_{lex} nebydlíme v bytě

R1: potřebuju koupit **housenky** -- R2: potřebuju koupit **housky**

R1: ta knížka, **což** jsi mi dal, je hezká -- R2: ta knížka, **kterou** jsi mi dal, je hezká

R1: **až** jsem udělal zkoušku, šel jsem na pivo -- R2: **když** jsem udělal zkoušku, šel jsem na pivo

R1: Vladimír je **lékařka** -- R2: Vladimír je **lékař**

R1: Přesto dopadlo to velmi **přírodně**. -- R2: Přesto to dopadlo velmi **přirozeně**_{lex}.

R1: **když** se učím, nemůžu mluvit -- R2: **i když** se učím

- opravu provedeme pomocí pavouka

R1: jsem **ruská** -- R2: jsem **Ruska**_{lex}

R1: jsem **Vietnam** -- R2: jsem **Vietnamec**_{lex}

R1: čtu tu knihu **dlouze** -- R2: čtu tu knihu **dlouho**_{lex}

R1: ve které žijou **světové** známé -- R2: **světově** známé osobnosti

5.9 Chyba ve významovém užití gramatické kategorie use

- chyby tohoto typu jsou způsobeny nevhodným užitím hodnoty nějaké morfologické kategorie. Tento typ chyby neřadíme k chybám lexikálním!

5.9.1 Neuzuální užití singuláru či plurálu



PŘÍKLADY:

R1: nemá rád **třešni** -- R2: nemá rád **třešně**

R1: bylo tam hodně **lidstev** -- R2: bylo tam hodně **lidstva**

R1: do polévky dáme **červené čočky** -- R2: do polévky dáme **červenou čočku**

R1: Ráda se dívá na **film**. -- R2: Ráda se dívá na **filmy**.

5.9.2 Chybné použití času



PŘÍKLADY:

R1: včera **bude sněžit** -- R2: včera **sněžilo**

- opravíme pomocí pavouka

R1: až mi to **dal**, budu spokojen -- R2: až mi to **dá_{lex}**, budu spokojen

5.9.3 Chybné užití vidu



PŘÍKLADY:

R1: celý večer **uvařím** večeři -- R2: celý večer **budu vařit** večeři

R1: **oslaví** narozeniny -- R2: **bude slavít_{use}** narozeniny

- opravíme pomocí pavouka

R1: Češi **schovají** všechno -- R2: Češi **uchovávají_{use + lex}** všechno

- značujeme obě chyby

5.9.4 Nevhodně užitý komparativ nebo superlativ



PŘÍKLADY:

R1: pošta je **nejvíc blízko** -- R2: pošta je **nejblíž**

R1: ten dům je **nejvíc hezký** -- R2: ten dům je **nejhezčí**

R1: moje tričko je **nejvíc nejhezčí** -- R2: moje tričko je **nejhezčí**

R1: mnohem **dobrý** -- R2: mnohem **lepší**

R1: Petr je **velký** než Pavel -- R2: Petr je **větší** než Pavel

5.9.5 Vokalizace předložek



PŘÍKLADY:

R1: **v** vodě-- R2: **ve** vodě

R1: **ve** lese-- R2: **v** lese

5.10 Chyba v zájmeném odkazu ref

- Chyby tohoto typu mohou obsahovat odkaz na slovo, s nímž se chybný tvar (chybné tvary) má shodovat, a to v rodě, čísle, osobě (nikoli v pádě!).
- Odkaz může vést i ke slovu, jež se nachází v jiné (nejčastěji bezprostředně předchozí) větě, než kde se nachází chybný tvar.
- Tento kód chyby se může vyskytovat v kombinaci s jinou chybou.



PŘÍKLADY:

R1: paní, **jenž** jsem se ptal, mi to ho moc neřekl. -- R2: paní, **jíž** jsem se zeptal, ...

R1: dal jsem to jemu i **jejího** bratrovi -- R2: dal jsem to jemu i **jeho**_{ref→jemu + agr→bratrovi} bratrovi

- dvě chyby na jednom slově

R1: Tomáš stojí před **jeho dům** -- R2: Tomáš stojí před **svým**_{ref→Tomáš} **domem**_{dep→před}

- z kontextu vyplývá, že dům je Tomášův

R1: Petr stojí před **její dům** -- R2: Petr stojí před **svým**_{ref→Petr + agr→dům} **domem**_{dep→před}

- z kontextu vyplývá, že dům je Petrův

-

R1: Lidé podle **té** žijou -- R2: Lidé podle **ní** žijou_{ref + lex}



Pod tento typ chyby NEPATŘÍ:

- užití zájmena první nebo druhé osoby místo zájmena zvrátého, takové chyby označujeme jen jako rflx
 R1: *dívám se na **moje** auto* -- R2: *dívám se na **svoje** auto*
 R1: *dívám se na **mě*** -- R2: *dívám se na **sebe**_{rflx}*
 R1: *dívám se na **mně*** -- R2: *dívám se na **sebe**_{dep} →_{na + rflx} → *dívám**

5.11 Sekundární chyba sec

- chyba, která vznikla teprve v důsledku jiné opravy, tzv. "zavlečená chyba"
- kód přidáváme jako doplňkový ke kódu, který signalizuje charakter chyby (např. sec+agr)



PŘÍKLADY:

R1: *dívá se na **americkém filmu*** – R2: *dívá se _{dep} → *dívá* **americký**_{sec+agr} → *film* + **film**_{dep} → *na**

- výraz *americkém* je v originálním textu správně ve vztahu k výrazu *filmu* (lokál sg). Protože však musíme opravit výraz *filmu* na *film* v důsledku valenční vazby slovesa *dívat se*, je nutné provést i opravu řízeného adjektiva. V tom případě je oprava *americkém* na *americký* sekundární, tj. vynucená opravou jiného tvaru.

R1: *stará se o **našich holčičkách*** -- R2: *stará se _{dep} → *stará* **naše**_{sec+agr} → *holčičky* **holčičky**_{sec+dep} → *o**

R1: *bojím se **před psem*** -- R2: *bojím se [**před**_{dep} → *bojím*] **psa**_{dep+sec} → *bojím**

*mají tam **dobrá obsluha*** -- *mají tam **dobrou**_{agr} → *obsluhu* **obsluhu**_{dep} → *mají**

5.12 Stylově nevhodný výraz styl

- kód přidáváme obvykle jako doplňkový ke kódu, který signalizuje charakter chyby (např. lex+stylOther)
- nedostatky v rejstříku; obecněčeské, knižní, nářeční výrazy

5.12.1 Obecněčeský výraz stylColl

- doplňková klasifikace napříč rovinami
- pro výrazy potenciálně kolokviální (tj. pro ty, o kterých anotátor uvažuje, že by mohly být kolokviální)



PŘÍKLADY:

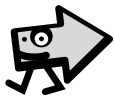
R0: *to byl hezkej den* -- R1: *to byl hezký_{incorInfl+stylColl} den*
 R1: *viděli jsme hezký holky* -- R2: *viděli jsme hezké_{agr→holky+stylColl} holky*
 R0: *oslavím narozeniny* – R1: *oslavím_{incorinfl+stylColl} narozeniny*



Aspekt kolokviality sledujeme primárně na morfému, příp. lexému. Nikoli na úseku textu.

5.12.2 Jiný stylově nevhodný výraz *stylOther*

- o také jako doplňková klasifikace chyby
- o knižní, nářeční, slangový výraz ap.
- o příp. i hyperkorektnost: *s kamarádami*



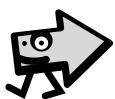
PŘÍKLADY:

R0: *pláči nad vejdělkem* -- R1: *pláči nad výdělkem_{incorStem+stylOther}* -- R2: *pláču_{stylOther} nad výdělkem*
 R0: *vona bula učitelka* -- R1: *ona_{stylColl} byla_{stylOther} učitelka*
 R1: *rozbil se mi hadr* -- R2: *rozbil se mi harddisk*

- *snuřasně inkn 7ex)*

5.13 Rozvrácená konstrukce *disr*

- o Tato chyba se může týkat celých odstavců, vět, částí vět.
- o V případě zcela nesrozumitelné konstrukce je možné
 - vůbec neemendovat na rovině 2: konstrukci převedeme na rovinu 1, zasažené části spojíme do polovičního pavouka a přiřadíme kód *disr*, na meziuzlu končí hrana pahýlem
 - interpretovat. V tom případě se jedná o naprostou interpretaci. Uděláme celého pavouka, přiřadíme na meziuzlu kód *DISR* a na R2 uvedeme svou vlastní interpretaci.



PŘÍKLADY:

- R0: *mám rád na skvělé hráju na kitaru* – R1: *mám rád skvělé hráju na kytar* – R2:
- R0: *kratka jakost vyborné ženy, které musíš dat pozornost* – R1: *krátká jakost výborné ženy, které musíš dát pozornost* – R2:



Pozor, abychom tuto chybu nenadužívali. Vždy se nejprve pokoušíme o emendaci a anotaci.

5.14 Problémová chyba *problem*

- Tímto kódem označujeme případy, kdy si určením chyby nejsme jisti. Vždy jako doplňující kód.

6 Kódy chyb

Vlastnosti uzlů pro popis chyb mezi rovinami 0-1 a 1-2 (v tabulce 1 a 2); -1 znamená "libovolná hodnota"

typ chyby	podtyp chyby	A/M	popis chyby	počet hran nahoru: min;max	počet hran dolů: min;max	počet odkazů: min;max	rovina
incor			nesprávný tvar (obecná kategorie pro neanotované případy oprav)				
	incorInfl	M	nesprávná flexe	1;1	1;1	0;0	R1
	incorBase	M	nesprávný kmen	1;1	1;1	0;0	R1
	incorOther	A	ostatní nesprávné tvary cizí / nově utvořené / neidentifikovatelné slovo	1;-1	1;-1	0;0	R1
fw			nově utvořené „české“ slovo	1;1	1;1	0;0	R1
	fwFab	M	nově utvořené „české“ slovo	1;1	1;1	0;0	R1
	fwNc	M	slovo z jiného jazyka	1;1	1;1	0;0	R1
flex		M	flexe u výrazů fw	1;1	1;1	0;0	R1
wbd			chybná hranice slov				
	wbdPre	M	prefix oddělený mezerou a předložka bez mezery	1;2	1;2	0;0	R1
	wbdComp	M	neoprávněně rozdělená/spojená kompozita	2;-1	1;1	0;0	R1
	wbdOther	M	jiné chyby týkající se hranice slov	1;-1	1;-1	0;0	R1
agr		M	narušení shody	1;1	1;1	0;-1	R2
dep		M	chyba ve vyjádření syntaktické závislosti	0;1	0;1	0;-1	R2
ref		M	chyba v zájmeném odkazu	1;1	1;1	0;1	R2
vbx		M	chyba v analytickém slovesném tvaru a složeném přísudku	1;-1	1;-1	0;1	R2
	cvf	A	chyba v analytickém slovesném tvaru	1;-1	1;-1	0;1	R2
	mod	A	chyba v konstrukci s modálním nebo fázovým slovesem	1;-1	1;-1	0;1	R2
	vnp	A	chyba ve sponově-jmenném přísudku (vč. pas. a rez.)	1;-1	1;-1	0;1	R2
rflx		M	chyba v reflexivním výrazu	0;-1	0;-1	0;-1	R2
neg		M	chyba v negaci	1;-1	1;-1	0;1	R2
odd		A	nadbytečné slovo	1;1	0;0	0;0	R2
miss		A	chybějící slovo	0;0	1;1	0;0	R2
wo		A	chybný slovosled	1;-1	1;-1	0;0	R2
lex		M	chyba v lexiku a frazeologii	0;-1	0;-1	0;1	R2
use		M	chyba ve významovém užití gramatické kategorie	1;-1	1;-1	0;1	R2
sec		M	sekundární, "zavlečená" chyba	1;-1	1;-1	0;-1	R2
styl			obecněčeský, knižní, nářeční tvar				

	stylColl	M	potenciální obecněčeský tvar knižní, nářeční, slangový ap.	0;-1	0;-1	0;-1	R1,R2
	stylOther	M	tvar/výraz	0;-1	0;-1	0;-1	R1,R2
disr		M	rozvrácená konstrukce	-1;-1	0;-1	0;0	R2
problem		M	problémová chyba	0;-1	0;-1	0;-1	R1,R2