

Jednoduchá anotace strukturálních metadat pro češtinu

-pracovní verze 5-

2.5.2005

Jáchym Kolář

jachym@kky.zcu.cz

Pozn.: Pokud se chcete odkázat na tato pravidla ve vědecké práci, prosím citujte následující publikaci:

KOLÁŘ, J., ŠVEC, J., STRASSEL, S., WALKER, C., KOZLÍKOVÁ, D., PSUTKA, J.: Czech Spontaneous Speech Corpus with Structural Metadata. –In: Proceedings of the 9th European Conference on Speech Communication and Technology, INTERSPEECH '2005, Lisboa, Portugal, 2005, pp. 1165-1168.

Obsah

1.	Úvod.....	2
1.1	Notace.....	2
2.	Výplně a neplynulosti.....	4
2.1	Jednotlivé typy výplní	4
2.1.1	Vyplněné pauzy	4
2.1.2	Discourse markers (DM, Označovače předělů v promluvě)	4
2.1.3	Discourse response (DM jako odpovědi v promluvě).....	6
2.1.4	Explicitní editační výrazy (EET).....	6
2.1.5	Odbočky a vsuvky (A/P).....	6
2.2	Více výplní za sebou	8
2.3	Výplně (DM) na hranicích syntaktických jednotek	8
3.	Editační neplynulosti (korekce obsahu promluvy).....	9
3.1	Anotace editačních neplynulostí	9
3.2	Příklady editačních neplynulostí	10
3.2.1	Repetice.....	10
3.2.2	Revize.....	10
3.2.3	Opakované začátky (restarty).....	10
3.3	Složité neplynulosti	11
3.3.1	Delreg obsahující více neplynulostí	11
3.3.2	Neplynulé korekce.....	11
3.3.3	Složité neplynulosti a opakované začátky.....	11
3.3.4	Více neplynulostí v jedné promluvě.....	12
3.4	Neplynulosti a spojky	12
3.5	Neplynulosti a výplně.....	12
4.	SU.....	13
4.1	Pravidla pro přepis typických jevů.....	14
4.1.1	Bezslovesné SU.....	14
4.1.2	Volně připojené větné členy.....	14
4.1.3	Presumptivní otázky (tag-questions).....	14

4.1.4	Otázky, které řečník směřuje sám na sebe a řečnické otázky	15
4.1.5	Souřadně spojené otázky	15
4.1.6	Oznamovací věty s tázací intonací	15
4.1.7	Vložené otázky	15
4.1.8	Fráze fungující jak kompletní výpověď	16
4.1.9	Upřesňování	16
4.1.10	Nekompletní SU	16
4.1.11	Souhlasy a nesouhlasy	17
4.1.12	Idiomatická spojení	17
4.1.13	Přímá a nepřímá řeč (citace) - Z!	18
4.1.14	Osamostatnělé vedlejší věty	18
4.1.15	Parcelace	18
4.1.16	Infinitivní konstrukce	19
4.1.17	Apozice (komplexní větné členy) – Z!	19
4.1.18	Juxtapozice vět – „uvozovací věty“	19
4.1.19	Anakolut	20
4.1.20	Turny, kterým chybí začátek	20
4.2	SU a souvětne konstrukce	20
4.2.1	Rozlišení souřadného a podřadného spojení	22
4.2.2	Odlišení souvětí od několikanásobného přísudku	22
4.2.3	Podřadná souvětí	23
4.2.4	Vedlejší věty připojené pomocí „takže“	23
4.2.5	Rozhodnutí o prioritě /, či /&	24
4.3	Odlišení konce SU(/.) a koordinace (/&) – Z!	24
4.3.1	Vedlejší věta závislá na více větách hlavních	24
4.3.2	Elipsy	24
4.3.3	Stažení souvětých konstrukcí	25
4.3.4	Spojení spojkou neboť	25
4.3.5	Souřadné spojení pomocí výrazů složených ze dvou částí	25
4.3.6	Souřadně připojená věta má zamlčený podmět a obsahuje pouze přísudek – ZRUŠENO!	25
4.4	Neplynulosti měnící strukturu SU	25
5.	Prozodicky vyjádřené SU	25
6.	Slova vyřčená řečníkem na pozadí - Backchannels	26
7.	Ostatní	26

1. Úvod

Cílem anotace je označit v již existujících doslovných přepisech zvukových nahrávek řeči několik druhů tzv. *strukturálních metadat*. Jde zejména o označení hranic syntaktických celků (SU) a anotaci tzv. *neplynulostí*. Konečným výsledkem by měly být prepisy, které jsou dobře čitelné pro člověka a mají vhodný tvar i pro další automatické zpracování. Jde zejména o označení hranic syntakticko-sémantických celků a anotaci tzv. *neplynulostí*. Prepisy obsahující anotaci metadat pak mohou být očištěny, tj. např. každý celek může být uveden na jedné řádce a mohou z něj být vypuštěna slova, která nenesou žádnou informaci či je řečník nechtěl pronést (vzápětí je opravil jinými slovy).

1.1 Notace

V tomto textu se bude používat následující notace, význam jednotlivých značek bude vysvětlen později.

[text] * - text je Delreg, * je bod přerušení

text – text je korekce (oprava)

{text} – text je odbočka/vsuvka

te~ - slovní fragment

text – text je výplň

∅ - v místě, kde by mohl být, není žádná hranice SU (pouze pro ilustrační účely v tomto návodu)

/ - konec oznamovací SU

// - prozodicky vyznačený konec oznamovací SU

/, - hranice klauze

/? - konec tázací SU

//? - prozodicky vyznačený konec tázací SU

/& - koordinace

/~ - konec nekompletní SU přerušené jiným řečníkem

/- - konec nekompletní SU svévolně přerušené řečníkem

Přímo v prepisech se ještě objevují další speciální symboly:

%text% - cizí slovo se zvláštní (nečeskou) výslovností

text - přeroknutí

text~ - nedokončené přeroknutí

~ - nedokončená věta nebo věta bez začátku (podle kontextu)

Mezi hranatými závorkami jsou zobrazeny tzv. *neřečové události*. Mezi ně patří neartikulované zvuky a různé šумы, které jsou buď způsobeny řečníkem nebo pocházejí z okolního prostředí. Jednotlivé události jsou vypsány v následující tabulce.

Neřečová událost	Značka
Šum nespecifikovaný	[NOISE]
Neartikulované zvuky, vyplněné pauzy „souhláskového typu“ - „mm“, „vv“	[MM][HESITATION]
Neartikulované zvuky, vyplněné pauzy – uh, er, ee atd.	[EE][HESITATION]
Souhlasné hm nebo m-hm	[HM]
Nesouhlasné m-m	[MM]
Hlasité dýchání včetně nádechu	[LOUD BREATH]
Smích	[LAUGHTER]
Kašel	[COUGH]
Mlasknutí (rty nebo jazykem)	[CLICK]
Ticho (relativně dlouhá pauza)	[SILENCE]

Pozornost je třeba věnovat zejména [HESITATION], [HM] a [MM], ostatní neřečové události nemají pro naše účely význam a je možné je při prepisech zcela ignorovat.

V tomto textu se často zmiňuje termín prozodie, pod kterým rozumíme to, jak řečník daná slova pronáší, tj. větnou intonaci (výšku hlasu), dělání pauz, časování (délka samohlásek) a hlasitost.

2. Výplně a neplynulosti

Výplň se dá charakterizovat jako úsek řeči, artikulovaný či neartikulovaný, který je možné z přepisu vymazat, aniž by to mělo za následek ztrátu informace. V tomto projektu se anotují následující výplně:

- vyplněné pauzy (VP),
- discourse markers (DM, „označovače v promluvě“),
- explicitní editační termíny (EET),
- odbočky/vsuvky (Asides/Parentheticals, A/P).

Žádný z těchto typů výplní nemění obsah tvrzení (promluvy), do které jsou vloženy. Anotace výplní zahrnuje identifikaci slov, které slouží jako výplň a určení o jaký typ výplně se jedná.

2.1 Jednotlivé typy výplní

2.1.1 *Vyplněné pauzy*

Vyplněné pauzy jsou neartikulované zvuky vydávané mluvčím při váhání. Nenesou žádný význam (jiný než indikaci váhání). V češtině to typicky bývají zvuky podobné „é“ nebo „mmm“, každý řečník má obvykle svůj charakteristický styl.

Za vyplněné pauzy se nepovažují zvuky, které mají v rozhovoru komunikační funkci (Např. odpovědi na otázky se zavřenými ústy - Hm, M-mm)

Vyplněné pauzy mohou také sloužit jako EET – pokud EET obsahuje pouze VP, je anotován pouze jako VP.

VP jsou již v transkripcích předznačeny jako *[HESITATION]* (váhání). Pokud je VP vyslovena jako *é* (nebo je její součástí zvuk podobný samohlásce), nechá se v přepisu beze změny. Pokud spíše připomíná *mmm*, *vvv*, *ddd* či jiný sled zvuků podobných souhláskám (bez samohlásek), je nutno tuto VP přeznačit na *[MM-HESITATION]*.

Pokud někde v nahrávce uslyšíte VP, která není v přepisu označena, pak ji včetně odpovídajícího typu doplňte.

2.1.2 *Discourse markers (DM, Označovače předělů v promluvě)*

„Discourse marker“ (DM) je slovo nebo fráze, která označuje strukturu mluveného projevu. DM posluchači signalizují řečníkům záměr označit předěl v promluvě. DM jsou neodmyslitelnou aktivní součástí mluveného projevu. Signalizují takové události, jako jsou změna řečníka, přebírání, udržování nebo vzdávání se kontroly v dialogu či začátek nového tématu. Nás budou zajímat pouze ty DM, které mají funkci výplně (tedy jakési „slovní vaty“) a jejichž vymazáním neztratíme žádnou informaci z rozhovoru.

Je prakticky nemožné sestavit pro daný jazyk vyčerpávající seznam DM, protože pokrývají velkou škálu konverzačních funkcí a je těžké je přesně definovat. DM také podléhají mnoha

dialektickým variacím a jsou také závislé na individuálních návycích každého řečníka. Příkladem DM v češtině mohou být:

tak, tak jako tak, dobře, dobrá, to víte, podívejte se, víte, víte co, no, tedy, takže, a nyní, prostě, vlastně, v podstatě, jaksi, mimochodem, jako, teda

Jako DM se budou označovat jen ty výrazy, které mají za účel strukturovat promluvu nebo držet řečníkovi slovo a nenesou samostatný význam (obsah). Rozlišení může být někdy velmi obtížné. V takovýchto nejasných případech se ponechají daná slova neoznačena.

Kromě kontextu může být vodítkem k odlišení DM i prozódie, zejména přítomnost pauzy.

Stejně výrazy se mohou objevovat jako DM i jako „obsahová slova“, je třeba vždy posoudit danou situaci z hlediska kontextu celého dialogu. Např. u výrazů jako *takže* je třeba vzít potaz, zda řečník chce vyjádřit vztah k předchozí větě nebo zda-li chce označit předěl v rozhovoru (např. změnit téma). (V následujících příkladech jsou DM vyznačeny tučně):

*Přesně tak se to stalo. x **Tak** tady je každá rada drahá.*

*Běžel jako vítr. x To **jako** není nic neobvyklého. → ve 2. případě je „jako“ parazitní výraz, pomocí kterého si řečník jen drží slovo.*

***Tak** myslím si, že rozhodující vliv má...*

*To **jako** není nic neobvyklého.*

No musím říct, že sem o tom nikdy neslyšel.

*A **jaksi** Čechům je také vlastní silná antimilitaristická tradice.*

***Takže** [pauza] Američané na jednu stranu stahují své vojenské jednotky z Evropy ...*

***Tak jako** mě tam zaujalo to, že pod tím slovem demokracie si představuje ty volby v parlamentu.*

***Čili** my se tady asi těžko shodneme, protože on mluví o něčem jiném.*

V konstrukcích typu

To znamená vašim kritériem není stabilita /.

se výrazy „*to jest*“ a „*to znamená*“ jako DM neoznačuje (viz pravidlo o juxtapozicích níže).

Problematické může být slovo *takže*. Toto slovo označíme jako DM, pokud zjevně nemá funkci spojky; obvykle tato DM provází delší pauza, případně intonační změna.

Dalším problematickým slovem je *tak*. Jako DM ho nebudeme označovat v příp. kdy jím uvozuje vedlejší větu např. *časovou* nebo *způsobovou*, protože přestože jeho použití není striktně vzato vhodné, již prakticky zlidovělo.

Až přijdu večer domů /, tak to udělám /.

Čím víc se snaží /, tak tím se mu lépe daří /.

Pokud vyhraje /, tak koupím bednu šampaňského /.

Když tenhle zápas srovnám s pondělní dohrávkou /, tak musím říct /, že na skoro neregulérním hřišti toho v utkání domácí Opavy se Spartou moc ke koukání nebylo.

Naopak na začátku věty (SU) má *tak* prakticky vždy funkci DM.

Tak to mi teda povězte /.

2.1.3 *Discourse response (DM jako odpovědi v promluvě)*

V mnoha případech jsou DM používány nejen ke strukturování promluvy, ale také k vyjádření odpovědi druhému řečníkovi. Tyto speciální případy DM se nazývají „Discourse responses“ – DR. DR mohou být jednotlivá slova nebo fráze a může se jich vyskytovat několik za sebou. Mluvčí obvykle používají DR k odpovědi na to, co právě řekla druhá osoba a zároveň takto signalizují snahu převzít kontrolu nad rozhovorem.

V případě přímých odpovědí na otázky se však o DR nejedná. Přímé odpovědi se označují jako oznamovací SU. DR vždy fungují také jako DM, zatímco přímé odpovědi jsou pouze odpovědi na otázky.

Ve velké většině případu se DR vyskytují na začátku turnu.

A: Myslím, že tohle zaměstnání vůbec nezvládá. Kdyby bylo po mém, tak ho propustím.

B: Podívejte se já nemůžu být tak tvrdý.

A: Řekl bych, že ho ho dříve nebo později chytí.

B: No já si nejsem tak jistý.

2.1.4 *Explicitní editační výrazy (EET)*

EET jsou výplně, které se objevují v neplýnulostech. EET obsahuje frázi, kterou řečník jednoznačně identifikuje neplýnulost. Obvykle se jedná o krátké fráze, jako jsou např. *myslím, pardon, promiňte, spíše, nebo* atd. Klasické EET se vyskytují poměrně řídko, většinou mají ve spontánní promluvě funkci EET vyplněné pauzy, ty se ovšem jako EET neanotují.

*Myslím, že se jedná o [stihom~] * nebo spíše extremní poruchu osobnosti /.*

*Tohle je [nejoddiskutov~] * pardon neoddiskutovatelný fakt /.*

EET se může vyskytovat i až za korekcí.

*Tohle je [koherentní] * konzistentní lépe řečeno stanovisko /.*

EET se mohou v rámci neplýnulosti vyskytovat i až za korekcí, ne jen před ní.

2.1.5 *Odbočky a vsuvky (A/P)*

Striktně vzato nejsou odbočky a vsuvky výplně, ale pro jednoduchost a podobnost jejich anotace je budeme mezi výplně zařazovat.

Odbočky se objevují, když mluvčí pronese krátký komentář k jinému tématu a pak se vrátí opět k původnímu tématu. Odbočka může být také adresována někomu, kdo není účastníkem aktuální konverzace. Odbočky jsou obvykle provázeny prozodickými příznaky - zvýšení/snížení výšky hlasu, zvýšení tempa či pauza.

A jeho rodiče byli {ted' nevím jak to říct} takový samorosti /.

Vsuvky jsou podobné odbočkám a také přerušují plynulý myšlenkový tok v promluvě, ale obvykle se týkají hlavního tématu. Vsuvky nejsou do věty mluvnický začleněny – nejsou v syntaktickém vztahu s jinou klauzí. Pro účely anotace se nebude mezi odbočkami a vsuvkami rozlišovat. Používání A/P by mělo být opatrné, nesmí se zaměňovat s např. vztažný větami atd. Za odbočky/vsuvky se nepovažují obvyklá slova nebo krátké fráze, které se vkládají do vět jako *myslím, prosím, řekněme, například, dejme tomu, tuším, nemýlím-li se, upřímně řečeno, řekl bych* atd. Tyto vložené výrazy nejsou obvykle provázeny stejnými prozodickými charakteristikami jako odbočky/vsuvky.

Jedná se o {používám jeho výraz} naprostý nedostatek komunikace /.

To se občas stane /, že člověk {nemělo by se to tedy stávat příliš často} udělá chybu /.

Podejte mi Ø prosím Ø ten kabát.

Tohle Ø myslím Ø nepůjde.

Uvnitř A/P je zakázáno používat jakékoliv SU symboly. Naopak se i uvnitř A/P označují výplně (DM, EET, Delregy, vyplněné pauzy atd.)

Pro ilustraci a upřesnění uvádím seznam spojení, která se budou brát jako lexikalizované (pokleslé) vsuvky a **nebudou** se nijak označovat. Tento seznam berte jako ilustrační, nikoliv vyčerpávající. Tato spojení jsou vždy dlouhá 1-2 slova, 3 a víceslovná spojení již označíme jako A/P.

tak řečeno
řekl bych
bych myslel
řekněme
chcete-li
myslím
myslím si
když už
prosím
například
dejme tomu
tuším
upřímně řečeno
nemýlím-li se
jednoduše řečeno

Pokud jste u daného spoje na pochybách, rozhodněte se dle případných prozodických změn.

Z jednoslovných a dvouslovných vsuvek se jako A/P anotují vsuvky typu *{cituji}, {omlouvám se}, {uvеду příklad}* - tedy takové, které do hlavní promluvy (myšlenky) evidentně nepatří.

V rámci odboček a vsuvek se mohou také vyskytnout vyplněné pauzy či DM, které se také označí obvyklým způsobem. Naopak hranice SU se uvnitř odboček/vsuvek neoznačují!

Odbočky a vsuvky se v těchto prepisech nebudou rozlišovat, anotují se jako ta samá událost.

Pokud je vsuvka vložena mezi dvě klauze, je třeba před ní vložit ještě odpovídající SU symbol.

Hnedka jsem věděl /, {abych to uvedl na pravou míru} že se děje něco nepatřičného /.

Zvláštním případem jsou **větné** vsuvky na začátku a konci SU. Tyto se v našem systému jako vsuvky označovat **nebudou**, ale podle svého charakteru se označí buď jako samostatná SU nebo se oddělí hranicí klauze (tj. budou se brát jako jakési uvozovací věty či samostatné SU – podle toho, co jejich struktura dovoluje). Jediná výjimka je uvedena na konci této podsekce. Pozor, stejné pravidlo platí i pokud se v dané SU před potenciální A/P vyskytuje pouze Delreg, DM nebo DR!

Vedlejší věty vyskytující se na začátku nebo konci SU typu „*jak jste včera řekl*“, „*jestli mi rozumíte*“, „*použiju-li jeho slova*“ budeme chápat jako jakési „uvozovací věty“ a budeme s nimi zacházet stejně jako s větami uvozujícími přímou řeč.

Takže sto jedna hlasů i s Pavlem Dostálem /, jak sme dneska četli //.

rozumím tomu tak /, že málo vlastně pro vymahatelnost toho práva /, jestli to můžu takle transponovat //.

Pokud se na začátku/konci vyskytne vsuvka ve tvaru odpovídajícímu větě hlavní, označíme ji jako samostatnou SU.

dovede nás to k historickému kolapsu /, používám jeho slova /.

Pozor, jako A/P se neoznačují větné vsuvky mezi dvěma hlavními větami souřadného souvětí. V těchto případech tvoří „vsuvka“ samostatnou SU.

ale seznamy sou prostě podkladem pro lustraci /, myslím teď ty vládní seznamy nepochybně /, a tyto seznamy nejsou přesné /.

pan Dostál tvrdí /, že asociační dohody s Evropskou unií nebyly vyjednány nejlépe /, dokonce až použil slovo nejhůře [LOUD_BREATH] /, a bojí se /, abychom nedopadli úplně špatně /.

Jedinou výjimkou, kdy se může A/P vyskytovat na konci SU jsou nevětné vsuvky delší než 2 slova.

jenomže lidé se obávají obrovského cenového nárazu {jeho slovy řečeno} /.

2.2 Více výplní za sebou

Řečníci často produkují více výplní za sebou. V tomto případě se každá výplň anotuje samostatně, a to i v případě, že výplně jsou stejného typu.

2.3 Výplně (DM) na hranicích syntaktických jednotek

Pokud se výplně vyskytují na hranicích syntaktických celků, je třeba určit, k jakému celku náleží. Pokud se po výplni (obvykle DM) nemění řečník, označí se jako náležející k celku,

který po ní následuje. Výjimkou je pouze případ, kdy výplň zvukově *zřetelně* náleží k SU před ní. Pokud je výplň v okolí /, nebo /&, umístí se tyto značky vždy před výplň. Pro zjednodušení platí pravidlo, že pokud je v okolí DM SU symbol s jedním lomítkem („/“, „/“, „/“, „/“, „/“, „/“, „/“, „/“) umístíme DM vždy až za hranici, naopak pokud je v okolí symbol se dvěma lomítky (//? nebo //.), necháme DM před symbolem, pokud je zjevné, že zvukově náleží k předchozí SU.

ono je to teď takový nejasný / . jo ale spolu s kolegy z výboru chystáme novou legislativní úpravu /.

ono je to teď takový nejasný jo // . ale spolu s kolegy z výboru chystáme novou legislativní úpravu /.

3. Editační neplynulosti (korekce obsahu promluvy)

Neplynulosti jsou nechtěná selhání řečníka při formulování a produkci promluvy. Editační neplynulosti jsou úseky řeči, ve kterých řečník koriguje nebo zcela mění obsah (části) své promluvy. Může také započatou promluvu zcela opustit. Struktura editační neplynulosti je následující

- Deletable region (vymazatelná oblast, Delreg, v literatuře též reparandum) – obsahuje původní úsek řeči, který řečník následně opravuje.
- Bod přerušení – Interruption point (IP) – bod přerušení je bod, ve kterém mluvčí přeruší Delreg pomocí EET, opakování, revize nebo opakovaného začátku. Všechny neplynulosti mají alespoň jeden bod přerušení na pravém konci Delregu. Existují i tzv. složité neplynulosti, které mají více než jeden IP.
- Editační fáze (Interregnum) – v tomto úseku si řečník uvědomuje, že se dopustil neplynulosti. Může obsahovat EET, DM či vyplněnou pauzu.
- Korekce (oprava) – obsahuje plynulý úsek promluvy, ve kterém řečník opravuje neplynulost. Opakované začátky korekce neobsahují.

Typickými druhy neplynulostí jsou repetice, revize a opakované začátky.

3.1 Anotace editačních neplynulostí

Delreg: Pro každou neplynulost se identifikuje jeho plný rozsah a označí se všechna slova, která do něj patří. (V tomto dokumentu se tyto úseky budou označovat pomocí hranatých závorek []). Pravidla pro identifikaci rozsahu Delregů budou popsána níže. V rámci Delregu je zakázáno používat jakékoliv SU symboly!

IP: Body přerušení na koncích Delregů jsou automaticky identifikovány a označovány anotačním programem. Vícenásobné IP v rámci jednoho Delregu musí být však *anotována manuálně*. V příkladech v tomto dokumentu se IP označují hvězdičkou *.

EET: EET se řadí mezi výplně, pravidla pro jejich přepis byla uvedena výše.

Korekce: Do korekce se zahrnou právě ta slova, kterými řečník nahrazuje slova a slovní fragmenty v Delregu. V jistých typech neplynulostí (opakované začátky) může být korekce prázdná.

Typ editační neplýnulosti: Typy neplýnulostí (revize, opakovaný začátek, repetice) se v tomto projektu neoznačují!

3.2 Příklady editačních neplýnulostí

3.2.1 *Repetice*

V případě repetice je Delreg zopakován v opravené části promluvy. Opakovaná část může být slovní fragment, kompletní slovo či více slov:

Jedná se o [bi~] [EE[HESITATION] bilaterální problém /.

*[Pane] * pane poslanče, prosím /.*

*[Ta věc] * ta věc se má následovně /.*

Je třeba mít na paměti, že ne všechna opakování jsou neplýnulosti, někdy může řečník zopakovat slovo nebo záměrně. Takové případy se jako repetice neoznačují, zajímají nás jen nechtěná selhání.

Ty náklady budou velmi velmi vysoké /.

3.2.2 *Revize*

Revize jsou velmi podobné repeticím. Rozdíl je v tom, že při nich řečník mění význam původní výpovědi místo toho, aby pouze opakoval již vyřčená slova. Revize může buď zcela změnit význam promluvy, nebo ho jen vyjasnit.

*Naše děti milují [kočky] * **chci říct** psy.*

*Chtěl vědět odjezdy [z Prahy v] * [EE[HESITATION] z Plzně v úterý.*

***No** rozhodli jsme se přistavit garáž / . [koupili jsme au~] a jak byla postavená /, koupili jsme auto na leasing /.*

Komplikací jsou občasné případy, kdy informace obsažená v Delregu není obsažena v revizi (korekci). To se stává nejčastěji v případě, když Delreg obsahuje jméno nebo jmennou frázi, ale korekce pouze zájmeno.

*Letěl jsem [do Bostonu] * **myslím** odtamtud v patnáct třicet.*

Revize tohoto typu jsou velmi vzácné, budeme je anotovat standardním způsobem, přestože to vede k malé ztrátě informace po potenciálním vymazání Delregu.

3.2.3 *Opakované začátky (restarty)*

V případě opakovaných začátků řečník původní formulaci zcela opustí a začne promluvu formulovat znovu. Obvykle je tento opakovaný začátek předcházen zřetelnou pauzou a může být provázen dalšími prozodickými jevy, jako je např. přenastavení výšky hlasu (pitch-reset). Korekce se v tomto případě považuje za prázdnou, původní formulace je přeformulována a její znalost není pro pochopení rozhovoru nutná.

*[Myslíš, že chce] * na jaký způsob mu to chutná /?*

[Je to] já jsem dřív žil na Moravě.*

*[Já bych] * chci jet do Brna.*

Nicméně někdy se stává, že původní formulace obsahuje nějakou informaci, která se již dále v rozhovoru neopakuje. Pro zachování konzistence přepisů se celá původní formulace stejně označí jako *Delreg*, přestože tím po jeho případném vymazání ztrácíme určité množství informace.

*[Bydlel jsem nedaleko od Pardubic~] * **no prostě** jsem pracoval pro tuhle firmu.*

Opakované začátky se někdy mohou plést s nekompletními SU (viz dále). Je třeba mít v patrnosti, že opakované začátky jsou omezeny na případy, kdy řečník opravuje sám sebe a není přerušen jiným řečníkem.

Opakovaný začátek se nemusí týkat pouze reformulace celé věty, ale často nastává případ, kdy řečník opustí pouze část (klauzi) složitějšího souvětí a pak pokračuje. Potom se do *Delregu* označí jen ta část (klauze) souvětí, která je opuštěna a do souvětí doplníme takové SU, jaké by tam patřily po vymazání *Delregu*.

*... kdy si banky zakládaly určité skupiny /. [které] * [EE[HESITATION]] a ty banky potom těmto skupinám půjčovaly //.*

3.3 Složitě neplynulosti

Ve spontánní řeči jsou časté série po sobě následujících neplynulostí.

3.3.1 *Delreg obsahující více neplynulostí*

Pokud *Delreg* (reparandum) obsahuje více neplynulých částí promluvy, je třeba jako *Delreg* označit celou neplynulou část s maximální možnou délkou. Tento úsek začíná levým okrajem první neplynulosti a pokračuje až k pravému okraji (IP) poslední neplynulosti. Celá tato část se označí jako jeden *Delreg*, přestože obsahuje části více neplynulostí. Poslední bod přerušení na pravém okraji *Delregu* je identifikován automaticky, ostatní IP je nutné doplnit ručně.

*No ale [ta vel~ * ten velkej * ten * ten] * ten podivnej pocit to se prostě nedá popsat.*

3.3.2 *Neplynulé korekce*

V některých případech korekce editační neplynulosti obsahuje další neplynulost:

*Jsem si jistý [ta], že [ta] * ta firma to umí udělat.*

*Letěl jsem [z Dallasu] * [EE[HESITATION]] z [Den~] * Denveru dneska ráno.*

Anotace překrývající se struktury složitých neplynulostí leží mimo rozsah přepisů našeho projektu. Proto se označí jen každý *Delreg* neplynulosti zvlášť. Jednoduchou metodou pro přepis tohoto jevu je označení všech výskytů mimo posledního výskytu každého slova v neplynulosti jako *Delreg*. Správný přepis je naznačen v předchozích příkladech.

3.3.3 *Složitě neplynulosti a opakované začátky*

Často se stává, že důsledkem složitě neplynulosti je opakovaný začátek.

*[bychom mohli být * být vy~] * **prostě** měli bychom mít vyšší cíle /.*

Repetice (být být) ústící v opakovaný začátek nemůže být sama o sobě anotována jako repetice, protože vnořené neplýnulosti nejsou povoleny. V takovýchto případech je třeba označit maximální rozsah Delregu (v uvedeném příkladě je to celý nepovedený začátek) a v něm označit všechny IP.

3.3.4 Více neplýnulostí v jedné promluvě

V jedné promluvě (větě) se může vyskytovat i více neplýnulostí, pro všechny se aplikují výše uvedená pravidla: označovat maximální rozsah Delregu, anotovat všechny vnitřní IP a považovat poslední výskyt každého slova v neplýnulosti za korekci (která se neanotuje).

*[Dělali * dělali] * **ohledně** těch razítek dělali vám [nějaká] * **nějaké** problémy /?*

*[Když jsem * když jsem] * **když** jsem viděl ty peníze [na tom] * **ne** na stole /, tak se mi **že** jo úplně zatmělo [EE[HESITATION] před vočima /.*

3.4 Neplýnulosti a spojky

Na krajích (začátcích) neplýnulostí se často vyskytují spojky. Tyto spojky se nezahrnují do rozsahu Delregu, pokud nejsou přímo sami o sobě „neplýnulé“ (tj. pokud je řečník neopakuje či nereviduje v korekci neplýnulosti). V následujících příkladech se spojka do Delregu nezahrnuje.

*Souhlasím, ale [já] * **já** vím /, že **é** hodně lidí je líných přemýšlet a jít volit.
Ne, protože [to bylo] * **ta** osoba to **prostě** věděla.*

V následujících příkladech je tomu naopak:

*[a když] * **a** jestli to nevíte, tak se se do toho prosím neplette.
Řekni si, [a] * **ale** stejně ti to nepomůže.*

3.5 Neplýnulosti a výplně

Delregy, které se vyskytují ve výplních (např. repetice vyskytující se v odbočce) se označují jako běžné Delregy.

***To** [vi~] **víte** takhle to nejde.*

Výplně objevující se v Delregích jsou samostatně označeny jako výplně, zůstávají ale uvnitř Delregu.

*[Myslím si [EE[HESITATION] že] * **myslím, že** to tak nejde.*

Výplně, které se objevují na levém nebo pravém okraji Delregu se do něj nezahrnují, ale anotují se zvlášť jako výplň.

*[EE[HESITATION] [odešel ten] * **přišel** ten člověk.
[Ale ta] * **Prostě** ta toho není schopna.*

Pokud se výplň vyskytne mezi dvěma sousedícími Delregy, je vyjmuta z rozsahu Delregů. Místo toho se oba tento případ označí jako dva Delregy s výplní stojící mezi nimi:

*SPRÁVNĚ: [já vím] * é [já nev~] * já nevím nic.*

*ŠPATNĚ: [já vím * é já nev~] * já nevím nic.*

Vyjímá se z Delregů pouze v případě, že je v rámci neplynulosti „mezi dvěma opakováními“. Naopak, pokud se výplň vyskytne uvnitř jedné neplynulosti (ne mezi dvěma), zahrne se do Delregu.

*[já * já opravdu é mám] * já opravdu nemám doma kočku.*

Pokud se před Delregem vyskytuje výplň, nezahrnuje se tato výplň do DelRegu.

*No [já si] * já si na to nepamatuju /.*

HESITATION [on šel k] v tý době on už tam byl /.*

4. SU

Klíčovým úkolem anotace je identifikace jednotek, které v rámci promluvy vyjadřují úplnou myšlenku. Tyto jednotky (SU) často odpovídají větě (souvětí), jindy věta obsahuje více SU. Každé slovo v promluvě musí být přiřazeno do nějaké SU. SU mohou být celé správně utvořené věty nebo fráze či jednotlivá slova. Mnoho krátkých SU, které nejsou plné věty nebo klauze (věty jako součásti souvětí), ale přesto jsou kompletní jednotky, slouží k řízení rozhovoru.

Důvodem anotace SU je vytvoření takových přepisů, ve kterých je informace čtenáři podána místo v dlouhých blocích v malých koherentních úsecích.

SU budou identifikovány na základě anotace jejich hranic. Budeme rozlišovat dva stupně označovaných hranic – interní v rámci SU a hranice SU.

Interní:

- /, - hranice klauze

Mám velký hlad /, protože jsem ráno nic nejedl /.

- /& - souřadné spojení vedlejších klauzí nebo hlavních klauzí, které nemohou stát samostatně

On mi tvrdil /, že tam nepůjde /& a že Petra taky zůstane doma /.

Bud' mi ty peníze dáš hned /& nebo s tebou už nikdy nebudu obchodovat /.

Hranice SU (externí):

- /. – oznámení (symbol // se používá pro oznámení s prozodicky vyznačeným koncem – viz část 5)

Dnes je pěkné počasí /.

- /? – otázka (symbol //? se používá pro otázku s prozodicky vyznačeným koncem)

Máš večer čas /?

- /~ – nekompletní SU přerušena jiným řečníkem

A: Mluvili jsme o tom člověku /, který včera /~

B: Počkejte /.

- /- – nekompletní SU svévolně opuštěná řečníkem

Vím o tom hodně /, protože to není /- oni musí bojovat s organizovaným zločinem /.

4.1 Pravidla pro přepis typických jevů

4.1.1 *Bezslavesné SU*

SU nemusí vždy obsahovat sloveso. Musí být však vždy jasné, že tato SU není syntakticky těsně svázána z předchozí – problémy mohou například způsobovat neplynulosti.

Vítám vás u sledování dnešního Radiofóra /. Hosté Jan Novák poslanec a Pavel Kučera stínový ministr obrany /.

Podle Pavla Rychetského {a cituji} je Václav Klaus v řadě svých postojů dokonce bližší sociální demokracii než ODS //. konec citátu //.

Častým případem bývá, že tázaný zopakuje téma otázky nebo uvede „název“ tématu, než o něm začne hovořit.

A: Jakým soupeřem je pro vás Liberec /?

B: No Liberec /. My jsme s nimi odehráli dva dobré zápasy /. Ale měli sme problémy s proměňováním šancí /.

Pokud po nějaké ukončené frázi v místě, kde lze napsat tečku, udělá řečník dlouho pauzu (řekněme větší než 0,5 vteřiny), použijeme symbolu „//.“. Pokud to ze syntaktické podstaty věty nejde, necháme místo navzdory dlouhé pauze neoznačené.

Jednoduše řečeno //. [SILENCE] politika nikdy nemůže principiálně vyloučit určitou míru eskalace //.

Přišel k nám ten starý Ø [SILENCE] vlčák /.

4.1.2 *Volně připojené větné členy*

Vymyslel si Ø a to předevečirem večer v koupelně při holení Ø zcela nový plán /.

Konečně dal gól Ø a to kupodivu hned Dominiku Haškovi /.

Tzv. volně připojené větné členy uvedené v předcházejících příkladech se žádným SU symbolem neoddělují.

4.1.3 *Presumptivní otázky (tag-questions)*

Jedná se o specifický typ otázek, ve kterých mluvčí předjímá platnost toho, co se v nich vyjadřuje. Označíme je jako otázky, presumptivní dovětek je oddělen hranicí klauze.

Vypadá to jako by zima nikdy neměla skončit /, že jo /?

Tys tam pracoval spoustu let /, ne /?

Vy jste patrně pan Novák /, že /?

To není dobrý nápad /, není-liž pravda/?

Pokud se nejedná o otázku (což bývá zřetelné z intonace), dovětek (že, jo, ...) se označí jako DM.

Přišli tam včera jo //.

4.1.4 Otázky, které řečník směřuje sám na sebe a řečnické otázky

Řečnické otázky a otázky, které řečník směřuje sám na sebe, se anotují stejně jako jiné otázky.

Dále se dostáváme k problému ústavního systému /. Má se do sněmovny volit poměrným nebo většinovým systémem /? Odpověď není jednoduchá /.

Třeba vyhraje /. není snad míč kulatý /?

ale jestli sme dojednali dobře nebo špatně /? těžko to posuzovat /, protože není podle čeho vysloveně //.

4.1.5 Souřadně spojené otázky

Pokud nenastala nějaká výše uvedená výjimka (v tomto případě je to ovšem více časté), souřadně spojené otázky (typicky spojkou „a“) označíme jako dvě tázací SU.

Tomáš půjde večer do divadla /? a Robert bude hlídat malého /?

Obvyklá elipsa je vynechání pomocného slovesa být v budoucím čase ve druhé větě.

Bude Tomáš hrát v divadle /& a Robert hlídat malého /?

4.1.6 Oznamovací věty s tázací intonací

Někteří řečníci pronášejí oznamovací věty se stoupavou intonací připomínající otázku. Když má promluva jasně funkci oznamovací, měla by být i takto označena. Ve sporných případech má přednost oznámení.

4.1.7 Vložené otázky

Když je otázka vložena ve větší klauzi, SU je označena podle vlastnosti celé věty.

Ředitel se ptal /, „Kdo si to vezme na starost (?)“ /.

4.1.8 Fráze fungující jak kompletní výpověď

Krátké fráze, které netvoří gramaticky kompletní věty, ale přesto mají funkci kompletní výpovědi (např. přímá odpověď na otázku) se označují jako oznamovací SU.

A: Na jak dlouho to vidíte /?

B: Na tři týdny /.

4.1.9 Upřesňování

Upřesnění se vyskytují na konci sdělení a přinášejí další informaci či opravují informaci již uvedenou. Je třeba dát pozor, aby nebyly zaměněny s neplynulostmi obsahujícími Delreg. Tato upřesnění se od původního sdělení oddělují hranicí klauze /,

Myslím /, že existují dva předpisy /, které to upravují /, vlastně tři /.

Když jsem byl malý /, bydleli jsme v Ostravě v malém domku /, v Slezské Ostravě /.

narůstá do určité úrovně u jednotlivých plodin /, do určité výnosové úrovně /.

4.1.10 Nekompletní SU

Když promluva netvoří gramaticky kompletní větu nebo frázi a nevyjadřuje kompletní myšlenku, označíme ji znakem nekompletní SU buď /- nebo /~. Pokud promluvu přeruší jiný řečník, použijeme symbolu /~. Pokud naopak řečník výpovědní konstrukci nedokončí svévolně, použijeme symbolu /-. Tj., symbol /~ se může vyskytovat pouze na konci turnu, symbol /- jak na konci tak i uvnitř turnu. Ve standardním textu by se podobné události většinou značily pomocí tří teček (...).

A: Když vložíte dost peněz do /~

*B: **Jasně jasně** ale to není jen o tom /.*

ty boty /, kerý měl schovaný pod botníkem /, vypadaly jak /- von byl takovej budižkničemu /.

Nekompletní SU se nesmí zaměňovat s neplynulostmi – opakovanými začátky. Pro rozhodnutí mezi opakovaným začátkem (tj. uzavřením celé nekompletní SU do Delregu) a /~ nebo /- se používají následující pravidla

1. Opakovaný začátek se nemůže vyskytovat jako poslední SU v turnu. V tomto případě se podle kontextu vždy použije /~ nebo /-.
2. Pokud řečník nedokončenou větu vzápětí opraví, znovu naformuluje, jedná se o opakovaný začátek. Pokud se k ní nevrací, jedná se o /-
3. SU typu /- musí vždy obsahovat alespoň jeden SU-interní symbol (tj. /, nebo /&) nebo užitečnou informaci, která již není následně v rozhovoru zopakována. To ovšem neznamená, že výskyt SU-interního symbolu okamžitě implikuje použití /-. Pokud se jedná jen o krátkou uvozovací větu, která nenese samostatný obsah („víte, že...“, Petr včera řekl, že ona...“, je možné ji celou zařadit do DelRegu.

4.1.11 Souhlasy a nesouhlasy

Reakce vyjadřující souhlas či nesouhlas s předřечníkem jako jsou *Ano, Ne, Jo, Hm* obvykle tvoří samostatnou SU.

A: *Můžeme podat odvolání /?*

B: *Ano / . to je možné / .*

A: *Stačí /?*

B: *[HM] / . pojd'me dá /.*

Pozor, neplést s větami typu:

Ne že by byl Zdeněk Škromach autorem výroku o duhové koalici /.

kde se po *ne* žádný symbol nezapisuje.

Slova jako *Ano, Ne, Jo* však mohou v dialogu hrát i jinou roli než jen vyjádření souhlasu či nesouhlasu. Mohou fungovat jako backchannel či jako DM. Platí základní pravidlo, že se nejedná o souhlas/nesouhlas, pokud mu bezprostředně nepředchází otázka. Uved'me si příklad imaginárního rozhovoru:

A: *Předpokládá se exekuce /.*

B: *Ano /?*

A: *Ano / . Již byl vydán příkaz /.*

B: *Ano Ø tak to je problém /.*

A: *Ano Ø je to nepříjemné /.*

B: *Tak to oni přijdou / . ano Ø a všechno seberou /?*

Pozor také na [HM] – to má funkci kompletní SU pouze pokud reprezentuje přímou odpověď na otázku, jinak jde o DR!

V případě že za odpovědí následuje jeho vysvětlení připojené pomocí podřadicí spojky, oddělíme jej pomocí /,

Ano / , protože on přijel až včera /.

4.1.12 Idiomatická spojení

Idiomatická (ustálená spojení) obsahující více sloves se neoddělují žádnou SU značkou.

Tady platí padni Ø komu padni /.

Ať dělám Ø co dělám / , nefunguje to /.

To je prašť Ø jako uhod' /.

Půjdu tam Ø děj se Ø co děj /.

4.1.13 Přímá a nepřímá řeč (citace) - Z!

Věta uvozující přímou řeč (citaci) tvoří spolu s přímou řečí jednu jednotku. Od sebe jsou odděleny pomocí /, Pozor, uvozovky se v prepisech řeči nepoužívají.

*Ministerstvo tvrdí /, my za to nemůžeme /.
Když odcházel /, řekl /, já se vrátím /.
Půjdu tam /, řekl mi včera.*

V rámci přímé řeči se používají všechna standardní pravidla označování souvětí. Výjimka nastává pouze v případě, že se citace skládá z několika vět (SU).

Pokud uvozovací věta předchází postupujeme následovně. Pokud je řeč „přímá“, je na uvozovací větu navázána pouze první SU, ostatní stojí samostatně. V následujících příkladech jsou naznačeny pro ilustraci uvozovky.

Ale premiér říká /, „nikdy jsem ho neviděl / . Vůbec toho člověka neznám / . Celá ta aféra je směšná / .“

U „nepřímé“ řeči záleží na její struktuře. Řídíme se zásadou „Pokud jste na pochybách, dělte“. Tj. ve sporných případech mají SU-externí symboly přednost před SU-interními.

Jiří říká /, že dražba je krajní řešení / . ale mohou k němu přistoupit.
vs.

Jiří říká /, že dražba je krajní řešení / . ale že k němu mohou přistoupit.

Pokud se uvozovací věta nachází mezi těmito větami, anotujeme přímou řeč podle následujícího případu (uvozovky jsou naznačeny pro ilustraci):

„Takhle to nepůjde“ /, řekl včera Tomáš / . „To bychom nikdy nic nevyřešili / . Svoboda slova se musí respektovat / .“

4.1.14 Osamostatnělé vedlejší věty

SU mohou tvořit i tzv. osamostatnělé vedlejší věty:

*Že jste si toho nevšiml /.
Aby už bylo jaro /.
Když to teď už nejde /.*

Jako samostatnou SU anotujeme vedlejší větu také v případech podobných následujícímu (vedlejší větu nelze navázat na žádnou hlavní – něco jako vysvětlovací poměr):

*jak takový pachatel reaguje /, pokud je neozbrojeným majitelem při činu přistižen /? **protože paní Skalová měla vlastně štěstí v neštěstí** / . vy ste těm pachatelům tváří v tvář ani v jednom případě nestála //.*

4.1.15 Parcelace

Z hlediska mluveného projevu, v němž mluvčí nemá vždy jasnou představu o perspektivě svého projevu, můžeme mluvit o dodatečném připojování prvků. Mluvčí touto segmentací vytváří vlastně několik eliptických výpovědí a každá z těchto výpovědí má vlastní intonační centrum. Pokud je parcelace opravdu velmi zřetelná (to jest řečník udělá zřetelnou pauzu), rozdělíme výpověď na několik samostatných celků.

Ty s ním chceš mluvit /? Sama /? Beze svědků /?

V případě, že není zřejmé, zda-li jde o parcelaci, nebudeme promluvu rozdělovat.

4.1.16 Infinitivní konstrukce

Infinitivní konstrukce, tj. spojení infinitivu s jeho rozvíjejícími větnými členy, tvoří jednu kompaktní SU.

Máte za úkol vytrhat na zahrádce všechny plevele /.

4.1.17 Apozice (komplexní větné členy) – Z!

Složky komplexních větných členů jsou v syntaktickém vztahu, kterému se říká apozice; označují v podstatě touž skutečnost. V psaném textu se tyto větné členy zpravidla oddělují čárkou, my je však nebudeme pomocí SU symbolů nijak oddělovat.

Jiří Malír, ředitel naší školy, je můj přítel /.

Byla zvyklá vstávat velmi brzy, už ve čtyři hodiny /.

Prezident České republiky, totiž Václav Havel, odcestoval na dovolenou /.

Výjimkou nejsou ani přístavky, před kterými se v psaném textu obvykle píše dvojtečka či pomlčka a řečník před nimi dělá krátkou pauzu – tyto přístavky budeme oddělovat pomocí /.

Často nás volal k nějaké drobné práci Ø zamést stodolu nebo přebrat jablka /.

parlament ve všem možném jedná veřejně Ø o důležitých zákonech, o registrovaném partnerství, o důchodové reformě /.

V apozičním vztahu bývají nejčastěji výrazy jmenné, ale stejně tak to mohou být (podle širokého pojetí Šmilauerova) slovesa, příslovce, ba i věta (klauze). Pokud je v tomto vztahu věta **a je vložená do jiné věty** (tj. po apozici ještě následuje nějaký větný člen první věty) oddělíme ji pomocí /, Pokud „větná apozice“ může stát samostatně, označíme ji jako samostatnou SU (opět platí zásada pro sporné případy „Dělte, kde můžete.“

Zvláštní zacházení se použije pro často používané konstrukce uvozující celou větu jako jsou „to jest“ a „to znamená“. Tyto fráze nebudeme brát jako věty, ale budeme je chápat jako částice.

řeka se kroutí /, tedy tvoří meandry /, blízko u lesa /.

řeka se kroutí /, to znamená tvoří meandry /, blízko u lesa /.

zde řeka to jest vodní tok tvoří meandry /.

údaje byly takové /, že nakonec soudy ty lidi osvobodily /. to znamená zbavily je té lustrace /.

4.1.18 Juxtapozice vět – „uvozovací věty“

Jako juxtapozici „uvozovací věty“ chápeme takové spojení hlavních vět, mezi nimiž není žádný z významových vztahů (slučovací, vylučovací, stupňovací atd.). Druhá věta je ve vztahu syntaktickosémantické determinace k větě první, ne však ve vztahu syntaktickoformálním (mluvnický se tedy jedná o vsuvku). Obvykle by po první větě mohla následovat spojka *že* bez změny významu. S těmito konstrukcemi budeme zacházet stejně jako s nepřímou řečí a její uvozovací větou, tj. oddělíme je pomocí hranice klauze /,

Já vím /, vy to nemáte rád /.
Řekněte upřímně /, je to pravda /?
vůbec se mu nedivím /, musím říct /.

Prosím vás pěkně /, je vůbec myslitelné /, že by někdo z koaličních poslanců si troufl za těchto okolností podrhnout svou vlastní vládu //?

Zvláštní zacházení se použije pro konstrukce uvozující celou větu jako jsou „*to jest*“ a „*to znamená*“. Tyto fráze nebudeme brát jako věty, ale budeme je chápat jako částice. Pozor, nebudeme je označovat ani jako DM!

To znamená Ø banka neotevře své pobočky /.
To jest Ø předčasné volby nejspíš nebudou /.

4.1.19 Anakolut

Častým jevem je ve spontánní řeči tzv. anakolut, tj. řečník dokončí větu (např. po vsuvce, vložené větě či neplynulosti) jinak, než je podle jejího začátku očekáváno (např. špatná shoda podmětu s přísudkem). K anakolutu typicky dochází v okolí vsuvky. Tyto anakoluty se nebudou nijak anotovat a anotátor se bude chovat tak, jako kdyby věta byla gramaticky správně.

pokud se skupina států {a nejsou to jenom Spojené státy je to Británie a další} rozhodnou použít sílu /, tak ...

todlensto je asi jedna z oblastí /, kde sociální demokracie , {protože momentálně máme koaliční vládu nikoliv sociálně demokratickou , ale koaliční,} tak v tédlenté oblasti ta sociálně demokratická menšinová vláda asi má určitý dluh /.

4.1.20 Turny, kterým chybí začátek

Turny, kterým chybí začátek, resp. jejich začátek je v předchozím turnu (obvykle překrývající se řeč) se anotují tak, že se jejich začátek bere v potaz.

PŘEKRÝVAJÍCÍ SE → (A: I vás překvapil . B: Na druhou stranu překvapil ~)
B: ~ i mě /, a také sem se tak pro deník Právo k tomu vyjádřil /.

4.2 SU a souvětne konstrukce

Pokud se jedná o SU větné podstaty (tedy ne o samostatné slovo nebo frázi), je její klíčovou vlastností výskyt hlavní věty (nezávislé klauze). SU je nejmenší možná jednotka promluvy, která může teoreticky stát sama o sobě jako samostatná věta bez změny významu. SU tedy nezávisí na žádné jiné klauzi pro kompletaci myšlenky.

Pokud je mezi klauzemi podřadicí vztah, označíme jejich hranici vždy jako /, Složitější rozhodnutí je nutné učinit, pokud je mezi klauzemi souřadicí vztah. Základní zjednodušené rozhodovací pravidlo je takové že:

Všechny hlavní věty, které mohou stát v textu samostatně (spolu se všemi vedlejšími větami na nich závislými) bez změny významu, budou tvořit samostatnou SU.

Nejdříve je tedy nutno rozhodnout o vztahu vět, vodítkem může být spojka.

Podřadicí spojky jsou například:

*že
protože
jelikož
když
přestože
třebaže
aniž
ačkoliv
byť
zatímco
než
pokud
jestliže*

Pozor, se spojkou *takže* budeme zacházet jako se spojkou souřadicí.

Vedlejší jsou i věty vztažné, např.:

To je ten muž /, kterého jsem včera viděl /.

Známkou vedlejší věty je i to, že je předcházena tzv. korelativním výrazem. Např. „... *proto, že ...*“, „... *proto, aby ...*“, „... *tak, že...*“, „... *tehdy, když ...*“, „... *tak, aby ...*“

Vedlejší věty připojené pomocí *což* a *takže* vyžadují zvláštní zacházení – viz níže.

Souřadicí spojky jsou například:

*a
nebo
ale
avšak
však
kdežto
ani
tudíž
nýbrž
či
čili*

vždyť
až

(neboť) – tato spojka je ovšem sémanticky podřadící, bude s ní zacházeno zvláštním způsobem

4.2.1 Rozlišení souřadného a podřadného spojení

Pro rozlišení souřadného a podřadného spojení (resp. hlavní a vedlejší věty) ve sporných případech existuje několik pravidel. Žádné z nich však není úplně stoprocentní.

Kritéria pro určení vedlejší věty:

1. Vedlejší větu je možno nahradit větným členem věty hlavní (substituční kritérium).

Může nám při tom pomáhat otázka skládající se z věty řídicí a tázacího zájmena nebo tázacího příslovce, popř. slovního spojení obsahujícího tázací zájmeno *jaký*. Např. v souvětí „Bylo pochopitelné, že s námi nesouhlasí.“ se můžeme ptát *Co bylo pochopitelné?* nebo *Jaká věc byla pochopitelná?*

2. O mnohých větách vedlejších platí, že mohou být jak v postpozici (za větou řídicí), tak v antepozici (před ní). Věta, která spolu se svým spojovacím výrazem může být jak v postpozici, tak v antepozici, je věta vedlejší. Z vázanosti věty na postpozici či antepozici však ještě neplyne, že jde o větu hlavní. Některé věty vedlejší se však užívají pouze v postpozici (např. věty účinkové s „... *tak, že ...*“, přirovnávací s *než*, věty časové s *až* atd.). Tato vázanost na postpozici je dána jejich některou specifickou vlastností.

3. Modálnostní kritérium – podřadně spojené věty nelze transformovat na tázací souvětí. Srov. souvětí „*Procházeli se po lese, a to pomohlo jejich zdraví.*“, kterou lze transformovat na „*Procházeli se po lese, a pomohlo to jejich zdraví?*“, se souvětím „*Procházeli se po lese, což pomohlo jejich zdraví.*“, které takto transformovat nejde.

4.2.2 Odlišení souvětí od několikanásobného přísudku

Několikanásobný přísudek se rozděluje pomocí /&. Rozlišení souvětí a věty s několikanásobným přísudkem je v mnoha případech velmi obtížné. Problémem také je, že na dělicí čáře mezi těmito dvěma jevy neexistuje v literatuře jednoznačná shoda.

Abychom dosáhli co největší konzistentnosti anotací, budeme pro naše účely za věty s několikanásobným přísudkem považovat věty, splňující jednu z následujících podmínek:

1. Přísudková slovesa mají společné rozvíjející členy (předmět, příslovečné určení, doplněk či vedlejší věta). Mezi slovesy musí být jasná významová souvislost.

Ovce žerou /& a spásají všechnu zelenou trávu /.

Nacpal /& a zapálil si dýmku /.

Večer rád leží a čte /.

Obsloužíme /& a poradíme vám k vaší plné spokojenosti /.

Ředitel si přeje /, abych Novákovi tlumočil /& a zdůraznil /, že se to už nesmí opakovat /.

2. Přísudky mají slovesa společná modální nebo fázová slovesa.

Musíte hrát /& a zpívat častěji /.
Budu odpočívat /& a číst tvé básně /.

Naopak o souvětí se jedná, když

1. Každý přísudek má své vlastní doplnění

Dobře se naobědval /. *a šel ven /.*
Učitel pokáral chlapce /. *a pochválil děvčata /.*

2. Predikátory mají různou časovou formu a mohou stát samostatně

Lhal jsi /. *lžeš /.* *a lhát budeš /.*

Naproti tomu:

Albánie je /& a zůstane naším přítelem /. (nejedná se o několikanásobný přísudek, ale nemůže stát samo o sobě)

3. Predikátory konstruuji věty s různou modalitou či ilokučným potenciálem (tvrzení, přání, rozkaz atd.)

Teď přijel /. *a už chce odjet /?*
Uč se /. *a nepropadneš /.*

V konstrukcích, ve kterých se opakuje vícekrát to samé sloveso, budeme používat koordinační symbol /&

Poutník šel /& a šel /.

4.2.3 Podřadná souvětí

Subordinace může být i vícenásobná.

Stalo se to včera /, *když mi řekl /,* *že mě chce propustit /,* *protože nejsem loajální k firmě /,* *která mi dává práci /.*

Pokud je věta závislá vložena do věty řídící, odděluje se symbolem /, z obou stran.

Ten Němec /, *který to prodává /,* *přijede zase zítra /.*
Toto auto /, *přestože se jedná o sportovní vůz /,* *má až překvapivě nízkou spotřebu /.*

4.2.4 Vedlejší věty připojené pomocí „takže“

Věty připojené spojkou *takže*, které se obvykle označují jako vedlejší (účinkové, důsledkové), budeme my zde označovat naopak jako věty hlavní a oddělíme je pomocí externích SU symbolů.

Přišel už večer /. *takže jsme neměli žádné obavy /.*

4.2.5 Rozhodnutí o prioritě /, či /&

Může nastat případ, že podle logiky jde na jedno místo vložit jak symbol /, tak i /& - např. konec vedlejší věty a zároveň koordinace následující hlavní věty s řídicí větou té věty vedlejší. Pro tyto případy se stanovuje pravidlo, že symbol /& má přednost.

Bud' přišel pan Novák /, který je za to zodpovědný /& nebo telefonoval pan ředitel /.

4.3 Odlišení konce SU(/.) a koordinace (/&) – Z!

Pokud odlišíme souřadné spojení od podřadného, nastává další krok, tj. odlišení hranice SU (/.) od souřadného spojení v rámci jedné SU (/&). Symbolu /& použijeme místo /, pokud nastane jeden z následujících případů:

1. V souvětí existuje taková věta vedlejší, která je závislá na obou větách hlavních.
2. Druhá věta je eliptická a nemůže stát sama o sobě.
3. Souřadné spojení je realizováno formálně souřadnou avšak sémanticky podřadící spojkou *neboť*.
4. Souřadné spojení je realizováno pomocí spojovacího výrazu skládajícího se ze dvou částí (dvojitou spojkou – např. „sice – ale“, „bud' – nebo“, „nejen – ale i“) rozděleného mezi obě hlavní věty.

Pro rozhodování ve sporných případech platí pravidlo: „Dělte, kde můžete!“, tj. hraniční (externí) SU symboly mají vždy přednost před interními, pokud neexistuje jasné pravidlo určující použití interního symbolu.

4.3.1 *Vedlejší věta závislá na více větách hlavních*

Pokud je na obou z dvojice hlavních vět závislá jedna vedlejší věta, jsou místo /, odděleny pomocí /&.

Jak to dopověděl /, sklopila ta dívka hlavu /& a začervenala se jako růže /.

4.3.2 *Elipsy*

Elipsa znamená vynechávání určitých výrazů ve větě. Pokud je v druhé větě souřadného souvětí vynechán (totožný) přísudek (nebo jeho část) a mezi větami není zřetelná pauza, tvoří pouze jednu SU a druhá věta je oddělena pomocí /&.

Nejdříve jsem uviděl psa /& o něco později i Marii /.

Já se musím dívat /, co bude v příští minutě v příští hodině /& a samozřejmě snažit se o to /, aby ten trend byl pozitivní //.

Jedná se o problém /& či nikoliv /?

Typickým příkladem použití elipsy je vyjádření příkladu a jeho exemplifikace či výjimky.

dovedla mu všechno odpustit /& i nevěru a lež /.

dovedla mu všechno odpustit /& jen lež ne /.

Tereza přinesla růže /& Katka tulipány /.

Občas se vyskytují i následující případy:

a ten známý scénář /, přišel /, píchlé gummy /, někdo ho na to upozornil /, a zatímco on tam s tím manipuloval /, tak byl okraden //.

4.3.3 Stažení souvětných konstrukcí

Jde o vynechávání různých elementů anteponované věty. Např. z věty „*Pak se říká /, že lidé nemají peníze /.*“ vznikne „*Pak že lidé nemají peníze /.*“ Podobně „*Kdo že tam byl /?*“. Stažená může být i věta vedlejší – „*Když tak na mě počkej /.*“ Tyto stažené věty tvoří jednu SU.

4.3.4 Spojení spojkou neboť

Pokud je souřadné spojení je realizováno formálně souřadnou avšak sémanticky podřadící spojkou neboť, použijeme symbol pro koordinaci /&.

Rohylnol bude na zvláštní předpis /& neboť ho zneužívají narkomani /.

4.3.5 Souřadné spojení pomocí výrazů složených ze dvou částí

Souřadné spojení je realizováno pomocí spojovacího výrazu skládajícího se ze dvou částí, které jsou obsaženy v obou hlavních větách, a tudíž žádná z těchto dvou vět nemůže stát sama o sobě. Opět použijeme /&. Příklady takovýchto výrazů jsou *sice ... ale, bud' ... nebo, nejen ... ale i.*

Bud' mi dáš ty peníze /& nebo s Tebou už nikdy nepromluví ano slovo /.
Sice to je teoreticky možné /& ale v praxi se to nepoužívá /.
Nejen Tomáš to dostal /& ale dokonce i Petr to dostal /.

4.3.6 Souřadně připojená věta má zamlčený podmět a obsahuje pouze přísudek – ZRUŠENO!

ZRUŠENO!

4.4 Neplynulosti měnicí strukturu SU

Může se stát, že na hranici SU se objeví Delreg, že se vyskytne takový Delreg, který mění strukturu syntaktických celků. Potom se za skutečnou strukturu považuje ta, která je zamýšlena po opravě a případná SU značka se vloží před Delreg.

*...aby si třeba vybrali jenom částku /, [*potřeb*~] *, kerou potřebují na nájem ...*

5. Prozodicky vyjádřené SU

Pro odlišení konců SU, které jsou silně vyjádřeny pomocí prozodických prostředků (délka pauzy, poklesnutí nebo stoupnutí hlasem, prodloužení závěrečných slabik) se zavádí symboly //, a //? (místo /, a /?). Symboly s dvojitými lomítky dostanou přednost tehdy, pokud je ze zvukové nahrávky zřejmé, že na tomto místě chtěl mluvčí ukončit větu (souvětí). Vodítkem je

zejména použití zřetelné pauzy, někdy postačí i silné intonační vyznačení. Pokud jste na pochybách, zvolte raději dvě lomítka //.

V místech, kde je v přepisu konec věty současně s koncem řádky, se obvykle používají symboly // a //? – obvykle po nich následuje pauza, protože původní anotátor na tato místa vložil časovou značku.

6. Slova vyřčená řečníkem na pozadí - Backchannels

Dalším jevem, který je třeba podchytit jsou slova vyřčená řečníkem na pozadí, tzv. backchannels. Jedná se o slova či fráze, kterými řečník, který v daný moment nemá slovo (není dominantní) dává dominantnímu řečníku najevo, že poslouchá. Nejčastějšími případy jsou *Ano* a *[HM]*.

7. Ostatní

Mimo výše uvedených úkolů je třeba ještě dát pozor na správnost přepisu. Pokud se v něm vyskytne zřetelná chyba, tj. chybí nebo přebývá slovo či je nějaké slovo zjevně špatně přepsáno, pak jej opravte. Všimněte si i chybějících VP [HESITATION]. Pokud má vyplněná pauza spíše souhláskový charakter, označte ji jako [MM-HESITATION] (viz výše – Vyplněné pauzy). Pokud není vyplněná pauza jednoduchá, ale řečník udělá v její polovině pauzu, je nutné aby v přepisu bylo [HESITATION] dvakrát za sebou.

Naopak interpunkci neopravujte.

V případě, že na některém místě nejste schopni rozhodnout, jaký symbol zvolit, označte toto místo značkou (*HD*), znamenající obtížné rozhodnutí (hard decision). Tak bude snadné tato místa v přepisu najít a správně je označit.

Souřadné spojení podle poměrů vět

Poměr slučovací

Dvě hlavní věty v poměru slučovacím (viz Grepl, Karlík – Skladba češtiny) se rozdělují na 2 samostatné SU symbolem (/.) V případě, že jsou propozice rozděleny vztažnými zájmeny (*což, načež, přičemž, čímž*), budou odděleny jako vedlejší věty (pokud před nimi není jasné intonační zakončení a velmi výrazná pauza).

Maminka prala /. a přitom si prozpěvovala /.

Maminka prala /, přičemž si prozpěvovala /.

Koupil jsem dolary /. a potom jsem je ukryl do polštáře /.

Koupil jsem dolary /, načež jsem je ukryl do polštáře /.

Výčet paralelních faktů:

Jednak to nevím /& jednak nejsem poslancem /& a jednak jsem měl dovolenou /.

V článku se jednak píše o Máchovi /&. jednak se vykládá o českém romantismu obecně /.

avšak:

Za prvé tohle jsem nikdy neřekl / za druhé o tohle tady vůbec nejde / a za třetí v žádném případě neodstoupím /.

Plynulý přechod:

Dlouho jsem přemýšlel / až jsem všechno pochopil /.

Dal gól / a tím rozhodl zápas /.

Dal gól /, čímž rozhodl zápas /.

Už jsem vše řekl / a tak teď budu zticha /.

Slučitelnost:

Stačí mu zatelefonovat / a on hned přijde /.

Hodnocení:

Jiřina mlčela / a to ho popuzovalo /.

Jiřina mlčela /, což ho popuzovalo /.

S tzv. nepravými větami vedlejšími se spojkou „aby“ budeme ovšem zacházet jako se souvětími podřadnými:

Tenkrát odešel z politiky /, aby se dnes vrátil /.

Odporovací poměr

Hlavní věty v odporovacím poměru se také dělí na 2 SU.

On je ještě nedospělý / ale zato se vyzná v technice /.

Škodil / ale také pomáhal /.

Maminka mi to zašila / jenže to stejně nedrží /.

Máš hledět do učebnice / a ty se místo toho díváš po děvčatech /.

Chystal se jít spát / ale vtom někdo zazvonil /.

Stařec byl unavený / přece však dívku vládně uvítal /.

Děti snědly oběd / jen salát nechaly /.

Neučil se / nýbrž místo toho hrál karty /.

Konfrontační poměr

U konfrontačního poměru je situace složitější. Záleží na tom, jak jsou věty v konfrontačním poměru spojeny.

a) bezespoječná spojení – oddělíme /.

Já se dřu / ty si odpočíváš /.

b) výrazy typu *naproti tomu*, *naopak* a výrazy oddělené spojkou *kdežto* – oddělíme /.

Počet korálů se zmenšuje / rostlin naproti tomu přibývá /.

Počet korálů se zmenšuje / kdežto rostlin přibývá /.

c) věty oddělené původně podřadicími spojkami *jestliže, když, aby, -li* – pro jednoduchost oddělíme symbolem /, stejně jako by uvozovaly vedlejší větu

Jestliže to byl těžký problém /, dnes je to už neřešitelný problém /.

Byl-li jen lhářem /, dnes je zlodějem /.

Sparta nakonec vyhrála /, když už prohrávala 2-0 /.

Vedli jsme /, abychom pak stejně prohráli /.

Stupňovací poměr

Pokud jsou hlavní věty spojeny ve stupňovacím poměru spojeny jednoduchou spojkou, rozdělíme je na 2 SU. Naopak, pokud jsou spojeny dvojitou spojkou (*nejen – ale, nejenže – nýbrž, nejenže – ale*) použijeme /&.

Znala její city /, ba dokonce prohlédla její úmysly /.

To pravidlo platí /, nadto platí i jeho modifikace /.

Nezná Karla /, tím spíše nebude znát Marii /.

Znala nejen její city /& ale prohlédla i její úmysly /.

Nejenže je krásná /& ale je i bohatá /.

Alternativní poměr

V případě, že jsou věty v alternativním poměru spojeny pomocí dvojitých spojek, oddělíme je znakem /&. Stejně tak použijeme /&, pokud se jedná o alternativní otázku se spojkami *nebo, či*.

Bud' je na dvoře /& nebo je ve sklepě /.

Je /& či není dobrým ministrem /?

Asyndetické spojení alternativních otázek oddělíme hranicí interogativní SU /?

Má mě ráda /? Nemá mě ráda /?

Pokud je alternativní poměr vyjádřen pouze jednoduchou spojkou, rozdělíme souvětí na 2 SU.

Máme před sebou velkolepého podvodníka /, nebo sledujeme zázrak /.

Věty typu „Raději A než B“ budeme brát jako podřadné spojení.

Raději bude trpět /, než se omluví /.

Vysvětlovací poměr

Věty hlavní spojené ve vysvětlovacím poměru se rozdělí na 2 SU.

Představoval si ho přesně /, znal totiž dobře tento typ lidí /.

To byl ještě zdravý /, však tenkrát také ještě nepil /.

Předpis je předpis /, vždyť to znáš /.

*Ze strojovny se začalo ozývat rachtání /. to motor už běžel naplno /.
Měl z toho radost i tak /. nešlo mu přece jen o peníze /.*

Výjimka nastává, pokud jsou věty spojeny spojkou *neboť*, před kterou uijeme */&* (viz výše). Důvodem je to, že tato spojka se dnes nejčastěji využívá jako téměř synonyma spojky *protože* (*poměr příčinný*). Její použití je však dnes naštěstí již jen velmi řídké.

Šli jsme se koupat /& neboť bylo krásné počasí /.

SU ukončíme pouze tehdy, pokud před ní předchází otázka (opět se je zde možné použít spojky *protože*)

Kdo přikázal střílet /? neboť (protože) já to nebyl /.