

Stručné poznámky k anotování textů Českého akademického korpusu  
Alla Bémová

**léto 2007**

Paralelní anotování stejného textu různými anotátory přináší zajímavou zkušenost. Teoreticky vzato, výsledek by měl být v podstatě identický, jelikož vychází ze stejných pokynů manuálu. Množství rozdílů je však značné. Neshody se týkají jak samotného **budování stromových struktur**, tak i **přiřazení analytických funkcí**.

Příčiny neshod by se daly stručně shrnout do několika bodů.

**- nedostatečné porozumění textu.**

Příklad : *Prezident Čs. republiky propůjčil mistru sportu, nadporučíku Fr. Venclovskému vyznamenání Za statečnost, za prokázanou osobní odvahu a příkladnou bojovnost při plavbě kanálem La Manche.*

Všechny anotátorky zavěsily rozvití - *za statečnost, za odvahu a bojovnost* jako Atribut na substantivum vyznamenání. Ale ve skutečnosti je to - *vyznamenání Za statečnost* (Atr) *propůjčil - za statečnost...a odvahu... (2x Adv-)*

**- nedostatečně zažité pokyny manuálu.**

Projevuje se to např. v zavěšení koordinace.

Příklad : *Do strany jsem vstoupil v roce XY po čtyřleté práci ve válcovnách... a jako elektrikář tam zařezávám až po dnešní den.*

vstoupil	a
a	
zařezávám	vstoupil zařezávám

Zvláště často je to vidět při přiřazení anal. funkcí :

Např. - *rozloučení na letišti (Adv)*, ale je to přívlástek neshodný, tedy - **Atr**

- *velvyslanec VDR v ČSSR (Adv)*, ale je to - **Atr**

Mají-li koordinované predikáty společný subjekt, zavěšuje se na koordinaci - *...s delegací se rozloučil a šťastnou cestu popřál velvyslanec Teodor Haš.* Tedy subjekt Haš je zavěšen na koordinační uzel.

Mají-li substantiva v koordinaci společné rozvití adjektivem, zavěšuje se toto adjektivum rovněž na koordinační uzel, např. - *odhaluje nové aspekty a dimenze.*

Váhání mezi **Adv** a **Obj** pomůže do značné míry vyřešit valenční slovník, i když byl vytvořen pro rovinu tektogramatickou. K identifikaci nevlastních předložek pomůže jejich seznam, avšak rozhodování, zda jde o užití předložkové nebo nikoli, zůstává na anotátorovi.

**- nedostatečné pokyny v manuálu.**

Zdá se, že nejsou např. pravidla pro zavěšení čárek, protože v počátečním procesu anotování jsme se intuitivně naučili zavěšovat je stejným způsobem.

Příklad: *Když v luxusním Ferrari, které koupil před týdnem, uháněl Baroš po komunikaci mezi Ženevou a Lyonem, naměřili mu policisté neuvěřitelných 271 kilometrů za hodinu.*

Uzavírá druhá čárka v pořadí jen větu vztažnou, nebo se vztahuje i k větě vedlejší časové? Ve stromové struktuře ji zavěsíme výše, tzn. jako uzavírající celý úsek - *Když v luxusním Ferrari, které koupil před týdnem.*

Asi nikde není výslovně řečeno, že pokud to jde, nemělo by nic viset na předložce, nebo že na modálních slovese může viset subjekt, ale ostatní rozvití by měla viset na slovese významovém atd.

- **samotné chápání syntaktické konstrukce může být homonymní, různě interpretovatelné nebo i sporné.**

K takovým jevům patří např. rozlišování děje a stavu u trpného tvaru (přestože existují kritéria).

Příklad: *Všechna naše setkání byla naplněna vřelým přátelstvím.*

Jedna možnost chápání věty: nikdo nic nenaplňoval, je to totéž, jako - *Setkání byla přátelská.*

Z toho plyne zavěšení i funkce: byla (Pred)

naplněna (Pnom)

Existuje i jiný pohled, větu lze totiž také parafrázovat - *Vřelá přátelství naplnila naše setkání.*

Odlišné zavěšení i funkce: naplněna (Pred)

byla (AuxV)

- **běžná lidská nepozornost.**

Stromová struktura, kterou anotátor dostává ke kontrole, je vlastně už vybudována. Ve více případech je vybudována dobře. Chyba se v takovém stromě přehlédne velmi snadno.

Podřadící spojka *že* měla anal. funkci - **Adv.** (je to pochopitelně AuxC)

slovu - *všichni* - v kontextu *Všichni vietnamští studenti...* záhadným způsobem zůstala funkce **Atv**, (je to Atr)

## **podzim 2007**

### **1. Rozlišení stavu a pasíva.**

V manuálu na str. 34 jsou sice uvedena kritéria pro takové rozlišení, avšak vedle případů jasných a zřetelných existuje množství situací, kde rozhodování je obtížné a kde je nutno se spoléhat na intuici.

Je-li takový predikát bohatě rozvitý a byla-li intuice anotátorů rozdílná, vzniká něco jako „dominový efekt“, při němž jedno pochybení vyvolává celý řetěz rozdílů v zachycení stromové struktury.

## 2. Nepravé (nevlastní) předložky.

Je to kategorie živá, je v pohybu. Předložkové užití sousloví jako -  
v souvislosti s+I, pod vlivem +G, v poměru k+D atd. se snadno přehledně.

Pochybnosti ani nevzniknou, a proto nemá anotátor žádný důvod ověřovat v seznamu, jestli to je nebo není nepravá předložka. Nicméně toto opomenutí vyvolává také „dominový efekt“, i když mnohem menší než předchozí.

Rovněž se musí dát pozor na jejich nepředložkové užití - z hlediska praxe... ale - z hlediska historického..., - v oblasti vědy... ale - v oblasti se vyskytly povodně.

Je dokonce tendence označovat za nepravou předložku sousloví, která nejsou uvedena jako předložky ani ve Slovníku češtiny ani v interním seznamu nepravých předložek v manuálu, např. -

v asociaci s +I (s18, věta 127), v jednotě s +I (s21, věta 39 a 42)

Menší korekce po konzultaci s J.Pa. - částici **až** lze označovat jako součást předložky, tedy AuxP a zavěšovat ji na předložku, před níž stojí.

Tedy jako u nepravých předložek - až (AuxP) po (AuxP) uzávěrce...

( KK pár nepravých předložek přidala, KG na ně občas zapomíná )

## 3. Objekt nebo adverbium.

Trvalý problém nejen při anotování. Hranice mezi objektem a adverbium není striktní. Je možné, že kritéria, která uplatňujeme v našem přístupu (práce J.Pa), se liší od lingvistické výchovy anotátorů. Je proto nutno ověřovat valenci v interním valenčním slovníku. Ve většině případů taková konzultace pomůže, ale i zde je nutná vlastní intuice.

Příklad : vycházet z budovy (Adv) , vycházet z teorie (Obj) , vycházet z příkladů, vycházet z proslavu ... ?

## 4. Doplněk.

Je třeba rozlišovat **doplněk určující (nevazebný)** - str. 89 manuálu a

**doplněk doplňující (vazebný, rekční)** - str. 60

Ten druhý má jiný způsob zavěšení - je to vlastně Obj a zavěšuje se na sloveso. Ve valenčním slovníku je obvykle uveden jako Eff. Seznam sloves, u nichž se vyskytuje doplněk vazebný, je sice uveden na str. 60, ale je evidentně neúplný. Valenční slovník je mnohem širší.

## 5. Rozdíl mezi AuxZ a AuxY.

Velmi zjednodušená interpretace :

Funkce **AuxZ** byla zavedena pro zvláštní druh adverbia, které se může vyskytovat u substantiv. Jsou to zdůrazňující částice.

Příklad AuxZ : **Již** Komenský zformuloval myšlenku...,

**Teprve** Komenský ....

**Pouze** Komenský ...

**Jenom** Komenský ...

Je jich asi kolem 30, mohou se vyskytnout i před predikátem...

**AuxY** jsou postojová adverbia, patří celé větě a vyskytují se hlavně u sloves.

Příklad : XY **určitě** přijde ..., XY **pravděpodobně** přijde..., ZY **asi** nepřijde

Pochopitelně jsou i homonymní případy.

Větu - *Například v IKEMu usilují o to, aby...* lze rozložit jako :

*Například v IKEMu / usilují o to...*, pak je to **AuxZ** na slově *IKEMu*, nebo *Například / v IKEMu usilují o to...*, a pak je to **AuxY** na predikátu *usilují*.  
Ještě jednou velmi zjednodušeně : **AuxY** by nemělo viset na substantivu. ( i když v jazyce nelze nic vyloučit )

#### 6. **AuxR / AuxT / Obj / Adv**

Např. **SE** u slovesa *ukázat se* může vystupovat ve třech funkcích - Obj, AuxR a AuxT. Někdy je skutečně obtížné se rozhodnout, o kterou funkci jde, ale ignorovat by se to nemělo. Lze ověřovat ve valenčním slovníku - pokud SE je AuxT, mělo by takové sloveso mít samostatné heslo.

Pozor na SI. V případech jako - Koupil si knihu - je to Obj. Ale je třeba si pamatovat, že ne každý dativ je adresát, tedy Obj na analytické rovině.

Může to být také benefaktum ( pro sebe), a pak takovému SI dáváme funkci Adv - omezit si, vymežit si, určili jsme si.

#### 7. **Strukturovaná koordinace**

Obecné pravidlo v manuálu říká, že se koordinační členy zavěšují na poslední nositel koordinace. To platí, pokud všechny členy koordinačního spojení jsou rovnoprávné. Ale ve skutečnosti může mít autor svoje důvody, proč člení koordinaci jiným způsobem, a je nutno k tomu přihlížet.

Velmi zjednodušeně, příklad : *Ve svém díle autor popisuje odvahu a statečnost, zbabělost a zradu, dobro a zlo*. Taková seskupení koordinace by se měla respektovat.

Je-li souvětí např. ze tri a více vět, stává se, že jen dvě mají společný subjekt nebo jiné společné rozvití. Nelze zde zformulovat žádné konkrétní pravidlo, jen si pamatovat, že obsah věty někdy vynutí jiné uspořádání, než pravidlo, že na vrcholu koordinačního spojení je poslední nositel koordinace.

Existují odklony od obecného pokynu.

#### 8. **Složené koordinační spojky**

Jsou sice v manuálu dány seznamem, ale za úplný výčet to považovat nelze.

V koordinačním sousloví - *jak... tak i* - nositelem koordinace bude slůvko TAK. Slova doprovázející koordinaci dostanou funkci AuxY, což znamená, že JAK a I budou mít AuxY.

Podobná situace nastává i u dalších složených spojek - *nejen (AuxY) ale (Coord) i (AuxY)*

Ještě jednu důležitou poznámku : ve větě - *Nejen způsob života ale i kultura pražského dělnictva jsou...* - slova NEJEN a I lze považovat za částice zdůrazňovací (AuxZ) , ale zároveň je lze považovat za součást složeného koordinačního výrazu. Přednost dáváme druhé možnosti, tedy součást koordinační spojky, a proto dáváme jim funkci AuxY. Stejně se postupuje i u jiných složených koordinačních spojek - přednost dáváme

koordinačnímu výrazu.

(tím se omlouvám KK, protože jsem všude opravila AuxZ na AuxY)

#### 9. **Zavěšení AuxZ**

Obvykle se zavěšuje na slovo, před nímž stojí. Ale není tomu tak vždy.

Příklad : *Na slavnosti byl přítomen i anglický vyslanec.*

Zavěšujeme zdůrazňující slůvko *I* na slovo *ANGLICKÝ* nebo na slovo *VELVYSLANEC* ?  
Záleží na předchozím kontextu.

Pokud se mluvilo o slavnosti za přítomností velvyslanců různých zemí, pak se zdůrazňuje slovo *ANGLICKÝ* a *I* se zavěsí na toto slovo.

Pokud se mluvilo např. o slavnosti na ambasádě, kde byli přítomní zaměstnanci, úředníci atd., a pak následuje zmíněná věta, částice *I* patří na slovo *VELVYSLANEC*.

Nedovedeme-li se rozhodnout, dáváme přednost širšímu pojetí a zavěšujeme částici na substantivum.

## 10. Jednotlivosti

- Spojka *NEBOŤ* je koordinační.

- Spojka *PROTO* je považována za koordinační, obvykle v pozici návaznosti na předchozí kontext. Jako spojky *A*, *ALE*, *VŠAK*, *AVŠAK*.

Stojí-li někdy hluboko ve větě, necháme ji tam. Dostane funkci AuxY, může mít rovněž funkci Adv.

- Větné úseky jako např. - *Jak řekl XY, situace se zlepšila.* - považujeme za Pred-Pa.

- Výraz typu - *od ... až do* - nepovažujeme za koordinační

- Výrazy typu - *některých, byť podstatných..., některých, i když podstatných...* - se zavěšují podle „šablony“ na str. 105

- Výraz začínající *AŤ* je ... - šablona na str. 158 (Pozor, také spojka AuxC)

- Výraz začínající *A TO* ... považujeme za volně připojený větný člen, řešení je na str.202. (Pozor, může být také Apos !)

Příklad : *Volby proběhly dobře, a to ve všech pražských obvodech...*

- Pro zachycení srovnání je celá řada „šablon“. Někdy vypadá přijaté řešení poněkud podivně, ale pokud bude zachycení všude stejné, lze to korigovat, pokud se přijde na lepší variantu

## 11. Nepozornost při čtení textu

Bohužel, tzv. hloupé chyby mají velký podíl na neshodách mezi anotátory.