# Workshop korpus znakového jazyka (5)

* Radka Nováková, Lucie Stanovská:
	+ na jednom semináři zkoušely podobně jako my elicitační práci a následné anotování
	+ inspirací pro ně byl polský korpus znakového jazyka
	+ výsledky představily minulý týden (anotovaný úsek dlouhý asi 2,5 minuty)
	+ hodně se zabývaly tím co a jak elicitovat
	+ na příští hodinu přijde Lucie Stanovská a povypráví nám o svých zkušenostech
		- měli bychom si dopodrobna projít nějakou vrstvu
	+ zajímat nás bude oblast slovních druhů (jak se to dělá v polském korpusu) a co použily při práci na svém materiálu
* Přehled slovních druhů, co používaly
	+ „substantivum nebo sloveso“ – elegantní řešení (řešili jsme, že ne pokaždé je to jasně rozpoznatelné, takže je tato kategorie pochopitelná)
	+ absence zájmen a číslovek
		- využívaly pouze inkorporaci číslovek, ale číslovky jako samostatnou kategorii ne
		- abychom znak chápali jako inkorporaci, musí mít svoji syntaktickou platnost
		- v ČZJ bohatě propracovaná struktura číslovek -> do základních slovních druhů není vhodné jednotlivé typy začlenit (je jich mnoho)
	+ pozn.: inkorporace adverbia – co je tím myšleno?
		- můžou to být různé modifikace základního slovesa?
		- např.: pracovat: aktivně × otráveně, psát: pomalu × rychle
	+ jsou přítomny jednotlivé typy sloves a pak je tam ještě kategorie „inkorporace objektu“
	+ -> asi by bylo vhodné značit inkorporaci v samostatné vrstvě (je to otázka pragmatiky a syntaktiky, inkorporace úplně není slovním druhem)
		- jíst (sloveso) + v další vrstvě (jablko, chleba atd.)
		- pozn.: Jaká by byla glosa? Jíst, nebo rovnou Jíst jablko?
			* v té inkorporační vrstvě spíše: označení ano×ne (inkorporace je přítomna × není přítomna)
			* v glose: jíst jablko (byl by to podtyp)
		- snažme se oprostit od sémantiky mluvené češtiny (pro nás je to jedení + něco) ale ve znakovém jazyce – chápeme jako samostatné znaky

Které slovní druhy budeme vyznačovat my?:

* substantivum (SUBST) + vrstva inkorporace (objektu)
* adjektivum (ADJ) + vrstva inkorporace
* zájmena (PRON)
* číslovky (NUM) + vrstva inkorporace
* verbum (VERB):
	+ zatím je nebudeme dále specifikovat
	+ asi je rozdělíme v samostatné vrstvě
* adverbia (ADV) + vrstva inkorporace
* konjunkce (KON)
* prepozice (PREP)
	+ pozn.: v Polsku nezaznamenávají vůbec spojky, předložky, částice
* partikule (PART)
	+ pozn.: pragmatické prostředky?
	+ Jak rozlišit gesto + např. citoslovečné ANO-NE –> budeme to začleňovat jako běžné částice?)
* pomocná slovesa (AUX)
	+ bude, bylo
* specifické znaky (SPEC)
* nepřímá pojmenování (INU)
	+ Budeme je vyčleňovat? Šli by zařadit do jiné kategorie (např. substantivum)
	+ objevuje se u nich konverze (např. HROB, NEMOC – nebude typické substantivum)
* klasifikátor (KLAS)
	+ je vždy součástí nějakého slovesného vyjádření × je lineárně identifikovatelný
	+ (AUTO – JEDE -> jak vyčleníme?)
	+ je otázkou, jestli ještě kromě inkorporace nevyčleňovat přítomnost klasifikátorů v samostatné vrstvě
		- pozn.: polský korpus vůbec neřeší
		- můžeme mít podtyp u inkorporace jako inkorporace klasifikátoru
	+ jak to řeší polský korpus?
		- všechno je to na rovině glos
		- je dobré, že si zkoušíme znaky začleňovat do jednotlivých kategorií
		- použijeme-li více vrstev, máme několik rovin obecnosti
			* má to přínos pro vyhledávací funkci v korpusu
			* výhodné i přes riziko, že v jedné obecnější vrstvě bude větší množství nejasných příkladů
* Specifikátor (SPEC)
* Prstová abeceda (SPELL)
* Inicializované znaky (INIC)
	+ chápána pro ty výrazy, co nemají samostatný znak a namístě využijeme inicializaci
	+ námitka: inicializace: dělá se to podle formy, nesedí to do roviny slovních druhů -> začlenit inicializaci do nové zpřesňující vrstvy?
		- pak by se ale nevyčleňovala ani prstová abeceda
* Úkol do příště:
	+ Popřemýšlet o následujících oblastech:
		- jak značit inkorporaci
		- klasifikátor + specifikátor,
		- prstová abeceda + inicializace,
		- verba (rozdělovat na druhy? Jak?)

Práce s ELANEM

* vyšla nová vrstva ELANU
* Zopakování správného postupu při vytváření nových vrstev a vztahů mezi nimi:
* 1) vytvoření řízeného slovníku
* 2) vytvoření typu (**TYPE**) – slouží k propojení slovníku a vrstvy,
	+ 3) přiřazení vrstvy k typu a během tohoto přirazení nastavení STEREOTYPU
		- nastavuje se během druhého kroku (vytváření TYPU)
		- stereotype: symbolic association
		- 
* 4) vrstva matka (PARENT TIER) – hlavní, která nemá žádnou nadřazenou vrstvu
	+ v našem případě vrstva glos pro každou ruku zvlášť (dvě základní)
	+ jako mateřskou vrstvu považujeme každou, která má podřazené nějaké další
	+ 
* 5) přirazení TYPu k mateřské vrstvě
	+ teď přeskočíme
* 6) dceřiná vrstva
* 
* 7) přirazení typu k dceřiné vrstvě
* 8) nastavení stereotypu (to nám označí, jakým způsobem se dceřiná vrstva váže k té mateřské)

Máme 5 druhů stereotypu:

* none:
	+ (žádný) vrstvy jsou na sobě nezávislé
* time subdivision:
	+ anotace v dceřiné vrstvě kopírují ty anotace ve vrstvě mateřské, ale můžeme je dále segmentovat, ale pouze na další stejně veliké intervaly
* symbolic subdivison:
	+ stejné hranice, ale možnost libovolné segmentace
* included in:
	+ můžou se vytvořit anotace libovolně dlouhé, ale pouze v mezích té původní anotace
* symbolic associtaion:
	+ anotace v dceřiné vrstvě plně kopíruje anotaci z mateřské vrstvy (má stejné hranice)

Dotazy:

* Jak udělat, když mám hranice znaku pro pravou ruku, ale jedná se o dvouruční znak –> jak převést hranici znaku do vrstvy pro levou ruku?
	+ vrstvy nejsou propojené, ale anotaci můžeme v přesném rozmezí zkopírovat
	+ anotace se označí, zobrazí se vybraná část (modrý pruh), do požadované kliknout pravým tlačítkem a dát **nová anotace** a anotace se zkopíruje
* prodloužení anotační vrstvy
	+ kliknu na místo kam až chci hranici prodloužit a použiji klávesovou zkratku ctrl+ enter
* AUTO JELO a ZASTAVILO (zastavení posunutím ruky): bude se rozdělovat jako dva samostatné znaky?
	+ buď by se dalo řešit vrstvou pro komentáře
	+ charakter toho pohybu může být popsán ve vrstvě pro lexikálně sémantické rysy toho slovesa
	+ vrstva pro komentáře má velký význam: když se třeba bude řešit, jestli vytvářet novou vrstvu, nebo modifikovat tu stávající, tak by ve vrstvě pro komentáře měly být přítomny všechny sporná místa
	+ záložka komentáře:
	+ 
* kam se znaky, co byly řečeny omylem (opravy)
	+ něco jiného (přeřeknutí – špatné nastavení ruky) a něco na konci mě napadne lepší (oprava formy)
	+ spíše na úrovni pragmatiky (nová vrstva pro opravy, a chyby)
* Původně vytvářené vrstvy? Co s nimi? (Ty co jsme si do teď zkoušeli) -> Elan nabízí možnost vrstvy skrýt

Na příští hodinu:

* dobré, abychom začali všichni dělat na narativech (vyprávění na základě zhlédnutého videa – „auto“)
	+ důležité, že budeme pracovat na podobném materiálu
	+ Pozn.: Honza Š. si bude muset nějaký narativ stáhnout z uložiště
* budeme pokračovat s tím, že budeme mít všichni stejné vrstvy a to:
	+ Gloss-R
	+ Gloss-L
	+ PoS-R
	+ PoS-L
* ideálně bychom měli zpracovat nějakou souvislou část promluvy (alespoň 30 sekund)
* zaměřit se na to, co nám připadá jako problematická místa
	+ sdělit je ostatním
* rezignujme na přesně hranice znaků
	+ radši větší množství glos s menší přesností (hranice znaku lze posléze upravit)
	+ vyzkoušení si určování slovních druhů (Dostačující? S čím je problém?)
* v případě, kdyby anotace byly totožné (dvouruční znak) – zatím bude stačit oanotovat pouze dominantní ruku
* když jednoruční znak u pasivní ruky nic nevyznačuji
* simultánní konstrukce: např. normálně znak dvouruční, ale z důvodu simultaneity použita pouze jednoruční podoba –> musíme pracovat s aktuální promluvou,
	+ když byl znak použit jako jednoruční –> přiřazen jako jednoruční
* překrývaní znaků: zůstává-li nám např. pasivní ruka ve tvaru a umístění do následujícího znaku
	+ bude u pasivní ruky prodloužená hranice
	+ u aktivní ruky hranice pouze pro aktivní ruku -> ačkoliv dvouruční znak, tak se hranice nebudou stoprocentně překrývat