

Univerzita Karlova v Praze

Filozofická fakulta

Ústav českého jazyka a teorie komunikace

Filologie – Český jazyk

M a g d a l é n a R y s o v á

**Diskurzní konektory v češtině
(Od centra k periferii)**

**Discourse Connectives in Czech
(From Centre to Periphery)**

Disertační práce

Vedoucí práce: prof. PhDr. Eva Hajičová, DrSc.

2015

Srdečně děkuji své školitelce prof. PhDr. Evě Hajičové, DrSc. za odborné vedení předkládané disertační práce, cenné připomínky, trpělivost a přátelský přístup. Zároveň děkuji RNDr. Jiřímu Mírovskému, Ph.D. za ochotnou technickou pomoc při zpracovávání dat *Pražského závislostního korpusu* a za změření anotátorské shody. Za pomoc s anotací jazykových dat vřele děkuji Mgr. Pavlíně Jínové a Bc. Veronice Pavlíkové.

Magdaléna Rysová

Předkládaná práce vznikla s podporou grantu Grantové agentury České republiky P406/12/0658 *Koreference, diskurs a aktuální členění v kontrastivním pohledu*, grantu Ministerstva školství, mládeže a tělovýchovy LH14011 *Vícejazyčná korpusová anotace jako podpora jazykových technologií řešených na Matematicko-fyzikální fakultě Univerzity Karlovy v Praze* a s podporou grantu Grantové agentury Univerzity Karlovy v Praze č. 36213 *Diskurzivní konektory v češtině řešeného na Filozofické fakultě Univerzity Karlovy v Praze*. Práce zároveň využívá jazykové zdroje získané v rámci projektu LINDAT-Clarin (projektu Ministerstva školství, mládeže a tělovýchovy České republiky; LM2010013).

Prohlašuji, že jsem disertační práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Sedlici dne 25. května 2015

Magdaléna Rysová

Klíčová slova

diskurz, diskurzní konektory, diskurzní vztahy, Pražský závislostní korpus, text

Key words

discourse, discourse connectives, discourse relations, Prague Dependency Treebank, text

Anotace

V práci se zabýváme popisem a analýzou diskurzních (tj. textových) konektorů v češtině v širším smyslu, tedy tím, jakými jazykovými prostředky je možné vyjadřovat v textu diskurzní vztahy. Výzkum přitom neomezujeme na předem stanovenou skupinu výrazů (danou například příslušností k určitým slovním druhům, jako jsou spojky či strukturující částice), ale snažíme se nalézt a obecně popsat všechny jazykové prostředky v češtině, které mají schopnost spojovat jednotlivé úseky textu v jeden koherentní celek. Zaměřujeme se především na méně probádané víceslovné konektivní struktury typu *to je důvod, proč; kvůli těmto skutečnostem; z těchto důvodů* atd., pro které užíváme označení **sekundární konektory** (za primární konektory považujeme především konektivní synsémantika typu *však, nebo, a, ale, proto* apod.).

Diskurzní konektory jsou velmi obecně chápány jako jazykové výrazy, které mají v textu schopnost signalizovat různé sémanticko-pragmatické vztahy. Existuje ovšem mnoho jazykových teorií, které se v konkrétním popisu diskurzních konektorů výrazně liší. Neexistuje tedy jediná, obecně přijímaná definice těchto výrazů, o jejichž podstatě a charakteristice se v lingvistice vedou diskuze. Cílem této práce je na základě podrobného jazykového výzkumu přispět k obecnému vymezení a definici diskurzních konektorů (nejen pro češtinu).

V naší práci nejprve popisujeme jazykové teorie zabývající se problematikou diskurzu (či textu). Zaměřujeme se přitom na jednotlivá pojetí diskurzních konektorů a na obecná kritéria, podle kterých můžeme tyto prostředky v jazyce definovat.

Podrobná analýza diskurzních konektorů (s důrazem na konektory sekundární) pak vychází především z anotovaných dat Pražského závislostního korpusu (PDT) obsahujícího psané publicistické texty, okrajově z Českého národního korpusu – z korpusu SYN2013PUB. V práci představujeme principy a výsledky anotace sekundárních konektorů v PDT (provedené pro účely této práce). Zaměřujeme se především na srovnání primárních konektorů a konektorů sekundárních podle četnosti, podle významových vztahů, které v textu vyjadřují, nebo podle toho, zda se vyskytují častěji mezivětně či vnitrovětně. Dále přinášíme srovnání anotace sekundárních konektorů v češtině a obdobných jazykových prostředků v angličtině (provedené na datech pensylvánského korpusu Penn Discourse Treebank) a stručně srovnáváme výskyt a užívání sekundárních konektorů v psané a mluvené češtině (s využitím korpusů PDT a DIALOG).

Z výsledků analýzy diskurzních konektorů se snažíme vyvodit o zkoumaných jazykových prostředcích obecné teoretické závěry. Zaměřujeme se přitom na definici diskurzních konektorů, stanovení principů, na základě kterých je možné konektory rozpoznat v textu, či na vymezení hranic mezi konektory primárními a sekundárními (především na základě jejich lexikální, syntaktické a sémantické charakteristiky).

Podrobný výzkum diskurzních konektorů a obecné závěry představené v této práci by mohly vést ke zlepšení porozumění významovým vztahům v textu, a tedy k většímu porozumění zákonitostem výstavby textu jako celku.

Annotation

The thesis focuses on description and analysis of discourse connectives in Czech in broader sense, i.e. by which language means it is possible to express sense relation within a text. The thesis is not limited to any parts of speech (like conjunctions or structuring particles) but it tries to find and describe all language means in Czech with the ability to connect two pieces or units of a text into one coherent complex. The thesis investigates discourse connectives in Czech with respect to the so called **secondary connectives** (i.e. mainly multiword phrases like *to je důvod, proč* – *that is the reason why*; *kvůli těmto skutečnostem* – *due to these facts* etc., in opposition to primary connectives like *však* – *however*, *nebo* – *or*, *a* – *and*, *ale* – *but*, *proto* – *therefore* etc.).

Discourse connectives are (in general terms) understood as language expressions that signal semantico-pragmatic relations within a text. However, there are many theories that significantly differ in the concrete description of these expressions. Therefore, there is not a generally accepted and universal definition of discourse connectives and their description and characteristics is still a matter of linguistic discussion. The aim of this thesis is to contribute to the general delimitation and definition of discourse connectives (not only in Czech) on the basis of a detailed language research.

Firstly, we describe various language theories focusing on description of discourse (or text). We concentrate especially on the individual approaches to discourse connectives and on the general criteria according to which we may define these expressions in language.

The detailed analysis of discourse connectives (with the emphasis on secondary connectives) is based especially on the annotated data of the Prague Dependency Treebank (PDT) containing written newspaper texts, marginally on the data of the Czech National Corpus (concretely SYN2013PUB). In the thesis, we introduce the principles and results of the annotation of secondary connectives in PDT (carried out for purposes of this thesis). We focus mainly on the comparison of primary connectives and secondary connectives according to their frequency in authentic texts, according to sense relations they express or according to whether they occur more often inter- or intra-sententially. We compare then the annotations of secondary connectives in Czech and the similar expressions in English (carried out on the data of the Pennsylvanian corpus Penn Discourse Treebank) and we shortly compare the occurrence and usage of secondary connectives in written and spoken Czech (in PDT and in the corpus DIALOG).

On the basis of the practical analysis of discourse connectives, we try to draw several theoretical conclusions about these expressions. We focus mainly on the definition of discourse connectives, on general principles describing how it is possible to detect these expressions in authentic texts, and on the delimitation of the boundaries between primary and secondary connectives (mainly with respect to their lexical, syntactic and semantic characteristics).

The detailed analysis of discourse connectives and the general conclusions presented in this thesis might improve the understanding of semantico-pragmatic relations in texts and in this sense also the understanding of a text structure as a whole.

Obsah

Anotace	6
Annotation	8
Obsah	10
Seznam zkratk	13
1 Úvodem	14
2 Cíle práce	16
3 Teoretické východisko	21
3.1 Kořeny textové lingvistiky ve světové literatuře	21
3.2 Kořeny textové lingvistiky v české literatuře	24
3.2.1 Diskuze o nadvětném útvaru – vymezení pojmu diskurz	27
3.2.1.1 Strukturní přístup k diskurzu	29
3.2.1.2 Funkční přístup k diskurzu	30
3.2.1.3 Diskurz jako fenomén přesahující využití jazyka v praxi	32
3.3 Co mohou spojovat diskurzní konektory	34
3.3.1 Teorie koherence	34
3.3.2 Teorie relevance	34
3.3.3 Diskurzní konektory z pohledu teorie koherence a teorie relevance	35
3.4 Vymezení jednotky diskurzu	38
3.4.1 Diskurzní argumenty v zachycení Pražského závislostního korpusu	39
3.4.2 Povrchová realizace diskurzních argumentů – verbální vs. nominální argumenty	40
3.5 Diskurzní vztahy	42
3.5.1 Diskurzní vztahy implicitní a explicitní	42
3.5.1.1 Pokusy o zpracování implicitních diskurzních vztahů ve velkých korpusech	43
3.5.2 Významová klasifikace diskurzních vztahů	44
3.5.2.1 Významové vztahy v textu v Pražském závislostním korpusu	47
3.6 Diskurzní konektory	54
3.6.1 Diskurzní konektory ve světové literatuře	54
3.6.2 Diskurzní konektory v české literatuře	59
3.7 Definice a dělení diskurzních konektorů	63
3.7.1 Syntaktická povaha diskurzních argumentů	64
3.7.2 Princip univerzálnosti – diskurzní konektory vs. volné konektivní fráze	68
3.7.3 Primární a sekundární konektory	71
3.7.3.1 Primární konektory	72
3.7.3.2 Sekundární konektory	73
3.7.3.3 Prostupnost hranic mezi primárními a sekundárními konektory	74
3.7.3.4 Úvahy ke vzniku primárních konektorů	75
3.7.3.5 K otázce plnovýznamovosti	78
3.7.3.6 Vzájemná nahraditelnost primárních a sekundárních konektorů	80
3.7.3.7 Souvýskyt primárních a sekundárních konektorů – složené konektory	83
3.7.4 Zpracování diskurzních vztahů a „sekundárních konektorů“ ve velkých korpusech	88
3.7.4.1 RST Discourse Treebank	88
3.7.4.2 Discourse Graphbank	89
3.7.4.3 Penn Discourse Treebank a navazující projekty	90
3.7.4.4 Další významné diskurzní korpusy	91
3.7.4.5 Zpracování „sekundárních konektorů“ ve velkých korpusech	91

4 Jazykový materiál	95
4.1 Pražský závislostní korpus	95
4.1.1 Anotace primárních konektorů v Pražském závislostním korpusu	98
4.2 Korpus DIALOG	101
4.3 Český národní korpus	102
5 Metody a postup práce	103
5.1 Anotace sekundárních konektorů v Pražském závislostním korpusu	104
5.1.1 Odlíšnosti od anotace primárních konektorů a další teoretické otázky a problémy	108
5.1.1.1 Anotace nominálních argumentů	108
5.1.1.2 Rozsah sekundárních konektorů	108
5.1.1.3 Hranice mezi sekundárním konektorem a argumentem	109
5.1.1.4 Dvouúrovňová anotace sekundárních konektorů	110
5.1.2 Anotační principy pro zpracování sekundárních konektorů (a volných konektivních frází) v PDT	113
5.1.2.1 Diskurzivní vztah mezi verbálními argumenty	113
5.1.2.2 Diskurzivní vztah mezi argumenty, z nichž je alespoň jeden nominální	113
5.1.2.3 Volné konektivní fráze	117
5.1.2.4 Slovesné sekundární konektory	119
5.1.2.5 Nové vztahy	121
5.2 Evaluace anotace	122
6 Výsledky práce	126
6.1 Charakteristika sekundárních konektorů na základě analýzy nově anotovaných dat PDT	126
6.1.1 Lexikální charakteristika sekundárních konektorů	128
6.1.2 Syntaktická charakteristika sekundárních konektorů	129
6.1.3 Sémantická charakteristika sekundárních konektorů	133
6.2 Porovnání výsledků anotace sekundárních konektorů v češtině a altlexů v angličtině na základě dat PDT a PDTB	135
6.3 Rozdíly mezi primárními a sekundárními konektory na základě analýzy dat PDT	137
6.4 Srovnání primárních a sekundárních konektorů podle četnosti; slovesa říkání s konektivní funkcí	139
6.5 Sémantické typy diskurzivních vztahů vyjádřené primárními a sekundárními konektory	144
6.5.1 Struktury typu <i>výjimkou je, příkladem je</i>	149
6.6 Rozšíření klasifikace diskurzivních vztahů vyjádřených sekundárními konektory	151
6.6.1 Vztah vyplývání (či vyvození výsledků)	151
6.6.2 Vyvození závěru	152
6.6.3 Zřetel	152
6.7 Diskurzivní vztahy mezivětné a nitrovětné	153
6.8 Vybrané sekundární konektory v psané a mluvené češtině	161
6.9 Testování univerzálnosti sekundárních konektorů	167
6.9.1 Stanovení „hranice univerzálnosti“ pro diskurzivní konektory	173
6.9.2 Testování předložkových frází typu <i>díky tomu</i> a <i>na základě těchto skutečností</i>	181
7 Závěry	185
7.1 Definice diskurzivních argumentů a konektorů	185
7.1.1 Charakteristika primárních a sekundárních konektorů	186
7.1.2 Hranice mezi primárními a sekundárními konektory	190
7.1.3 Vznik primárních konektorů	192
7.1.4 Oslabování lexikálního významu u sekundárních konektorů	192
7.2 Výsledky a evaluace anotací PDT – primární konektory, sekundární konektory, volné konektivní fráze	193
7.3 Nejčastěji používané primární a sekundární konektory v PDT	195
7.4 Srovnání sémantických typů diskurzivních vztahů pro primární a sekundární konektory	196

7.4.1 Možnosti rozšíření klasifikace diskurzivních vztahů vyjadřovaných sekundárními konektory.....	197
7.5 Srovnání diskurzivních konektorů v rámci mezivětných a vnitrovětných vztahů	198
7.6 Srovnání sekundárních konektorů v psané a mluvené češtině.....	199
7.7 Testování vhodnosti sekundárních konektorů pro různé kontexty – princip univerzality	200
Závěrem	204
Literatura.....	205
Přílohy	215
Příloha 1: Seznamy konektorů	216
Příloha 1.1: Abecední seznam forem primárních konektorů z Pražského závislostního korpusu.....	220
Příloha 1.2: Frekvenční seznam forem primárních konektorů z Pražského závislostního korpusu	224
Příloha 1.3: Abecední seznam forem sekundárních konektorů z Pražského závislostního korpusu.....	229
Příloha 1.4: Frekvenční seznam forem sekundárních konektorů z Pražského závislostního korpusu	233
Příloha 1.5: Frekvenční seznam forem primárních a sekundárních konektorů z Pražského závislostního korpusu s ohledem na vyjadřované významové vztahy.....	238
Příloha 2: Ukázka textu z Pražského závislostního korpusu s anotací významových diskurzivních vztahů vyjádřených konektory	262

Seznam zkratk

#neg	větná negace
altlex	alternativní lexikální vyjádření diskurzního konektoru
C. k.	Cohenova kappa
CONC	přípustka
COND	podmínka a výsledek podmínky
CONFR	konfrontace
CONJ	konjunkce
CONJALT	konjunktivní alternativa
CORR	rektifikace
DISJALT	disjunktivní alternativa
EQUIV	ekvivalence
EXEMPL	exemplifikace
EXPLICAT	explikace
F_COND	nepravá podmínka a výsledek podmínky
F_OPP	nepravý kontrast
F_REASON	nepravá příčina a důsledek
GENER	generalizace
GRAD	gradace
NP	nominální fráze
OPP	opozice
PDT	Prague Dependency Treebank, Pražský závislostní korpus
PDTB	Penn Discourse Treebank
PRECED	nesoučasnost (předčasnost a následnost)
PURP	účel
REASON	příčina a důsledek
RESTR	restriktivní opozice
RST	Rhetorical Structure Theory
SDRT	Segmented Discourse Representation Theory
SPEC	specifikace
SYNCHR	současnost
VP	slovesná fráze

1 Úvodem

Porozumění psaným textům je úzce spojeno s komunikačními potřebami všech uživatelů jazyka. Zároveň bezprostředně souvisí s vyrozuměním vzájemných vztahů mezi jednotlivými úseky či částmi textu. Předkládaná disertační práce se proto zaměřuje na to, jaké jazykové prostředky v češtině mají schopnost vztahy mezi jednotlivými částmi textu vyjadřovat a pomáhat tak čtenáři lépe a rychleji textu porozumět.

Podle mnohých autorů (srov. např. Halliday a Hasanová, 1976) mají lidé přirozenou schopnost odlišit od sebe ucelený text a náhodný shluk vět na sebe nenavazujících; jinými slovy, existují různé atributy, které dělají text textem. Mezi tyto významné atributy koherentního textu patří významová soudržnost a srozumitelnost, tj. jednotlivé věty textu na sebe navazují a jsou propojeny různými významovými vztahy. Obecně tedy můžeme říci, že souvislý text představuje síť vzájemných vztahů, ve které žádný úsek nezůstává osamocen.

Tyto vztahy pak mohou být v textu vyjádřeny explicitně pomocí konkrétních jazykových prostředků (srov. např. *proto* v příkladu *Mám rýmu. Nepůjdu proto do školy.*) nebo implicitně, tj. daný diskurzí vztah není vyjádřen jazykovým výrazem, ale pouze vyplývá z významu příslušných úseků textu (srov. např. *Mám rýmu. Nepůjdu do školy.*).

Mnohé jazykové studie (srov. např. Irwinová a Pulverová, 1984; Dolková, 2010) upozorňují na skutečnost, že lidé lépe rozumí textům, ve kterých jsou vzájemné vztahy mezi jednotlivými úseky vyjádřeny explicitně, pomocí konkrétních výrazů. Je tedy zřejmé, že je zapotřebí daným jazykovým prostředkům věnovat pozornost.

Tyto prostředky, nazývané nejčastěji diskurzí konektory, bývají mnohdy chápány jako synsémantické výrazy (srov. např. Halliday a Hasanová, 1976; Martin, 1992; Knott, 1996; Stede a Neumann, 2014;), tj. často bývají omezeny na určité slovní druhy (především spojky typu *nebo*, *ale*, *a* či částice typu *totiž*, *zkrátka*).

Vzájemné diskurzí vztahy mezi jednotlivými úseky textu ovšem nemusí být vyjádřeny pouze spojkami či některými druhy částic. Ukazuje se, že škála výrazů s konektivní funkcí je mnohem bohatší (srov. např. Hoffmannová, 1983, 1984; Prasadová a kol., 2010; Rysová, 2012, 2013b, 2015; M. Rysová a K. Rysová, 2014). Například diskurzí příčinně-důsledkový vztah nemusí být vyjádřen pouze výrazy typu *proto* nebo *a tak*, ale i víceslovnými spojeními typu *z tohoto důvodu*; *to je důvod, proč*; *důsledkem je* atd. Podobně mohou být vyjadřovány i další diskurzí vztahy, srov. např. spojení *podmínkou je* pro významový vztah podmínky, *to znamená* pro vztah explikace apod.

Předmětem naší práce je tedy podrobně popsat (na základě anotovaných korpusových dat) tyto „jiné“ jazykové prostředky (většinou víceslovné struktury), které v mnohých pojetích nebývají chápány jako diskurzní konektory, ale které spojovací funkci v rámci diskurzu zcela jistě mají a množinu diskurzních konektorů podstatně obohacují.

Problematika diskurzu a výrazů vyjadřujících diskurzní vztahy je v současné době aktuálním tématem na poli světové lingvistiky. Mnohé studie (srov. např. Sanders a Noordman, 2000; Blakemoreová, 2002; Prasadová a kol., 2010; Johannsen a Søgaard, 2013; apod.) se zabývají právě otázkou, na základě jakých obecně platných kritérií můžeme diskurzní konektory definovat.

V českém prostředí se problematikou diskurzu a ukazateli diskurzních vztahů zabývali a zabývají zejména Josef Hrbáček (1994), Jana Hoffmannová (např. 1983, 1984), Jarmila Tárnyiková (2002) či Ondřej Pešek (2011).

Předkládaná disertační práce se snaží na tyto studie navázat a obohatit dosavadní výzkum o analýzu anotovaných korpusových dat. Domníváme se, že detailní analýza jazykových prostředků vyjadřujících diskurzní vztahy v češtině (chápaných v širokém smyslu) provedená na rozsáhlých jazykových datech může být pro českou lingvistiku přínosem.

2 Cíle práce

Diskurzí vztahy mohou být v textu vyjadřovány různými způsoby. Obecně můžeme vyčlenit dva základní typy diskurzích vztahů – implicitní¹ (nevyjádřené konkrétními jazykovými prostředky, ale pouze vyplývající z obsahu daných výpovědí či vět, srov. příklad 1²) a explicitní (vyjádřené konkrétními jazykovými prostředky, srov. příklady 2–13). Škála těchto jazykových prostředků je přitom značně široká, viz následující příklady.

- (1) *Petr je nemocný a nemůže jet. Výlet se ruší.*
- (2) *Petr je nemocný a nemůže jet. Proto se výlet ruší.*
- (3) *Petr je nemocný a nemůže jet. Výlet se tedy ruší.*
- (4) *Petr je nemocný a nemůže jet. Výlet se tak ruší.*
- (5) *Petr je nemocný a nemůže jet. Výlet se tím ruší.*
- (6) *Petr je nemocný a nemůže jet. Výlet se tímto ruší.*
- (7) *Petr je nemocný a nemůže jet. Na základě toho se výlet ruší.*
- (8) *Petr je nemocný a nemůže jet. Z tohoto důvodu se výlet ruší.*
- (9) *Petr je nemocný a nemůže jet. Kvůli tomu se výlet ruší.*
- (10) *Petr je nemocný a nemůže jet. Kvůli této skutečnosti se výlet ruší.*
- (11) *Petr je nemocný a nemůže jet. Kvůli jeho nemoci se výlet ruší.*
- (12) *Petr je nemocný a nemůže jet. Kvůli němu se výlet ruší.*
- (13) *Petr je nemocný a nemůže jet. Důsledkem je zrušení výletu.*

Ve všech uvedených příkladech je diskurzí vztah příčiny a důsledku³ vyjádřený pokaždé jiným způsobem, pomocí různých výrazových prostředků: *proto, tedy, tak, tím, tímto, na základě toho, z tohoto důvodu, kvůli tomu, kvůli této skutečnosti, kvůli jeho nemoci, kvůli němu, důsledkem je*.

¹ Někteří autoři (srov. např. Taboadová, 2009) ovšem existenci čistě implicitních vztahů zpochybňují – diskurzí „signály“ je podle nich nutné hledat na různých jazykových rovinách. Na výstavbě textu se tak podle Taboadové podílejí např. také slovesný čas, způsob, užití určitého, resp. neurčitého slovesného tvaru apod.

² Příklady uváděné v této práci pochází především z Pražského závislostního korpusu (zejména příklady v kapitolách představujících anotaci dat PDT a její výsledky), v ostatních kapitolách jsou zároveň příklady nalezené internetovým vyhledávačem Google. V některých případech jsou tyto příklady pro lepší přehlednost upraveny, zejména kráceny.

³ V PDT, jehož data jsou v této práci využita, je vztah příčiny a důsledku chápán poměrně široce (nerozlišuje se například mezi příčinou a důvodem). Příčina a důsledek jsou zde chápány jako jeden typ vztahu, liší se pouze směrem, kterým je vztah veden.

V předkládané disertační práci se zabýváme výrazovými prostředky, které v textu vyjadřují diskurzní vztah explicitně. Ne u všech je ale stejná míra explicitnosti vyjádření, resp. míra jeho konkrétnosti (srov. *proto* vs. *kvůli jeho nemoci*). Explicitně vyjádřené výrazové prostředky signalizující diskurzní vztahy tedy chápeme jako škálu s různými stupni abstraktnosti/obecnosti. V této práci se zaměřujeme na to, které typy výrazů ve škále můžeme označovat jako diskurzní konektory a proč (např. zda jsou diskurzním konektorem výrazové prostředky *proto* i *kvůli jeho nemoci*), na základě jakých kritérií či principů je můžeme třídit a klasifikovat a v jaké míře se jednotlivé typy těchto výrazových prostředků podílejí na celkové koherenci textu.

Cílem práce je tedy přispět k popisu výstavby textu v češtině, a to především k popisu jazykových prostředků, které mají schopnost vyjadřovat diskurzní vztahy v textu, tedy k popisu diskurzních konektorů. Práce se přitom zaměřuje na výrazy morfologicky, lexikálně i syntakticky neomezené (typu *důsledkem je, příčinou bylo, to je v rozporu s tím* atd., dané výrazy označujeme sekundární konektory), které v mnohých pojetích (srov. např. Stede a Neumann, 2014; Halliday a Hasanová, 1976; Hrbáček, 1994 atd.) nejsou mezi diskurzní konektory řazeny.

Dílčím cílem práce je srovnat různé teorie a přístupy k diskurzu a k diskurzním konektorům v českém i světovém měřítku a přispět novými poznatky (na základě analýzy rozsáhlých korpusových dat) k obecnému vymezení těchto výrazů.

Výzkum byl uskutečněn na ručně anotovaných datech Pražského závislostního korpusu (PDT) (tj. téměř na 50 000 vět), který v současné době obsahuje i (dosud nepublikovanou) rozšířenou anotaci diskurzních vztahů, tedy vztahů vyjádřených právě sekundárními konektory, tj. výrazy typu *podmínkou je; to je důvod, proč* atd. Dílčím cílem je tuto anotaci (provedenou pro účely naší práce) krátce představit a srovnat její výsledky s výsledky dřívější diskurzní anotace primárních konektorů (srov. Poláková a kol., 2012a), tj. především jednoslovných konektorů typu *ale, proto, zatímco* apod. V práci se tedy pokusíme poukázat na možnosti, které zpracování diskurzních vztahů ve velkém měřítku nabízí, uvést teoretické otázky, které mohou na základě anotací vyvstat, a představit způsoby, kterými je možné je řešit.

Dílčí kroky pro dosažení stanovených cílů:

1. V první části práce se budeme zabývat teoretickým srovnáním českých i zahraničních přístupů ke zpracování diskurzu. Stěžejní otázkou bude vymezení a definice diskurzních konektorů. Zaměříme se tedy především na to, jakým způsobem a podle jakých kritérií je možné tyto výrazy obecně charakterizovat a dělit do různých skupin, které dohromady tvoří (morfologicky, syntakticky i lexikálně) velmi heterogenní množinu výrazů (od jednoslovných spojek typu *ale* přes ustálenější spojení typu *stručně řečeno* po celé klauze typu *důsledek je jednoduchý*) atd.

V práci budeme vycházet z českých i zahraničních studií, které se ve vymezení diskurzních konektorů značně liší. Někteří autoři (srov. např. Halliday a Hasanová, 1976; Martin, 1992; Knott, 1996; Stede a Neumann, 2014) definují diskurzní konektory v užším smyslu – jako uzavřenou skupinu výrazů, které patří mezi neohebné slovní druhy, především spojky, příslovce a částice (např. *protože, i, nebo, také, poté* atd.), jiní (srov. např. Hoffmannová, 1983, 1984; Prasadová a kol., 2010; Asrová a Dembergová, 2013) vymezují konektory v širším smyslu, tj. jako výrazy morfologicky, lexikálně i syntakticky neomezené – zdůrazňují tedy především funkci daných výrazů v textu, ne jejich slovnědruhovou příslušnost.

V předkládané disertační práci se pokusíme na základě anotovaných korpusových dat přispět k této diskusi novými poznatky, podle kterých by bylo možné jazykové prostředky vyjadřující diskurzní vztahy v textu lépe vymežit. Pokusíme se tedy najít kritéria, podle kterých lze například stanovit škálu diskurzních konektorů podle míry jejich gramatikalizace, vhodnosti jejich užití v různých kontextech apod.

2. V práci dále představíme, jakým způsobem jsme zpracovávali sekundární konektory v Pražském závislostním korpusu. Na základě studia české i zahraniční literatury (především jednotlivých pojetí diskurzních konektorů) provedeme nejprve anotaci vzorku dat PDT, na kterém budeme sledovat chování sekundárních konektorů v reálných textech. Na základě zkušeností s autentickými texty se pokusíme stanovit obecné anotační principy, podle kterých budeme sekundární konektory zpracovávat v celém PDT (tj. na přibližně 50 000 vět; anotace budou provádět tři anotátoři). Následně proběhne kontrola anotací a měření anotátorské shody, na základě které bude možné anotaci vylepšovat do konečné podoby. (Podobná anotace v češtině nebyla dosud provedena, a pokud je nám známo, nebyla provedena ani jinde v zahraničí – tyto výrazy do jisté míry zachycuje pouze pensylvánský korpus Penn Discourse Treebank, ovšem ne v takovém měřítku jako Pražský závislostní korpus – viz kapitola 3.7.4.5.)

Průběžným cílem této části výzkumu je tedy na základě zkušenosti s reálnými texty sestavit anotační principy pro zpracování sekundárních konektorů ve velkých korpusech (v našem případě v korpusu PDT) a tyto principy představit.

3. Na získaných datech PDT provedeme vlastní výzkum diskurzních konektorů v češtině, zaměřený na konektory sekundární.

a) Na základě analýzy získaných jazykových dat se pokusíme sestavit charakteristiku sekundárních konektorů po stránce lexikální, syntaktické a sémantické (ve srovnání s primárními konektory), pokusíme se popsat jejich vlastnosti a chování v reálných textech a dále obecně vymezit kategorii diskurzních konektorů a jasně definovat rozdíly mezi primárními a sekundárními konektory.

b) V práci dále srovnáme výsledky anotace sekundárních konektorů v češtině a obdobných jazykových prostředků v angličtině (na datech korpusu Penn Discourse Treebank – PDTB). Korpusy PDT a PDTB obsahují publicistické texty, počet anotovaných vět je také srovnatelný (PDT obsahuje přibližně 50 000 vět, PDTB 49 000 vět). Oba korpusy tak nabízejí vhodný jazykový materiál pro srovnávací studii.

c) V práci představíme podrobnou analýzu sekundárních konektorů z hlediska jejich frekvence v reálných textech – zaměříme se především na rozmanitost jednotlivých struktur (tj. např. zda se v češtině častěji používá *důvod je*, *důvodem je* nebo *důvod spočívá*), jaké sekundární konektory se v češtině používají nejčastěji a s jakou frekvencí (opět ve srovnání s primárními konektory), jaké zvláštnosti vykazují některé sekundární konektory apod.

d) Dále se zaměříme na to, jaké významové vztahy v textu vyjadřují sekundární konektory – jaké typy vztahů sekundární konektory vyjadřují nejčastěji, jaké méně často, zda sekundární konektory vyjadřují stejné významové vztahy jako konektory primární, zda existují významové vztahy častěji vyjadřované sekundárními konektory než konektory primárními, příp. za jakých podmínek.

e) V práci dále popíšeme, v jaké míře vyjadřují sekundární konektory diskurzní vztahy mezivětné, v jaké míře vnitrovětné, a zaměříme se na faktory, které mohou mezivětné či vnitrovětné vyjádření diskurzního vztahu ovlivnit (jako je např. možná souvislost

s významovými vztahy v textu apod.). Výsledky analýzy opět porovnáme s primárními konektory.

f) Dále se pokusíme srovnat výskyt, charakteristiku a chování vybraných sekundárních konektorů v psané a mluvené češtině (na základě dat PDT a korpusu mluvené češtiny DIALOG).

g) Na závěr v práci představíme možnosti vyhodnocování míry vhodnosti vybraných diskurzních konektorů jako ukazatelů diskurzních vztahů pro velké množství různých kontextů. Jinými slovy, pokusíme se testovat vybrané konektory v různých kontextech a zjišťovat, pro jaké procento z nich jsou dané konektory vhodnými jazykovými prostředky pro vyjádření příslušného diskurzního vztahu. Obecně se tedy zaměříme na to, jaké faktory ovlivňují výběr vhodného konektivního prostředku, tj. podle jakých kritérií mohou mluvčí volit příslušný výraz, jaká existují možná omezení tohoto výběru apod.

Předkládaná disertační práce by tak měla na základě podrobné analýzy současného jazykového materiálu, a to psaného i mluveného, přinést nové poznatky o možnostech vyjadřování diskurzních vztahů a celkově přispět k prohloubení popisu výstavby textu v češtině.

3 Teoretické východisko

V následující části práce představíme hlavní teoretické přístupy k textu a jeho výstavbě v českém a světovém měřítku. Blíže se zaměříme na vztah mezi pojmy text a diskurz⁴ a na představení dalších zásadních termínů, které úzce souvisejí s hlavním tématem této práce, tj. s diskurzními konektory. Budeme se tedy zabývat především vymezením základní jednotky diskurzu (v souvislosti s obecnou diskuzí o tom, jaké entity jsou spojovány diskurzními konektory), klasifikací diskurzních významových vztahů a vlastním vymezením diskurzních konektorů.

3.1 Kořeny textové lingvistiky ve světové literatuře

Za jednoho z nejvýznamnějších zakladatelů textové lingvistiky bývá obvykle považován holandský vědec Teunen A. van Dijk – srov. např. jeho díla *Some aspects of text grammar* (1972), *Pragmatic connectives* (1979), *Textwissenschaft* (1980b), *Studies in the pragmatics of discourse* (1981) aj.

V sedmdesátých letech se van Dijk zabývá především popisem textových struktur a zaměřuje se na textovou sémantiku (u popisu diskurzních konektorů pak na jejich sémantické a pragmatické aspekty). Jeho hlavním tématem je tzv. textová gramatika (v originále *text grammar*, v kontrastu ke gramatikám větným – tzv. *sentence grammar*) sledující gramatické jevy a souvislosti mezi nimi v rámci větších celků. V tomto ohledu upozorňuje na rozdíl mezi „formálními konstrukty gramatik“ a reálnými texty – srov.: „Notice however that SENTENCES, strictly speaking, are formal constructs of a grammar and not empirical entities“ (van Dijk, 1979: 3).

Podle van Dijka by proto neměla být základní jednotkou lingvistického zkoumání věta, ale text – protože přirozená řeč se většinou neskládá pouze z jedné věty, ale sekvence vět, které dohromady tvoří diskurz – srov.: „In our opinion, such sentence-based descriptions of the structures underlying utterances are inadequate. Sentences, just as clauses or syntagms, have less significant empirical, e.g. psychological, reality than texts. Only in some cases a sentence may correspond to the empirical unit of the utterance. We therefore introduce the

⁴ Vědní disciplína věnující se studiu textu z lingvistického pohledu je v evropské tradici většinou označována jako textová lingvistika (*text linguistics*). V americké tradici se používá spíše pojem analýza diskurzu (*discourse analysis*).

concept of TEXT as the basic linguistic unit manifesting itself, as DISCOURSE, in verbal utterances.“ (van Dijk, 1979: 3)

Velkou pozornost věnuje van Dijk studiu tzv. makrostruktur, tj. hloubkovým obsahovým strukturám textu – na začátku osmdesátých let vychází např. jeho kniha *Macrostructures. An interdisciplinary study of global structures in discourse, interaction, and cognition* (1980a). Studium diskurzu a textu přitom podle něj leží na pomezí několika vědeckých disciplín – lingvistiky, psychologie, sociologie a logiky. Sám se v pozdějších dílech zabývá spíše pragmatikou a sociologií.

Mezi významné osobnosti textové lingvistiky patří dále János Sándor Petőfi, který ve svých pracích navazuje na Chomského teorii generativní gramatiky. Obdobně jako pro van Dijka je pro Petőfiho významným tématem textová gramatika (srov. 1969, 1971, 1973 aj.) a dále vztah textologie a sémiotiky (2004, 2011).

Dalšími významnými vědci zabývajícími se textovou lingvistikou ve světovém měřítku jsou Wolfgang Dressler a Rober-Alain de Beaugrande. Svá stěžejní díla publikují především v sedmdesátých a osmdesátých letech (srov. např. *Einführung in die Textlinguistik*, 1972 či *Introduction to Text Linguistics*, 1981). Text je podle jejich pojetí záležitostí komunikace – aby tato komunikace proběhla úspěšně, musí splňovat sedm kritérií textovosti: jsou jimi koheze (povrchová manifestace koherence), koherence (významová spojitost na úrovni hloubkové struktury)⁵, intencionalita, akceptabilita, informativnost, situativnost, intertextualita.

Významnou prací věnující se textu a textové lingvistice je dále kniha *Cohesion in English*⁶ (1976), jejímiž autory jsou přední britští lingvisté Michael Alexander Kirkwood Halliday a Ruqaiya Hasanová. V úvodu knihy vyjadřují autoři myšlenku, že většina lidí má přirozenou schopnost říci o sekvenci několika vět, zda se jedná o ucelený text či náhodný shluk vět na sebe nenavazujících. Jedná se podle nich o určitý přirozený instinkt. Autoři se následně pokoušejí tento „instinkt“ opřít o vědecké zkoumání, tj. kladou si za cíl odpovědět na otázku, co dělá text textem. Jednou ze zásadních vlastností textu je podle nich koheze, kterou vymezují ze sémantického hlediska, tj. kohezní text se vyznačuje vzájemnou interpretační závislostí jednotlivých textových jednotek. Jinými slovy, jedna jednotka textu předpokládá druhou, tj. bez pochopení předchozího úseku textu nemůže být správně

⁵ V některých pracích jsou od sebe odlišovány termíny koheze a koherence – koheze je chápána především jako formální (resp. povrchová) soudržnost textu, koherence jako soudržnost významová (hloubková). V naší práci používáme souhrnný termín koherence ve významu gramatické i významové soudržnosti textu.

⁶ Halliday a Hasanová zde vymezují kohezi ze sémantického hlediska – jako vlastnost textu, ve kterém interpretace jedné jednotky textu závisí na jiné (tj. daná jednotka textu předpokládá druhou).

interpretována ani jednotka následující. Text si tedy můžeme představit jako síť vztahů mezi jednotlivými úseky. Autoři v knize zavádějí pojmy „kohezní pouto“ (z anglického *cohesive tie*) pro vztah mezi dvěma jednotkami textu a „kohezní řetězec“ (z anglického *cohesive chain*) pro vyšší kohezní celek tvořený několika kohezními pouty.

Halliday a Hasanová dále dělí kohezi na gramatickou a lexikální. V rámci gramatické koheze se věnují popisu reference (osobní, ukazovací a srovnávací), substituce (může být nominální, verbální, větná), elipsy (nominální, verbální, větná) a spojovacích prostředků (v naší terminologii konektorů). V rámci lexikální koheze vydělují opakování (tj. opakování významu použitím stejného slova, synonyma, slova nadřazeného, podřazeného či obecného) a kolokace. Vzhledem k tomu, že se v této práci zabýváme výzkumem diskurzních konektorů, nebudeme ostatním prostředkům textové koheze věnovat bližší pozornost.

Dalším významným světovým autorem zabývajícím se textovou lingvistikou je Michael Hoey (srov. např. díla *Signalling in discourse*, 1979; *Patterns of Lexis in Text*, 1991 či *Textual Interaction: An Introduction to Written Discourse Analysis*, 2001), podle kterého je text vytvářen a organizován přítomností určitých slov či gramatických rysů v každé větě, které vybízejí čtenáře podívat se do okolního textu, aby je mohl správně interpretovat. Koheze je podle Hoeyho v textu zajišťována explicitními jazykovými prostředky, ať už lexikální či gramatické povahy.

U zmíněných autorů se přitom setkáváme se dvěma zásadními termíny – text a diskurz. V některých pojetích jsou dané pojmy považovány za synonyma, v jiných jsou striktně odlišovány. Z tohoto důvodu věnujeme vymezení těchto termínů samostatnou kapitolu (srov. kapitola 3.2.1).

3.2 Kořeny textové lingvistiky v české literatuře

Textová lingvistika se jako samostatná vědecká disciplína začala v českém prostředí formovat v 70. letech 20. století, nicméně zkoumání nadvětných jevů bylo předmětem zájmu předních českých lingvistů mnohem dříve.

Jazykovým jevům přesahujícím rámec věty se podrobněji věnovali již někteří lingvisté Pražské školy a posléze její pokračovatelé. Ve třicátých letech publikuje Vladimír Skalička článek *Promluva jako lingvistický pojem* (1937), ve kterém se zabývá nadvětnými útvary navazujícími na linii hlásky, tvary, slova, věty. Nadvětný útvar nazývá promluvou, kterou definuje jako „totální sémiologickou reakci“ (Skalička, 1937: 163). Skalička se ve svém článku zabývá především vztahem promluvy k nižším, souřadným a vyšším jednotkám a upozorňuje na skutečnost, že v dosavadní lingvistické teorii se nadvětným útvarům věnuje málo pozornosti a lingvistická analýza většinou nejde za hranici věty. V šedesátých letech pak Skalička publikuje článek *Syntax promluvy (enunciace)* (1960). V tomto článku již nechápe promluvu jako nejvyšší jazykový útvar (jako v uvedeném článku z roku 1937), ale nad promluvu klade tzv. „sdužení promluv“ (Skalička, 1960: 241), pro které používá termín diskurz. Ten je podle něj „konstruován pomocí jakési syntaxe“ (Skalička, 1960: 242). Již u Skaličky se tedy objevuje myšlenka nadvětné syntaxe, na kterou později navazují další lingvisté (srov. např. Hrbáček, 1994; viz níže).

Obecně se výzkum nadvětných jevů v českém prostředí následně ubírá různými směry. Někteří lingvisté se zaměřují na analýzu literárních textů, jiní se zabývají popisem specifických pravidel výstavby textu (tj. tzv. nadvětnou syntaxí) či různých stylů textu. Později se do popředí textové lingvistiky dostává komunikačně pragmatický pohled (text je chápán jako nástroj jednání účastníků komunikace v průběhu komunikační situace).

V jedné linii zkoumání nadvětných jevů autoři věnují pozornost především zapojení lingvistických metod do analýzy literárních textů. Mezi hlavní představitele tohoto pohledu na text patří např. Jan Mukařovský (*Kapitoly z české poetiky*, 1948), Roman Jakobson (srov. např. *Poetická funkce*, 1995), Lubomír Doležel (*O stylu moderní prózy*, 1960 či *Narativní způsoby v české literatuře*, 1993), Miroslav Červenka (*Významová výstavba literárního díla*, 1992) či Alena Macurová a Petr Mareš (*Text a komunikace. Jazyk v literárním díle a ve filmu*,

1993), obdobně v zahraničí např. Wolf-Dieter Stempel (1971) či Siegfried J. Schmidt (1973). Textová lingvistika je v tomto směru chápána jako disciplína na pomezí mezi jazykovědou a literární vědou (lingvistickými metodami autor analyzuje a interpretuje literární text).

Dlouhou tradici má v českém prostředí také výzkum aktuálního členění větného. To je v některých koncepcích považováno za jev na úrovni věty, někteří autoři jej ale chápou jako součást nadvětných jevů. Ve svých pracích mu věnuje pozornost především Vilém Mathesius (srov. např. díla *O tak zvaném aktuálním členění větném*, 1939 či *Jazykozpytné poznámky k řečnické výstavbě souvislého výkladu*, 1943). Na Mathesia navazují ve svých dílech například Jan Firbas (pracující s pojmem funkční perspektiva větná) v Brně a František Daneš a Petr Sgall v Praze. Zatímco Firbas (srov. např. 1962, 1978, 1992) a Daneš (srov. např. 1957, 1968) zvolili pro svůj výzkum aktuálního členění tradiční přístup k jazyku, Sgall (srov. např. 1964, 1967) ve svých dílech uplatňuje formální popis jazyka. Aktuální členění větné je pak významnou součástí tzv. funkčního generativního popisu (jazykové teorie formulované Petrem Sgallem a jeho žáky – srov. např. Sgall, 1964, 1967; Sgall, Hajičová, Buráňová, 1980; Sgall, Hajičová, Panevová, 1986; Hajičová, Partee, Sgall, 1998 aj.).

V rámci textově-gramatického pojetí textové lingvistiky se do popředí zkoumání dostává popis pravidel či mechanismů, které se uplatňují při výstavbě textu. V této souvislosti se hovoří o strukturním popisu nadvětné či textové syntaxe (můžeme se setkat i s pojmy hypersyntax, makrosyntax či textová gramatika) snažícím se hledat v textu systémové jevy. Textová syntax je přitom jakýmsi protějškem větné syntaxe – zatímco větná syntax se zabývá tím, jak se ze slov skládají věty, textová syntax zkoumá, jak se z vět skládají nadvětné útvary (pro které jsou užívány různé termíny, jako je např. text či diskurz). Zkoumáním nadvětné syntaxe se zabývají např. Bohumil Palek (srov. např. *Cross-reference. A study from hyper-syntax*, 1968), Karel Hausenblas (srov. např. *Gramatická syntax textu*, 1979), Jana Hoffmannová (srov. např. Hausenblas a Hoffmannová, *Od syntaxe k stavbě textu*, 1982) či Josef Hrbáček (*Text, promluva a nadpromluvové útvary*, 1991; *Nárys textové syntaxe spisovné češtiny*, 1994), ze slovenských lingvistů např. Jozef Mistrík (*Hypersyntax a štylistika*, 1978).

Nadvětné jevy se zároveň v chápání některých lingvistů ustalují jako samostatná rovina jazyka (srov. stěžejní práci Františka Daneše *Zum Status der Textebene*, 1983 či *Věta*

a text, 1985). Do popředí zkoumání textové lingvistiky se rovněž dostává rozlišení textu ve smyslu *langue* a *parole*. Analýza textu ve smyslu *parole* je předmětem zájmu především stylistiky, která má v českém prostředí hlubší kořeny – srov. zejména práce Karla Hausenblase (např. *On the characterization and classification of discourses*, 1964; *Výstavba jazykových projevů a styl*, 1971), Aloise Jedličky, Věry Formánkové a Miloslavy Rejmánkové (*Základy české stylistiky*, 1970), kteří zároveň reflektují některé systémové rysy textu, či nověji práce Jany Hoffmannové (*Stylistika a ... Současná situace stylistiky*, 1997).

Při chápání textu na úrovni *langue* a *parole* se rozlišuje hloubková a povrchová rovina textu – tato teorie se opírá o existenci tzv. textových vzorců, které popisuje zejména van Dijk (1972). Někteří autoři se přitom zaměřují na zkoumání sémantiky textu, tj. na významovou stavbu textu a jeho jednotek (srov. např. Daneš a Viehweger, 1983).

V sedmdesátých letech dochází k tzv. komunikačně-pragmatickému obratu ve světové lingvistice, který významně ovlivňuje i směřování textové lingvistiky u nás. Lingvistický výzkum se tedy obrací ke komunikačním aktům, recepci jazykových projevů, mluvené komunikaci apod. Významnou studií týkající se dané tematiky v češtině je kniha *Komunikace a čeština* Jana Kořenského (1992), která přináší například typologii komunikačních situací, popis různých aspektů komunikačního chování, zdůrazňuje nutnost osvojování zásad, pravidel a norem v komunikaci apod. V poslední době se komunikaci a kultuře vyjadřování věnovali např. Světlá Čmejrková (srov. 1999, 2000 aj.) či Petr Kaderka (srov. např. 2013).

Pro účely naší práce jsou z oblasti textové lingvistiky stěžejní především studie zabývající se problematikou diskurzních konektorů, tj. jazykových prostředků spojujících textové jednotky a vyjadřujících mezi nimi určité významové vztahy. Systematicky se popisu konektorů v češtině věnuje především Jana Hoffmannová (srov. např. 1983, 1984), dále František Daneš (1985), Josef Hrbáček (1994), Přemysl Adamec (1995), Marek Nekula (1996) nebo Ondřej Pešek (2011) aj. Více k problematice diskurzních konektorů a jejich popisu v českém a světovém kontextu viz kapitola 3.6.

3.2.1 Diskuze o nadvětném útvaru – vymezení pojmu diskurz

Při zkoumání jevů přesahujících rámec věty je nejprve nutné stanovit, jaký jazykový útvar je předmětem našeho zájmu. V obecné rovině můžeme říci, že se jedná o útvar nad větou, v literatuře se tedy často setkáváme s pojmem „nadvětný útvar“ – pro něj bývá užíváno různých termínů, např. text, diskurz nebo komunikát. Rozbor jednotlivých termínů a zamyšlení nad vhodností jejich užívání přináší ve svém článku *Text, komunikáty a jejich komplexy. Zamyšlení pojmoslovné* Karel Hausenblas (1984). Stručně jednotlivé pojmy představíme.

Sám Hausenblas používá pro základní jednotku sdělování a komunikování termín komunikát, který přebírá od Romana Jakobsona (1960) z jeho schématu komunikačních složek. Zároveň ovšem říká, že „komunikace neprobíhá jen v izolovaných jednoduchých komunikátech, nýbrž velmi často i v složených celcích“ (Hausenblas, 1984: 6). Tyto celky nazývá komplexy komunikátů. Výhodou tohoto termínu je podle Hausenblase to, že neodkazuje pouze k oblasti písemné či mluvené (a je tedy v tomto ohledu neutrální) a zahrnuje prostředky verbální i neverbální (jako příklady komunikátů uvádí plakáty, písňe apod.).

Dalšími termíny, které Hausenblas ve svém článku rozebírá a které se v tradiční lingvistice používají pro základní jednotku komunikování (a jsou tedy terminologickými konkurenty pojmu komunikát), jsou promluva, jazykový projev, text či diskurz (případně ještě ne zcela rozšířený termín enunciace navrhaný Vladimírem Skaličkou (1960) jako mezinárodní ekvivalent pojmu promluva).

S termínem promluva (ve smyslu nejvyšší jazyková jednotka či jednotka nad větou) pracuje ve třicátých letech Vladimír Skalička (1937) (ačkoli v pozdějších letech nad promluvu staví sdružení promluv, které nazývá diskurzem). Podle Hausenblase je nevýhodou pojmu promluva to, že svou motivací ukazuje k oblasti mluveného jazyka.

Dále Hausenblas rozebírá pojem jazykový projev, který podle něj označuje jednoduché komunikáty i některé jejich komplexy – srov. např. jazykový projev monologický a dialogický (Hausenblas, 1984). Podle Hausenblase je však tento pojem spojený spíše se stylistikou.

Termín text má podle Hausenblase výhodu v tom, že je mezinárodní, nevýhodou je ovšem jeho mnohoznačnost. Text může být například chápán jako písemný verbální výtvar, základní jednotka verbálního i neverbálního sdělování (písemná i ústní) či ve smyslu parole jako jazykový proces, manifestace jazyka (Hausenblas, 1984).

Posledním zmiňovaným pojmem v Hausenblasově článku je diskurz (jako základní jednotka sdělování). Jak bylo uvedeno výše, tento pojem používá již Vladimír Skalička (1960) pro označení vyšší jednotky stojící nad promluvou (jedná se tedy v tomto smyslu o sdružení promluv). Termín diskurz používá ve svých pracích dále např. Petr Sgall (1973). Nevýhodou pojmu diskurz je ovšem podle Hausenblase to, že na rozdíl např. od textu nemá v českém prostředí tradici.

Článek Karla Hausenblase o jednotlivých termínech, jimiž se pojmenovávají nadvětné útvary, je cenný z několika hledisek. Přináší především srovnání několika zásadních lingvistických pojmů označujících základní jednotky sdělování a jejich komplexy, které se ovšem zároveň v literatuře objevují v různých významech. V závěru článku Hausenblas upozorňuje na skutečnost, že vývoji v lingvistice je třeba přizpůsobovat i pojmový a terminologický aparát (Hausenblas, 1984), tj. že k vývoji nedochází pouze v lingvistice, ale i v oblasti terminologie. V posledních letech dochází k takovému vývoji u jednoho základního pojmu textové lingvistiky – u pojmu diskurz.

Zájem textové lingvistiky se v současné době obrací (mimo jiné) ke zkoumání významových vztahů mezi jednotkami nadvětných útvarů (ať už je nazýváme jakkoli) a k analýze jazykových prostředků, které mají schopnost tyto vztahy vyjadřovat. Ve světové literatuře bývá tato oblast lingvistiky nazývána diskurzí analýzou (z anglického *discourse analysis*). Slovo diskurz se tedy stává často používaným termínem textové lingvistiky (patrně nejvíce pod vlivem anglo-americké literatury). Znovu tedy vyvstávají otázky, jak pojem diskurz vymezovat a jaký je jeho vztah k dalšímu základnímu pojmu této lingvistické oblasti, tj. k textu. Srov. např. práce Jarmily Tárnyikové, která se zabývá systematickým popisem základních termínů textové lingvistiky v evropském i americkém kontextu v knize *From Text to Texture* (2002).

Tárnyiková upozorňuje na skutečnost, že se oba pojmy v literatuře objevují v mnoha různých významech, a to v souvislosti s tím, v rámci jaké lingvistické teorie jsou používány. Z tohoto důvodu se v jednotlivých pracích vyskytují samy nebo vedle sebe, jako synonyma i antonyma, jedno může zahrnovat druhé apod. (srov. Tárnyiková, 2002: 20). Na začátku každé práce, která tyto termíny používá, je tedy nutné je nejprve významově vymezit. V některých případech je podle Tárnyikové text spojován s psaným jazykem, diskurz s mluveným, v jiných je text chápán jako abstraktní výtvar (funkčně analogický k termínu *langue*), zatímco diskurz jako jeho konkrétní manifestace (tj. na úrovni *parole*). Zároveň může

být text vykládán jako pojem statický, tj. výtvar, diskurz jako pojem dynamický, zdůrazňující proces vytváření textu. Text ovšem může být chápán rovněž jako mluvený i psaný a může zahrnovat jednoho a více mluvčích/pisatelů, tj. texty mohou být např. monologické i dialogické apod. (srov. Tárníková, 2002: 19, 20).

Podle Tárníkové je termín diskurz častěji používaný v americké tradici, text v evropské. Na tomto místě cituje amerického lingvistu Roberta E. Longacra, který ve své knize *The Grammar of Discourse* napsal, že to, co on označuje jako diskurz, je v Evropě nazýváno textem (srov. Longacre, 1983: 337). Sama Tárníková chápe ve své knize oba termíny jako synonyma, resp. používá text jako zastřešující termín pro text a diskurz zároveň (srov. Tárníková, 2002: 21).

Vzhledem k tomu, že slovo diskurz a diskurzní analýza se začínají častěji objevovat i v současné české lingvistice (srov. např. Poláková a kol., 2012a) – patrně v reakci na lingvistiku světovou, pokusíme se v následující části práce různé významy termínu diskurz popsat a vysvětlit.

Pojem diskurz bývá v odborné literatuře používán v různých významech. V některých případech bývá diskurz definován poměrně široce či vágně. Podle americké lingvistky Ellen Princeové je diskurz (či diskurzní analýza) dokonce jedním z nejméně přesně vymezených termínů na poli lingvistiky – srov. „Discourse analysis is without doubt one of the most widely used and loosely defined terms in the entire field of linguistics.“ (Princeová, 1988: 164)

Různými pojetími diskurzu se důkladně zabývá např. Deborah Schiffrinová ve své knize *Approaches to Discourse* (1994). Diskurz analyzuje především ze dvou pohledů, a to strukturního (formálního) a funkčního.

3.2.1.1 Strukturní přístup k diskurzu

Z hlediska strukturního přístupu je diskurz vnímán v zásadě jako synonymum k textu – tj. jako manifestace jazyka přesahující rámec věty. Podstatou lingvistického zkoumání je v tomto případě organizace jazyka nad úrovní věty – tj. předmětem zájmu se stává především to, v jakém sledu se jednotlivé úseky (či jednotky) textu vyskytují, jaké jsou mezi nimi vzájemné vztahy a jak jsou tyto vztahy vyjádřeny.

Strukturní přístup tedy předpokládá, že je diskurz rozložitelný na menší jednotky. Jeden z prvních lingvistů zabývajících se diskurzní analýzou Zellig Sabbetai Harris (srov. např. *Methods in Structural Linguistics*, 1951 či *Discourse Analysis*, 1952) viděl organizaci

diskurzu v linii *morfém – slovo – klauze – věta – diskurz*. V tomto ohledu můžeme podle některých autorů (srov. např. Salkie, 1995) organizaci diskurzu srovnat s gramatikou – v úvodu své knihy *Text and discourse analysis* (1995) Salkie říká, že nejlepší způsob, jak porozumět diskurzu, je srovnat ho s jinou lingvistickou oblastí, a to s gramatikou – zatímco gramatika podle něj zkoumá, jak jednotlivá slova tvoří ucelené věty, podstatou diskurzu je to, jak jednotlivé věty vytvářejí ucelený text – srov.: „Grammar is basically about how words combine to form sentences; text and discourse analysis is about how sentences combine to form texts.“ (Salkie, 1995: 9)

Diskurz tedy vykazuje vnitřní hierarchickou strukturu a v tomto ohledu je analogický k větě. Výzkum se v tomto případě zaměřuje na identifikaci vztahů mezi jednotlivými úseky (či jednotkami) textu, a to analogicky k identifikaci syntaktických vztahů v rámci věty (srov. např. Hovy, 1990). Diskurzní konektory jsou pak v tomto pojetí (srov. např. Halliday a Hasanová, 1976; Blakemoreová, 2002) definovány na základě toho, do jaké míry mají schopnost tyto vztahy signalizovat, a konečným cílem diskurzní analýzy je vztahy mezi úseky textu popsat a klasifikovat.

Podstatou tohoto přístupu k diskurzu je tedy skutečnost, že žádný úsek textu nezůstává osamocen či nezapojen, ale je vždy v nějakém vztahu k předchozí části (či částem) – pokud by text nějakou nezapojenou jednotku obsahoval, byla by narušena jeho soudržnost a srozumitelnost, tj. jedny ze základních vlastností každého textu. O jaký typ vztahu se jedná a jakým způsobem je vyjádřen, je předmětem zkoumání diskurzní analýzy.

3.2.1.2 Funkční přístup k diskurzu

Druhým zásadním přístupem k diskurzu, který ve své knize Schiffrinová uvádí, je přístup funkční. Předmětem tohoto přístupu je podle Schiffrinové jazyk v praxi či jazyk v kontextu (v zahraniční literatuře se používají termíny *language in use* nebo *language in context*). Podle Ralpa Fasolda (1990) je v tomto případě analýza diskurzu analýzou jakéhokoli aspektu používání jazyka – srov.: „The study of discourse is the study of *any* aspect of language use.“ (Fasold, 1990: 65)

Podobný přístup zastávají i další významní lingvisté (srov. např. Brown a Yule, 1983; Schiffrinová, 1994), kteří chápou diskurz jako systém, prostřednictvím kterého se realizují konkrétní funkce. Z tohoto hlediska nemůže být analýza diskurzu omezena na popis lingvistických forem nezávisle na jejich účelu a funkci (srov. Brown a Yule, 1983).

Strukturní a funkční pojetí se tedy zásadně liší v definici textu. Jak již bylo řečeno, ve strukturním přístupu je text chápán v podstatě jako synonymum k diskurzu. Funkční pohled ovšem oba pojmy odlišuje. Text je chápán jako výsledek komunikace, zatímco diskurz jako samotný proces (srov. např. Widdowson, 2007).

Rozdíl mezi strukturním a funkčním přístupem k diskurzu vidí Schiffrinová (1994) i v pojetí diskurzních konektorů. Zatímco strukturní pohled podle ní chápe diskurzní konektory jako ukazatele vztahů mezi strukturními jednotkami textu, ve funkčním pojetí jsou konektory vymezeny jako ukazatele vztahů mezi jednotkami chování (používá zde termín *units of behaviour*). Jinými slovy, zatímco ve strukturním přístupu je diskurz chápán jako abstraktní objekt analogický k větě, ve funkčním přístupu je chápán jako produkt lidského chování (Schiffrinová, 1994).

Toto dělení Schiffrinové do jisté míry problematizuje Diane Blakemoreová ve své knize *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers* (2002). Podle ní je dělení diskurzu na strukturní a funkční poněkud zjednodušující. Nesouhlasí například s tím, že Schiffrinová označuje jako zástupce strukturního přístupu Noama Chomského. V rámci strukturního přístupu podle Schiffrinové by totiž Chomsky musel zastávat názor, že podstatou lingvistiky je zkoumání věty. S tímto názorem ovšem Blakemoreová zásadně nesouhlasí, protože Chomsky chápe větu pouze jako abstraktní objekt (srov. např. Chomsky, 1986).

Podle Blakemoreové se dělení přístupů k diskurzu na strukturní a funkční dotýká jiné zásadní dvojice pojmů Chomského (1986), a to je jazyk jako externí (tzv. *E-language*) a interní objekt (tzv. *I-language*). Podle Blakemoreové byl jedním ze zásadních přínosů Chomského v lingvistice přechod od studia jazyka jako externího objektu ke studiu jazyka jako interního systému, tj. zjednodušeně právě přechod od zkoumání věty ke zkoumání lingvistické kompetence uživatelů jazyka. Podle Blakemoreové by tato myšlenka měla mít zásadní vliv také na vývoj výzkumu diskurzu – tj. vývoj by se měl posunout od studia samotného diskurzu ke studiu diskurzní kompetence – tedy schopnosti lidí integrovat své lingvistické i nelingvistické znalosti do úspěšné komunikace. Srov.: „The move in the case of discourse would be a move from the study of discourse itself (whether this be defined in the structural terms advocated by Harris (1951) or Hovy (1990)) to the study of discourse competence, or in other words, to the study of the knowledge that underlies our ability to integrate linguistic and non-linguistic knowledge for successful communication.“ (Blakemoreová, 2002: 155)

Z tohoto hlediska je tedy patrné, že chápání diskurzu se u jednotlivých autorů může značně lišit a hranice mezi strukturním a funkčním přístupem nejsou ostré či zcela jednoznačně vymezené. Např. anotace diskurzních vztahů v Pražském závislostním korpusu se zaměřuje na zachycení vztahů mezi jednotlivými úseky textu, tj. přiklání se k tomu, že je diskurz dělitelný na menší jednotky, které jsou provázané vzájemnými významovými vztahy (tato skutečnost je podle Blakemoreové (2002) či Schiffrinové (1994) podstatou strukturního přístupu), tyto vztahy ale zároveň hledá v reálných textech, tj. produktech lidského chování a myšlení.

Strukturní a funkční přístup nejsou ovšem jedinými pohledy na význam a vymezení diskurzu. Existuje např. i následující pojetí (Schiffrinová ho ve své knize neuvádí), které chápe diskurz jako fenomén přesahující využití jazyka v praxi.

3.2.1.3 Diskurz jako fenomén přesahující využití jazyka v praxi

V tomto pojetí se autoři (srov. např. Coupland a Jaworski, 2006) snaží vymezení diskurzu posunout ještě o krok dále, tj. za užití jazyka v praxi. Diskurz se podle nich týká sociální, politické i kulturní scény – reflektuje a zároveň formuje sociální uspořádání a interakci uvnitř společnosti – srov.: „Discourse is beyond language in use. Discourse is language use relative to social, political and cultural formations – it is language reflecting social order but also language shaping social order, and shaping individuals’ interaction with society.“ (Coupland a Jaworski, 2006: 3)

Předmětem zkoumání diskurzu proto není pouze jazyk. Kromě lingvistiky se pojem diskurz vztahuje také na filosofii, sociolingvistiku, geografii, antropologii, psychologii, politologii atd. Ačkoliv každá z těchto oblastí diskurzu klade důraz na něco jiného, podle Couplanda a Jaworského je diskurz zásadním klíčem pro porozumění společnosti a společenským interakcím stejně jako pro porozumění jazyku samotnému.

Je tedy zřejmé, že diskurz bývá v odborné literatuře vymezován různě. Existuje řada přístupů a definic, které se v mnohých aspektech překrývají, v jiných naopak vzájemně liší. Coupland a Jaworski ve své knize *The Discourse Reader* (2006) vyjadřují myšlenku, že ať už je diskurz cokoli a jakkoli konkrétně či abstraktně je definován, většina autorů se shodne alespoň v tom, že má něco společného s jazykem, významem a kontextem.

„Whatever discourse is, and however concretely or abstractly the term is used, there will be at least agreement that it has focally to do with language, meaning and context.“ (Coupland a Jaworski, 2006: 11)

V předkládané disertační práci budeme termín diskurz (diskurzí) užívat v úzkém smyslu, tj. jako synonymum k textu (textový). Chápeme jej tedy jako dělitelný útvar skládající se z menších jednotek (k bližšímu vymezení jednotky diskurzu – viz kapitola 3.4) s tím, že podstatou diskurzí analýzy je zpracování významových vztahů mezi jednotkami diskurzu a zkoumání jazykových prostředků, které mají tyto vztahy schopnost vyjádřit (tj. především diskurzí konektorů).

3.3 Co mohou spojit diskurzí konektory

Jak již bylo řečeno výše, podle některých přístupů je podstatou diskurzí analýzy hledání a interpretace diskurzí vztahů v rámci textu. Jednotliví autoři se ovšem neshodují v tom, mezi čím, resp. mezi jakými entitami, můžeme příslušné významové vztahy hledat. Touto otázkou se zabývá např. Villy Rouchota ve své knize *Discourse connectives: what do they link?* (1996). Rouchota zde popisuje dva zásadní přístupy k diskurzu a k diskurzí konektorům s ohledem na to, mezi jakými jednotkami vyjadřují příslušný diskurzí vztah – prvním je tzv. teorie koherence (v originále *coherence theory*), druhým teorie relevance (v originále *relevance theory*).

3.3.1 Teorie koherence

Teorie koherence říká (v návaznosti na Hallidaye a Hasanovou, 1976), že základními vlastnostmi textu jsou koheze a koherence, tj. jednotlivé úseky (či jednotky) textu jsou vzájemně propojeny určitými vztahy. Porozumění těmto vztahům pak přispívá k porozumění celému textu a úlohou diskurzí konektorů je pomoci čtenáři tyto vztahy lépe identifikovat (srov. např. Mann a Thompsonová, 1986; Fraser, 1990; Sanders, Spooren a Noordman, 1992; Hovy, 1990 či Knott a Dale, 1994).

3.3.2 Teorie relevance

Teorie relevance navazuje na některé stěžejní myšlenky výkladu komunikace Herberta Paula Grice, tj. že výpověď automaticky vzbuzuje u čtenáře/posluchače očekávání, které vede k porozumění významu zamýšleného pisatelem/mluvčím (srov. Grice, 1961).

Teorie relevance tedy předpokládá, že autor má obvykle na mysli jistou interpretaci svého výroku a očekává, že posluchač dojde ke stejnému výkladu. Aby posluchač k této zamýšlené interpretaci dospěl, musí ovšem daný výrok „zpracovat“ ve správném kontextu (srov. Sperber a Wilsonová, 1986; Wilsonová a Sperber, 1993).

Podle teorie relevance jsou tedy základem komunikace vodítka, která autor poskytuje adresátovi. Adresát pak z nich pomocí vzájemně sdílených znalostí rozkrývá autorův komunikační záměr (srov. Pokorný, 2010). Mezi tato vodítka patří právě také diskurzí konektory.

Pro teorii relevance je zároveň významnou opozicí dvojice pojmový (konceptní) vs. procedurální význam (v originále *conceptual vs. procedural meaning*). Zatímco slova s pojmovým významem přispívají k vlastnímu obsahu výpovědi, slova s procedurálním významem zasazují pojmy do kontextu a pomáhají čtenáři/posluchači pochopit celkový význam výpovědi a záměr pisatele/mluvčího. Z tohoto hlediska chápe Rouchota diskurzní konektory jako slova s procedurálním významem.

O tom, zda diskurzní konektory nesou pojmový či procedurální význam, se ovšem v odborné literatuře vedou diskuze (srov. také Blass, 1990; Blakemoreová, 1992, 2002; Wilsonová a Sperber, 1993 aj.). Zatímco někteří autoři (Rouchota, 1996; Fraser, 1999) konektory striktně vymezují jako výrazy nesoucí pouze význam procedurální, podle jiných (srov. např. Blakemoreová, 2002) mohou mít konektory i význam pojmový. To znamená, že podle Blakemoreové nejsou diskurzní konektory jednotnou, ale heterogenní skupinou – srov.: „Finally, as the last section has shown, not all expressions classified as discourse connectives/markers can be analysed in terms of procedural encoding. There are so-called discourse markers which encode constituents of conceptual representations. The picture that emerges from this approach is one in which there is no single category of discourse markers.“ (Blakemoreová, 2002: 185)

Podle našeho názoru je v souvislosti s touto diskuzí klíčové to, jak jednotliví autoři diskurzní konektory vymezují. Zatímco Rouchota konektory chápe jako „malá“ slova (v originále *little words*) typu *ale, a, zatímco* a Fraser jako slova patřící pouze mezi spojky, příslovce a předložkové fráze, pojetí Blakemoreové je značně širší. Jako diskurzní konektory chápe i výrazy typu *jinými slovy, to jest* či *výsledkem je*, které podle ní nesou význam pojmový, protože mohou být součástí sémanticky komplexních struktur jako *řící jinými slovy* (v originále *to put in other words*) a přispívat tak k celkovému pojmovému obsahu dané výpovědi (srov. Blakemoreová, 2002). (Více k diskurzním konektorům v kapitole 3.6.)

3.3.3 Diskurzní konektory z pohledu teorie koherence a teorie relevance

Z hlediska analýzy diskurzních vztahů se obě teorie výrazně liší především ve výkladu toho, jaké entity spojují diskurzní konektory – podle teorie koherence konektory spojují dva úseky textu (podle Rouchoty většinou realizované dvěma na sebe navazujícími větami či souvětími), zatímco podle teorie relevance vyjadřují vztah mezi výpovědi

a situačním kontextem, a to verbálním i neverbálním (v závěru své práce Rouchota používá místo slova kontext spojení „specifický inferenční proces“).

V tomto ohledu Rouchota přistupuje k teorii koherence spíše kriticky – podle něj tato teorie nedokáže vysvětlit, jak je možné, že se konektory často vyskytují na samém začátku diskurzu (např. v úvodní větě textu), tj. případy, ve kterých je narušen „klasický“ model *výpověď – konektor – výpověď*. Rouchota ve své práci uvádí několik takových příkladů a vykládá je podle teorie relevance s tím, že teorie koherence žádné relevantní vysvětlení nenabízí. Jedná se např. o následující výpověď (přeloženo do češtiny):

(1) (Kontext: Peter přichází domů z práce ve tři hodiny odpoledne.)

Mary: *Ale dnes jsi měl mít odpolední směnu.*

(Rouchota, 1996: 6)

Ale dnes jsi měl mít odpolední směnu je první vyřčenou výpovědí v dané situaci, a tedy nemůže navazovat na žádnou výpověď předchozí. Konektor *ale* zde proto nemůže spojit dva realizované (tj. v tomto případě vyřčené, verbální) úseky diskurzu. Tento příklad tak podle Rouchoty vyvrací pojetí konektorů v teorii koherence (podle které konektory spojují dvě realizované jednotky diskurzu). Teorie relevance naopak podle něj přináší relevantní vysvětlení – konektor *ale* zde vyjadřuje vztah mezi daným výrokiem a situačním kontextem. Tj. Peter přišel z práce dříve, než Mary očekávala. Tato situace ji zaskočila. Daným výrokiem tedy Mary reagovala na vzniklou situaci.

V tomto ohledu Rouchota vyjadřuje nesouhlas s výkladem konektorů podle teorie koherence s tím, že skutečnost, že se někdy konektory vyskytují v textu mezi dvěma jeho úseky, je pouze náhodná. Zároveň ve své knize říká, že nejčastější výskyt diskurzních konektorů je právě v první výpovědi celého diskurzu. Toto své tvrzení již ovšem nijak nedokládá.

V Pražském závislostním korpusu (PDT) jsou diskurzní vztahy zpracovány spíše podle teorie koherence – tj. vztahy (vyjádřené konektory) jsou zde značeny mezi dvěma úseky textu. Ovšem případy, ve kterých jeden z daných úseků „chybí“ (tj. i příklady, ve kterých se konektor vyskytuje v první větě textu), Pražský diskurzní korpus nepomíjí. Všechny tyto výskyty jsou v anotaci důsledně značeny.

V datech PDT se vyskytuje 277 podobných případů⁷ – srov. například nadpisy článků *A nejistota zůstává* nebo *Kdo je tedy lhář?* Počet diskurzních konektorů v PDT, které stojí mezi dvěma úseky textu, je 21 416 (viz kapitola 6.1). Je tedy zřejmé, že korpus psaných novinových textů nepodporuje Rouchotovu myšlenku, že většina diskurzních konektorů nestojí mezi dvěma úseky textu, tj. že diskurzní konektor většinou odkazuje k situaci.

Na druhou stranu rozdíl mezi tím, zda diskurzní konektor navazuje na předchozí text či situaci, může podle našeho názoru souviset také s tím, zda se jedná o text mluvený či psaný⁸. Všechny příklady, které Rouchota uvádí na podporu teorie relevance, se týkají mluveného diskurzu – nastane vždy nějaká situace (Peter si bere jídlo, přijde domů z běhání, čte si časopis apod.), na kterou Mary reaguje. Dané situace tedy mají pokaždé alespoň dva přímé účastníky – jeden svým chováním vyvolá reakci toho druhého.

V mluvených projevech se často setkáváme s tím, že mluvčí má nutkání začínat svou promluvu konektorem (zřejmě má potřebu na něco navázat, tj. zasazuje svoji promluvu do situačního kontextu) – srov. *Takže dobrý den / Tak začneme*.

⁷ Tj. výskytů diskurzních konektorů (primárních a sekundárních), které mají pouze jeden (vyjádřený) diskurzní argument – k vymezení pojmu diskurzní argument viz kapitola 3.4.1.

⁸ Resp. s tím, zda jsou autor i adresát sdělení přímými účastníky komunikační situace.

3.4 Vymezení jednotky diskurzu

Jak již bylo řečeno výše, v naší práci budeme termín diskurz užívat jako synonymum k textu, tj. chápeme ho jako útvar dělitelný na menší jednotky, které jsou vzájemně provázány různými typy vztahů. V této kapitole se pokusíme vymezit a definovat tyto „menší jednotky“. Jednotliví autoři (chápující diskurz rovněž jako dále dělitelný útvar) se totiž ve vymezení jednotky diskurzu značně liší.

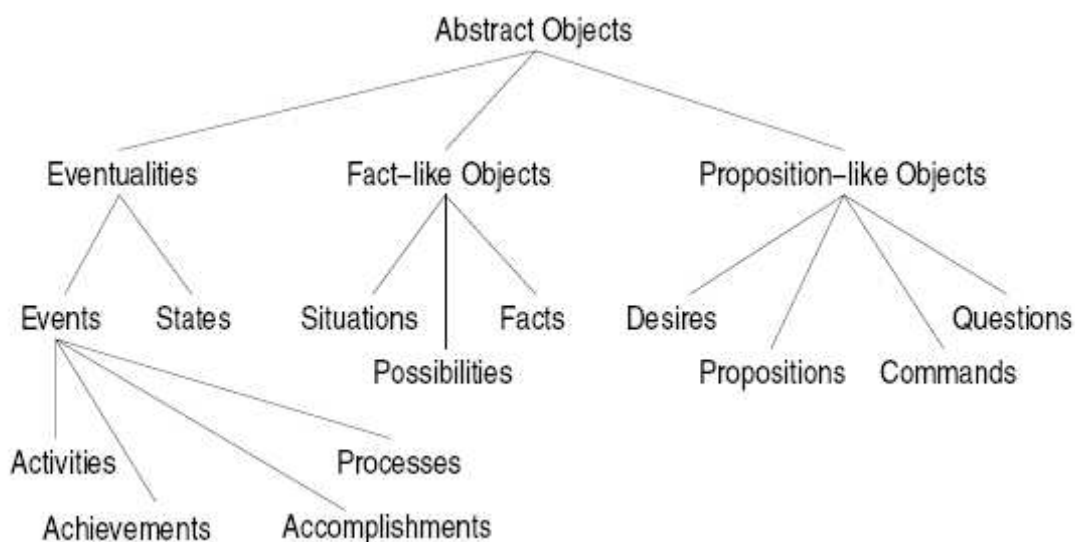
Někteří autoři ji definují z hlediska syntaktického, někteří z hlediska fonetického. Jako jednotka diskurzu tak může být označována např. věta (srov. Harris, 1951, 1952), výpověď (Redekerová, 1991; Fraser, 1996) či intonační jednotka (Chafe, 1987). Podrobně se jednotlivými přístupy k vymezení jednotky diskurzu zabývá Barbara Johnstonová ve své knize *Discourse analysis* (2002). Johnstonová shrnuje jednotlivé definice jednotky diskurzu a přináší jejich základní charakteristiku z různých pohledů. Podle ní bývají jako jednotka diskurzu v odborné literatuře označovány např. věta (obsahující podmět a přísudek a vyjadřující ucelenou myšlenku), výpověď (tj. konkrétní realizace věty), intonační jednotka (skupina slov vyjadřující ucelenou myšlenku často v podobě věty či fráze ohraničené z obou stran pauzami), tónová skupina (úsek řeči obsahující jeden hlavní přízvuk), tzv. centrum pozornosti (v originále *center of interest*, soubor několika intonačních jednotek odpovídající přibližně jedné predikaci) atd. Existují ovšem i další přístupy – např. v rámci zachycení diskurzních vztahů se v pensylvánském korpusu Penn Discourse Treebank (srov. Prasadová a kol., 2008) či v Pražském závislostním korpusu (srov. Bejček a kol., 2013) používá pro jednotku diskurzu označení diskurzní argument (viz níže).

Je tedy patrné, že vymezení jednotky diskurzu není zcela jednoznačné. Dianne Blakemoreová k tomu ve své knize *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers* říká, že se všechny zmíněné přístupy shodují alespoň v jedné věci – ať už za diskurzní jednotku označují cokoli, diskurz je podle nich dělitelný na menší úseky – srov.: „In all these approaches it is assumed that there is such a thing as discourse and that it can be divided up into units“. (Blakemoreová 2002: 152)
Podstatou diskurzní analýzy je pak hledat mezi danými úseky významové vztahy a prostředky, které je vyjadřují.

3.4.1 Diskurzní argumenty v zachycení Pražského závislostního korpusu

Vzhledem k tomu, že pro náš výzkum diskurzních konektorů používáme data Pražského závislostního korpusu, přebíráme i terminologii, která je pro popis těchto dat užívána (srov. např. Poláková a kol., 2012a), tj. pro jednotky diskurzu termín diskurzní argumenty.

Diskurzní argumenty jsou zde definovány jako abstraktní objekty (srov. Asher, 1993), které mohou být hierarchizovány (srov. obrázek 1).



Obrázek 1: Hierarchie abstraktních objektů podle Ashera (1993: 57)

Abstraktní objekty mohou být v textu realizovány klauzemi (či jejich posloupností) nebo jinými syntaktickými strukturami, jako je např. nominální fráze.

Diskurzní argumenty jsou vymežovány podle tzv. principu minimality (srov. Prasadová a kol., 2007). Podle tohoto principu patří do diskurzního argumentu pouze výpovědi (nebo jejich části), které nesou informace nezbytně nutné pro interpretaci příslušného diskurzního vztahu. Všechny ostatní jsou pouze doplňkové (tyto doplňkové části textu, které nepatří do vlastního argumentu, nejsou v Pražském závislostním korpusu značeny; v tomto ohledu se PDT liší od jiných, podobných korpusů zpracovávajících diskurzní vztahy – srov. např. pensylvánský korpus Penn Discourse Treebank: Prasadová a kol., 2008).

3.4.2 Povrchová realizace diskurzních argumentů – verbální vs. nominální argumenty

V Pražském závislostním korpusu jsou v současné době (srov. Poláková a kol., 2012b) zachyceny pouze takové diskurzní vztahy, které jsou vyjádřeny konektory v užším smyslu (tedy hlavně spojkami, příslovci a částicemi) spojujícími dva verbální argumenty (tj. argumenty, které obsahují sloveso v určitém tvaru), a to vnitrovětně (viz příklad 14 s konektorem *a tak*) i mezivětně (viz příklad 15 s konektorem *proto*).

- (14) *Největším nedostatkem slovenského celku byla nedůraznost, a tak Švédům stačil k vítězství jen poloviční výkon.*
- (15) *Přestup do Evropy je pro každého hráče z Jižní Ameriky velkým krokem do neznáma. Proto si musí najít klub, kde mají pro jeho aklimatizační problémy pochopení.*

V rámci výzkumu víceslovných konektivních struktur (především typu *důsledkem je, důvodem je, podmínkou je* apod.) jsme se ovšem neomezovali na syntaktickou realizaci příslušných argumentů, tj. zohledňovali jsme i diskurzní argumenty v podobě nominálních frází (podle Ashera (1993) je nominalizace jedním z typů abstraktních objektů) – srov. příklad 16:

- (16) *V roce 1993 klesl počet porodů na 748 500, což bylo o 1,3 procenta méně než v roce 1992.*
Důsledkem je stárnutí obyvatelstva.

V příkladu 16 vyjadřuje konektivní struktura *důsledkem je* diskurzní vztah příčiny a důsledku mezi verbálním argumentem *V roce 1993 klesl počet porodů...* a nominálním argumentem *stárnutí obyvatelstva*.

Uvedený příklad tedy dokládá, že některé víceslovné konektivní struktury se v reálných textech chovají jinak než „klasické“ konektory typu *proto, ale* apod., a tudíž na ně nemůžeme důsledně aplikovat stejné anotační zásady. Podle našeho názoru není významový rozdíl mezi následujícími příklady⁹:

⁹ Nebereme-li v úvahu případný rozdíl ve stylistické platnosti (složených) konektivních výrazů *a proto* a *důsledkem je*.

(17) *V roce 1993 klesl počet porodů na 748 500, což bylo o 1,3 procenta méně než v roce 1992.*

A proto obyvatelstvo stárne. /

Důsledkem je, že obyvatelstvo stárne. /

Důsledkem je stárnutí obyvatelstva.

Ve všech třech případech je explicitně vyjádřený diskurzni vztah příčiny a důsledku. Skutečnost, zda může být druhý argument vyjádřen větnou klauzí (*obyvatelstvo stárne*) nebo jmennou frází (*stárnutí obyvatelstva*), pak závisí na syntaktické povaze příslušného konektoru (např. *a proto* vyžaduje, aby byl druhý argument vyjádřen větnou klauzí – tj. není možné říci **A proto stárnutí obyvatelstva* – tato skutečnost byla pravděpodobně jedním z důvodů, proč byly diskurzni vztahy v Pražském závislostním korpusu nejprve zachycovány pouze mezi argumenty verbálními). Některé konektivní struktury (typu *důsledkem je*) ovšem nominalizaci diskurzniho argumentu dovolují.

Je tedy zřejmé, že některé konektivní struktury se chovají jinak než „klasické“ konektory typu *nebo, a, avšak* atd. a tuto skutečnost je při zpracování diskurzni vztahů třeba zohlednit. Z tohoto důvodu jsme se rozhodli neomezovat diskurzni argumenty jejich syntaktickou povahou (na rozdíl od pojetí diskurzni argumentů v Pražském závislostním korpusu při prvotní anotaci diskurzni vztahů), ale vymezovat je čistě sémanticky (v souladu s Asherovým vymezením abstraktních objektů).

3.5 Diskurzí vztahy

3.5.1 Diskurzí vztahy implicitní a explicitní

Jak jsme popsali výše, jednotlivé úseky (argumenty) textu jsou vzájemně provázány diskurzími významovými vztahy. Tyto vztahy mohou být v textu vyjádřeny explicitně (tj. pomocí konkrétních jazykových prostředků) nebo implicitně (tj. příslušný diskurzí vztah lze vyvodit pouze z obsahů obou diskurzí argumentů) – srov. příklady 18 a 19:

(18) *Mám rýmu. A proto zítra nepůjdu do školy.*

(19) *Mám rýmu. Zítra nepůjdu do školy.*

V příkladech 18 a 19 je diskurzí vztah příčiny a důsledku mezi argumenty *mám rýmu* a *zítra nepůjdu do školy*. V příkladu 18 je tento vztah explicitně vyjádřen konektorem *a proto*, v příkladu 19 pouze vyplývá z obsahů příslušných argumentů (tj. v textu není žádný jazykový prostředek, který by čtenáři pomohl vztah příčiny a důsledku interpretovat). Jak dokazují některé studie (srov. např. Irwinová a Pulverová, 1984), čtenář/posluchač rozumí lépe textům, ve kterých jsou diskurzí vztahy vyjádřeny explicitně. Diskurzí konektory tedy působí jako jistá vodítka pro čtenáře, kteří díky nim dokáží textu snáze a rychleji porozumět. Přítomnost diskurzí konektorů může zároveň zabránit víceznačnosti (srov. Dolková, 2010).

Implicitnost a explicitnost diskurzí vztahů přitom nelze chápat jako striktní binární opozici, ale spíše jako širokou škálu, která začíná prostým kladením vět vedle sebe (tzv. juxtapozicí, viz příklad 19) a končí komplexními souvětnými schémata obsahujícími výrazy s konektivní funkcí (srov. Jelínek, 1995).

U konkrétních textů pak můžeme mluvit o různém stupni explicitnosti. Podle M. Jelínka (1995) vyplývají vztahy mezi jednotlivými úseky textu často z kontextu, a proto může být jejich explicitní vyjadřování vnímáno jako příznaková nadbytečnost. K té mají podle něj sklon především odborné texty. Volba vhodné míry explicitnosti je pak jednou ze základních stylistických dovedností. Podle Jelínka je rozdíl mezi vyjadřováním diskurzí vztahů implicitně či explicitně natolik výrazný, že je možné hovořit dokonce o implicitním a explicitním stylu.

Některí autoři ovšem existenci čistě implicitních vztahů zpochybňují – M. Taboadová (2009) vyslovuje hypotézu, že neexistují diskurzí vztahy, které by nebyly vyjádřeny

žádnými explicitními jazykovými signály. Kromě diskurzních konektorů považuje za tyto signály, které se podílejí (v konkrétních příkladech) na výstavbě textu, např. také slovesný čas, způsob, užití určitého, resp. neurčitého slovesného tvaru, užití vzažné věty, některých interpunkčních znamének (např. pomlčky) apod. Taboadová tedy diskurzní signály chápe v širokém smyslu, dále je dělí na morfologické, syntaktické, sémantické a pragmatické. Podle jejího pojetí tak žádný diskurzní vztah není čistě implicitní (tj. odvoditelný pouze z obsahů argumentů), ale vždy je signalizován nějakým explicitním jazykovým prostředkem, i když ne vždy diskurzním konektorem, ale právě např. určitou mluvnickou kategorií apod. Hlavní myšlenkou Taboadové tedy je, že diskurzní signály je zapotřebí hledat na různých jazykových rovinách.

3.5.1.1 Pokusy o zpracování implicitních diskurzních vztahů ve velkých korpusech

Jak již bylo řečeno, čtenáři mnohdy lépe rozumí textu jako celku, pokud jsou v něm jednotlivé diskurzní vztahy vyjádřeny pomocí konkrétních jazykových prostředků (tj. explicitně). Tato skutečnost pak platí i pro anotátory diskurzních vztahů ve velkých korpusech, protože anotátor je v první řadě také čtenář. Ve většině projektů zaměřených na zpracování diskurzních vztahů na rozsáhlých jazykových datech se proto také zachycují hlavně vztahy explicitní – srov. např. CSTNews Corpus (Cardoso a kol., 2011), Pražský závislostní korpus (Bejček a kol., 2013), Potsdam Commentary Corpus (Stede a Neumann, 2014), korpus DisFrEn (Crible, 2014), LOCAS-F (Degandová a kol., 2014) aj. Při zpracování vztahů implicitních totiž hrozí vysoká míra neshody mezi jednotlivými anotátory.

Přesto existují projekty, které se o zachycení implicitních vztahů v textu pokoušejí. V první řadě se jedná o pensylvánský korpus Penn Discourse Treebank (Prasadová a kol., 2008). Ten zachycuje diskurzní vztahy různého druhu – vyjádřené explicitními konektory v užším smyslu (typu *proto*, *nebo*, *pak* atd.), konektory v širším smyslu (autoři projektu pro ně používají termín alternativní lexikální vyjádření diskurzních konektorů – zkráceně altlexy – typu *to je důvod*, *proč*) i vztahy vyjádřené implicitně (více k anotaci korpusu Penn Discourse Treebank v kapitole 3.7.4.3). Zpracování implicitních vztahů zde ovšem bylo omezeno pouze na vztahy mezi sousedními větami, značeny nebyly ani implicitní vztahy vnitrovětné či vztahy přesahující jeden odstavec. Při takovémto omezení bylo dosaženo poměrně vysoké mezianotátorské shody (85,1 %) na rozpoznání diskurzního vztahu a rozsahu

diskurzních argumentů a 72 % na dosažení explicitního konektoru, který by se pro vyjádření daného typu implicitního vztahu nejlépe hodil (srov. Miltsakakiová a kol., 2004).

Omezení implicitních vztahů pouze na sousední věty je ovšem značně zjednodušující – srov. malou sondu ke zpracování implicitních vztahů provedenou na datech Pražského závislostního korpusu (viz Poláková a kol., 2013). Shoda dvou anotátorů při zpracování implicitních vztahů (bez jakéhokoli omezení) ve třech textech obsahujících dohromady 100 větných celků byla v tomto případě výrazně nižší – 49 % na rozpoznání diskurzního vztahu. Je tedy zřejmé, že zpracování implicitních diskurzních vztahů na velkých korpusových datech může být problematické (přítomnost konektivního prostředku pomáhá příslušný typ vztahu snáze interpretovat a může zabránit víceznačností) – srov. příklady 20, 21 a 22 z *Příruční mluvnice češtiny* (1995: 749):

(20) *Bude pršet, zůstaneme doma.*

(21) *Bude pršet, proto / z tohoto důvodu zůstaneme doma.*

(22) *Jestli bude pršet, tak zůstaneme doma.*

Příklad 20 je ilustrací implicitního diskurzního vztahu. Je ovšem otázkou, o jaký významový typ vztahu se jedná (viz např. vztah příčiny a důsledku nebo vztah podmínky). Tuto ambiguitu diskurzního vztahu může vyřešit právě přítomnost explicitního diskurzního konektoru (v příkladu 21 *proto* či *z tohoto důvodu*, v příkladu 22 *jestli_tak*).

3.5.2 Významová klasifikace diskurzních vztahů

Mezi stěžejní témata vztahující se k analýze diskurzu patří bezesporu diskuze o klasifikaci diskurzních vztahů. Existuje přitom celá řada teoretických přístupů, které se ve vymezení diskurzních vztahů výrazně liší. Jednotlivé významné teorie a základní principy klasifikace v této kapitole blíže představíme.

Většina autorů bere jako základ k interpretaci významových vztahů výsledný text, ve kterém se jednotlivé vztahy snaží najít a popsat (tj. po předchozím procesu vytváření textu autorem). Vyhledání a interpretace diskurzního vztahu je tedy na straně posluchače/čtenáře (srov. kapitola 3.3.1). Se zajímavou myšlenkou přichází v tomto ohledu Jerry R. Hobbs (1979), který definuje diskurzní vztahy na základě vlastního vytváření textu, tj. ze strany

mluvčího/pisatele. Podle Hobbse vypadá proces vytváření koherentního textu následovně: autor vytvoří první úsek textu, posléze vybere vhodný diskurzí vztah (z pomyslného seznamu všech vztahů) a teprve poté vyprodukuje druhý úsek textu. Tímto způsobem jsou postupně vytvořeny všechny vzájemně provázané úseky textu tvořící jeden koherentní celek. Podobné přístupy ovšem počítají s tím, že existuje konečný a jasně definovatelný seznam všech významových diskurzí vztahů. Jednotliví autoři se nicméně ve výčtu tohoto seznamu ne vždy shodují. Odlišnosti se týkají celkového počtu jednotlivých vztahů i jejich definice – např. Hovy (1990) vymezuje sedmdesát významových vztahů, Mann a Thompsonová (v rámci Rhetorical Structure Theory (RST), 1988) šestnáct, Asher a Lascarides (v rámci své teorie Segmented Discourse Representation Theory (SDRT), 2003) sedm základních atd. Někteří autoři (srov. např. Prasadová a kol., 2008) vymezují diskurzí vztahy hierarchicky, tj. vyčleňují několik základních, obecnějších vztahů (např. vztahy časové), které lze dále členit na vztahy konkrétnější (např. vztah současnosti).

Některá pojetí přesný počet vztahů nespecifikují – například Asher ve svých dřívějších pracích (1993) uvádí, že povaha jednotlivých vztahů nedovoluje jejich celkový výčet, protože pro různé typy diskurzů jsou důležité jiné „sady“ vztahů. Asher nicméně představuje výčet typů vztahů, které jsou podle něj důležité (jsou jimi elaborace, vysvětlení, příklad, účel, pokračování, důsledek, zobecnění a kontrast). Podobně Hobbs počet typů diskurzí vztahů blíže nespecifikuje, pouze konstatuje, že je „poměrně malý“ (srov. Hobbs, 1979: 3).

Blakemoreová (2002) upozorňuje na skutečnost, že se jednotliví autoři liší i v definici některých vztahů – jako příklad uvádí vztah elaborace (která se podle ní vyskytuje u většiny autorů, ale která není vymezována jednotně). Zatímco např. Hobbs (1979) zahrnuje do elaborace i případy přeformulování (stejně jako opravy, repetice apod.), Mann a Thompsonová (1988) definují elaboraci a přeformulování jako dva samostatné vztahy.

Se zcela odlišným přístupem k vymezení diskurzí významových vztahů přicházejí Sanders, Spooren a Noordman (1992), kteří je definují z psycholinguvistického hlediska. Podle nich totiž není pravděpodobné, že by měl autor v podvědomí všechny typy nejrůznějších vztahů a že z kognitivního hlediska je výhodnější generovat celou sadu jednotlivých vztahů na základě kombinace čtyř jednoduchých prvků (přeloženo z anglického originálu):

- a) *základní operace (kauzální či doplňující)*
- b) *zdroj koherence (sémantický či pragmatický)*
- c) *polarita (negativní či pozitivní)*
- d) *pořadí segmentů (základní či nezákladní)*

(Sanders, Spooren a Noordman, 1992: 52)

Podle tohoto pojetí (jak uvádějí sami autoři) může být každý konektor přiřazen právě k jednomu typu vztahu, tj. diskurzní konektory jsou přiřazovány k jedné, zpravidla široce vymezené kategorii vztahů – např. konektory *přesto, nicméně, ale, však* jsou řazeny do obecné kategorie adverzativních (či kontrastivních) vztahů, konektory *proto, tak, tudíž* do kategorie vztahů kauzálních atd.

Podobné přístupy ovšem kritizuje Blakemoreová (2002). Podle ní totiž nejsou schopny reflektovat významové rozdíly (i když někdy velmi jemné) mezi jednotlivými konektory. Blakemoreová například detailně analyzuje rozdíly mezi anglickými konektory *but (ale), however (šak)* a *nevertheless (nicméně)* s tím, že jejich prosté přiřazení k adverzativním vztahům je zjednodušující a ne zcela dostatečné. Jak Blakemoreová ovšem sama připouští, detailní rozlišování mezi konektory může mít také svá úskalí. Pokud bychom upustili od jakékoli míry zobecnění jednotlivých vztahů, v podstatě každý konektor by mohl vyjadřovat svůj vlastní diskurzní vztah (viz Blakemoreová, 2002: 162).

Z českých autorů se vymezením typů diskurzních vztahů zabývá např. Josef Hrbáček (1994), který ve své knize *Nárys textové syntaxe* vymezuje čtyři základní typy vztahů (používá pro ně termín mezivýpovědní vztahy obsahově sémantické), a to vztahy enumerativní (výčtové), explikativní (vysvětlovací), temporální (časové) a kauzální (příčinné).

Jana Hoffmannová v *Mluvnici češtiny 3* nabízí podrobnější dělení (1987: 627–629). Podle ní je struktura celistvého textu tvořena sítí vztahů různého charakteru. Některé typy vztahů mezi jednotkami textu se přitom uplatňují již ve struktuře souvětí jako vztahy mezi jednotlivými větami – jedná se např. o vztah kopulativní, adverzativní (tj. kontrast, konfrontace, srovnání), významovou gradaci, kladení alternativních možností vedle sebe, připojení vysvětlujícího dodatku k předchozímu úseku textu, dále o vztahy kauzální (konkrétně o vztahy vyjadřující příčinu, důvod, důsledek, účel, podmínku apod.), časové (tj.

například současnost či nesoučasnost) či prostorové. Podle *Mluvnice češtiny 3* je v textu výrazná především síť vztahů časových, prostorové vztahy jsou zastoupeny méně často.

Všechny vyjmenované vztahy se uplatňují především mezi jednotlivými větami. *Mluvnice češtiny 3* uvádí další typy diskurzních vztahů, které se vyskytují spíše mezi většími jednotkami či úseky textu. Jedná se o exemplifikaci (či ilustraci), korekturu (rektifikaci), analogii, přirovnání, specifikaci, definici, zobecnění, explikaci, argumentaci či interpretaci faktu. Jednotka vztažená např. k celému textu (či jeho podstatné části) může vyjadřovat vyvození výsledků, závěr textu, shrnutí apod.

3.5.2.1 Významové vztahy v textu v Pražském závislostním korpusu

Vzhledem k tomu, že v této práci vycházíme z anotovaných dat Pražského závislostního korpusu, přebíráme klasifikaci jednotlivých diskurzních vztahů z anotačního manuálu pro zpracování diskurzu v PDT (viz Poláková a kol., 2012a). Klasifikace diskurzních vztahů (sestavená pro konektory v užším smyslu, tj. především synsémantika typu *proto, ale, nebo, také* atd.) je přitom kombinací některých výše zmíněných teoretických přístupů.

Stejně jako v klasifikaci vztahů Prasadové a kol. (2008) (podle které byla anotována rozsáhlá data v pensylvánském korpusu Penn Discourse Treebank) jsou i v Pražském závislostním korpusu chápány významy diskurzních vztahů hierarchicky. V klasifikaci diskurzních vztahů pro PDT jsou zároveň od Prasadové a kol. převzaty čtyři základní typy či obecné kategorie, tj. vztahy časové, kauzální, kontrastivní a vztahy navazovací (z anglického *expansion*). Dále jsou v PDT rozlišovány jednotlivé konkrétní typy vztahů, které ovšem mohou být vždy zařazeny do jedné ze zmíněných základních kategorií.

Jednotlivé typy vztahů se od klasifikace Prasadové a kol. (2008) (a dalších, především anglofonních autorů) ovšem liší – v Pražském závislostním korpusu je např. navíc vztah explikace či gradace. Pro rozlišení konkrétních typů diskurzních vztahů je tedy spíše využita klasifikace vztahů představená v *Mluvnici češtiny 3* (1987) (i když PDT nezahrnuje všechny vztahy, které jsou v mluvnici uvedené).

V této části práce představíme všechny typy významových vztahů v textu, které jsou v Pražském závislostním korpusu zachycovány. Konkrétní typy vztahů uvedeme vždy v rámci jejich základní, obecné kategorie (příklady ke každému vztahu jsou vybrány z manuálu pro zpracování diskurzních vztahů v Pražském závislostním korpusu (Poláková a kol., 2012a) a článku shrnujícího diskurzní anotaci v PDT (Zikánová a kol., 2015), v závorce je vždy uvedena zkratka, která je pro daný vztah v tomto korpusu používána).

I) Časové vztahy

a) Současnost (SYNCHR) – časový vztah současnosti vyjadřuje, že dva děje probíhají v čase zároveň – srov. příklad 23:

(23) *Město postihla krize a nezaměstnanost. Zároveň začala nová éra svobodných celních zón.*

b) Nesoučasnost (předčasnost a následnost) (PRECED) – časový vztah nesoučasnosti vyjadřuje, že jeden děj proběhl dříve než druhý – srov. příklad 24:

(24) *Po činu z místa utekl. Později se přihlásil na policii.*

II) Kauzální vztahy

Tato kategorie zahrnuje příčinné vztahy v širokém smyslu, tj. obsah prvního diskurzního argumentu implikuje obsah druhého. Obecně tedy kauzalita popisuje vzájemné souvislosti mezi jevy.

c) Příčina a důsledek (REASON) – diskurzní vztah příčiny a důsledku je v Pražském závislostním korpusu chápán poměrně široce (nerozlišuje například mezi příčinou a důvodem). Příčina a důsledek jsou zde chápány jako jeden typ vztahu (liší se pouze směrem, kterým je vztah veden). V tomto ohledu se klasifikace diskurzních vztahů Pražského závislostního korpusu liší např. od *Mluvnice češtiny 3* (ta textové kauzální vztahy diferencuje na příčinu, důvod, důsledek apod., tj. tyto vztahy jsou zde chápány jako samostatné typy (1987: 629), zatímco v Pražském závislostním korpusu jsou sloučeny pod jeden typ vztahu) či *Příruční mluvnice češtiny* (ta příčinu a důvod chápe rovněž jako dvě samostatné podskupiny patřící do „určení příčinných v širokém smyslu“ (1995: 469–483)). Srov. příklad 25:

(25) *Dostal výpověď. Pracoval totiž nezodpovědně.*

d) Nepravá příčina a důsledek (F_REASON) – nepravá příčina a důsledek vyjadřuje vztahy pragmatické, tj. založené na mimotextových jevech, zkušenosti účastníků komunikace apod. Nejedná se tedy o objektivní vlastnost textu (viz např. Hrbáček, 1994: 27) – srov. příklad 26:

(26) *Petr je doma, protože se svítí v oknech.*

Příklad 26 můžeme interpretovat následujícím způsobem – na základě toho, že se svítí v kuchyni, usuzuji, že je Petr doma. Světlo v kuchyni tedy není vlastní příčinou toho, že je Petr doma.

e) Podmínka a výsledek podmínky (COND) – srov. příklad 27:

(27) *Ceny centrálně vyráběného tepla by mohly být i nižší. Ale výrobci tepla musí začít hodně investovat.*

f) Nepravá podmínka a výsledek podmínky (F_COND) – opět se jedná o vztah pragmatický, tj. vztah podmínky je srozumitelný prostřednictvím inference (nevyjádřeného obsahu, který si posluchač/čtenář doplní) – srov. příklad 28:

(28) *Jestliže se někteří levicoví politici snaží porovnávat úroveň dotací u nás a v EU, je to působení podmínky velmi zavádějící.*

g) Účel (PURP)

(29) *Věra chodí pravidelně cvičit. Chce totiž zhubnout.*
(= Věra chodí pravidelně cvičit za tím účelem, aby zhubla.)

h) Explikace (EXPLICAT) – vztah explikace je jedním z diskurzních vztahů, který není u některých (především zahraničních) autorů samostatně vyčleněn (srov. např. Prasadová a kol., 2008) (důvodem může být to, že například angličtina nemá pro vztah explikace typický explicitní konektor – na rozdíl od českých výrazů *totiž*, *vždyť*, *přece*, které nemají v angličtině přímý ekvivalent, a jsou proto obtížně přeložitelné). V explikativním vztahu vysvětluje druhý argument obsah argumentu prvního – tj. jedná se o vysvětlení jevu či názoru uvedením detailu, který dané tvrzení podpoří. Důležitým aspektem explikativního vztahu je tedy významová podobnost obou argumentů (viz příklad 30), přičemž obsah prvního argumentu je obecnější, druhého konkrétnější (srov. Rysová, 2013a). Významným prvkem explikace bývá také subjektivnost (prvek obhajování – obhajujeme jisté tvrzení uvedením detailu, který naše tvrzení vysvětluje).

- (30) *Včerejší porada ministrů o státním rozpočtu na rok 1995 dopadla víc než dobře. Václav Klaus ani Ivan Kočárník totiž nenašli v Kramářově vile nikoho, kdo by se s nimi chtěl prát o ideu vyrovnaného rozpočtu.*

III) Kontrastivní vztahy

U kontrastivních vztahů jsou obsahy obou diskurzních argumentů různé – jsou tedy dávány do protikladu nebo jsou vzájemně porovnávány.

ch) Konfrontace (CONFR) – u vztahu konfrontace jsou proti sobě kladeny různé obsahy dvou diskurzních argumentů (které si ovšem neodporují), přičemž je zdůrazněna právě jejich odlišnost – srov. příklad 31:

- (31) *Klasická teplárna nebo kondenzační elektrárna nemá účinnost vyšší než 33 %. Kogenerační jednotka má naopak účinnost až 87 %.*

i) Opozice (OPP) – vztah opozice vyjadřuje, že obsahy daných diskurzních argumentů si odporují – srov. příklad 32:

- (32) *Chtěla bych pracovat jako doktorka. Nemám na to ale vzdělání.*

j) Nepravý kontrast (F_OPP) – vztahy nepravého kontrastu (stejně jako nepravé příčiny a nepravé podmínky) jsou založené na inferenci, tj. mimotextovém obsahu, který si čtenář/posluchač musí doplnit. O vztah kontrastu jde tedy v tomto případě pouze formálně – srov. příklad 33 (ve kterém se jedná o vztah mezi dvěma implikacemi).

- (33) *Podle vedoucího výroby Miloše Přiklopila má Seba rozpracovanou celou řadu zakázek. Zákazníci však vyvíjejí velký tlak na snižování cen tkanin.*

k) Restriktivní opozice (RESTR) – restriktivní opozice vyjadřuje vztah výjimky, tj. platnost prvního argumentu je omezena obsahem druhého argumentu – srov. příklad 34:

- (34) *Každá krajina má svou krásu. Jenom ji musíte umět vidět.*

l) Přípustka (CONC) – ve vztahu přípustky implikuje obsah prvního argumentu určité očekávání spojené s následujícím argumentem (srov. příklad 35 – předpokládá se, že při metrové vrstvě sněhu a za silného větru nebude fungovat letecký ani silniční provoz normálně), toto očekávání ovšem není naplněno (srov. *Mluvnice češtiny 3*, 1987: 490).

(35) *Když přistál letoun v Montrealu, ležela kolem letiště asi metrová vrstva sněhu a foukal silný vítr. Přesto fungoval letecký i silniční provoz normálně.*

m) Rektifikace (CORR) – u vztahu rektifikace druhý argument opravuje či modifikuje obsah prvního argumentu – srov. příklad 36:

(36) *Na druhé straně stát není soukromým majetkem ústavních orgánů. Ale je to veřejnoprávní instituce.*

n) Gradace (GRAD) – u vztahu gradace dochází ke stupňování obsahu prvního argumentu (obsah druhého argumentu převyšuje očekávání spojené s obsahem prvního argumentu) – srov. příklad 37:

(37) *Státní úředníci nemají dostatečný respekt. Nemají navíc snad ani představu o požadavcích Listiny základních práv a svobod.*

IV) Navazovací vztahy

Navazovací vztahy určují obecně především to, jakým způsobem obsah příslušného argumentu navazuje na obsah celého předchozího úseku textu, tj. zda jej shrnuje, specifikuje, uvádí příklad atd. Některé typy těchto vztahů (srov. např. generalizace) se vyskytují spíše mezi většími úseky textu (např. celými odstavci).

o) Konjunkce (CONJ) – vztah konjunkce vyjadřuje prosté navazování na předchozí kontext – srov. příklad 38:

(38) *Dánsko plně podporuje budoucí zapojení ČR do Unie. Obě strany jsou také spokojeny s růstem vzájemného obchodu.*

p) Exemplifikace (EXEMPL) – srov. příklad 39:

(39) *Jde o ceny průměrné, které se začnou v rámci státu výrazně lišit podle druhu paliva, používaného k výrobě tepla. Například v Brně lze očekávat letos průměrnou cenu až 180 Kč.*

q) Specifikace (SPEC) – u vztahu specifikace mají oba diskurzní argumenty podobný obsah, v prvním je ovšem obsah obecnější, v druhém konkrétnější. Druhý argument tedy přináší detail či podrobnost ke sdělení v prvním argumentu – srov. příklad 40:

(40) *V souladu se západními vzory je možná i omezená preference soukromého pojištění před sociálním pojištěním. Konkrétně, pokud si výdělečně činná osoba zaplatí dostatečně vysoké soukromé pojištění, bude se moci ze sociálního pojištění „vyvléknout“.*

r) Ekvivalence (EQUIV) – u vztahu ekvivalence dochází k opakování téhož obsahu u obou diskurzních argumentů (druhý argument říká jinými slovy totéž co první) – srov. příklad 41:

(41) *Dnes nebo zítra se v dolní komoře polského parlamentu – v Sejmu – očekává hlasování, které bude mít vážné politické důsledky, ať už dopadne jakkoliv, tj. bude-li zákon odmítnut či přijat.*

s) Generalizace (GENER) – ve vztahu generalizace shrnuje druhý diskurzní argument obsah prvního (který bývá často realizován větším počtem vět, např. i celým předchozím odstavcem) – srov. příklad 42:

(42) *Byl nesnášenlivý, netrpělivý a panovačný. Prostě to byl takový malý tyran.*

t) Konjunktivní alternativa (CONJALT) – konjunktivní alternativa označuje vztah dvou diskurzních argumentů vyjadřujících alternativy, které je možné chápat ve slučovacím i vylučovacím vztahu – srov. příklad 43:

(43) *Můžeme jít do kina. Nebo bychom mohli jít na večeři.*

u) Disjunktivní alternativa (DISJALT) – disjunktivní alternativa označuje vylučovací vztah mezi dvěma diskurzními argumenty – srov. příklad 44:

(44) *Chovej se tu slušně. Nebo sem nechod'!*

Jak již bylo zmíněno výše, klasifikace významových vztahů v textu byla v rámci jejich zpracování v Pražském závislostním korpusu sestavena pro konektory v užším smyslu (tj. synsémantika typu *ale, i, proto* atd.).

Vzhledem k tomu, že náš výzkum proběhl na datech Pražského závislostního korpusu, přebíráme tuto klasifikaci i pro vztahy vyjádřené konektivními strukturami typu *z tohoto důvodu, podmínkou je* atd. Jsme si ovšem vědomi toho, že zkoumané struktury jsou heterogennější skupinou než konektivní synsémantika (jedná se o jmenné, verbální či předložkové struktury, ustálená spojení atd.). Můžeme tedy předpokládat, že i škála významových vztahů, které tyto výrazy vyjadřují, může být bohatší a že bude nutné výše představenou klasifikaci vztahů (na základě zkušenosti s konkrétními texty) případně upravit a rozšířit.

Při zpracování diskurzních vztahů vyjádřených konektory v širším smyslu (tj. především víceslovnými frázemi typu *hlavní příčinou bylo* apod.) tedy anotátoři vybírali příslušné typy vztahů z výše uvedené klasifikace. Pokud ovšem klasifikace vhodný typ vztahu neobsahovala, označili anotátoři daný příklad poznámkou s návrhem „nového“ typu vztahu, o který by měla být zpětně rozšířena či obohacena (viz kapitola 6.6).

3.6 Diskurzí konektory

Diskurzí analýza může být v rámci různých teoretických přístupů (srov. např. Prasadová a kol., 2008, 2010; Stede a Neumann, 2014 aj.) založena na vyhledávání a identifikaci konkrétních jazykových prostředků, které mají v textu funkci diskurzí konektorů (tj. může vycházet od zpracování konkrétních výrazů a směřovat od nich k obecnějšímu popisu strukturace diskurzu). Teoretické vymezení těchto prostředků (diskurzí konektorů) je tedy pro dané přístupy zcela zásadní. Diskurzí konektory jsou obecně chápány jako explicitní ukazatele významových vztahů v textu, nicméně detailnější, obecně přijímaná definice diskurzí konektorů zatím neexistuje. V následující kapitole proto představíme hlavní přístupy k diskurzím konektorům v české i světové literatuře a posléze se pokusíme tyto jazykové prostředky vymezit pro naši práci.

3.6.1 Diskurzí konektory ve světové literatuře

Pojetí diskurzí konektorů se u jednotlivých autorů výrazně liší. Skutečnost, že konektory nejsou snadno definovatelné, je patrná již z toho, že pro tyto výrazy neexistuje ve světové literatuře jediný, všeobecně zavedený a používaný termín – jednotliví autoři nazývají dané výrazy většinou podle toho, z jakého hlediska je analyzují nebo o jak rozsáhlé množině hovoří. Setkáme se tedy například s termíny (diskurzí) konektory (z anglického *discourse connectives*, Blakemoreová, 1987, 1992, 2002; Rouchota, 1996; Celleová a Huartová, 2007), diskurzí částice (*discourse particles*, Schourup, 1985; Fischerová, 2006), diskurzí ukazatele (*discourse markers*, Schiffrinová, 1987; Urgellesová-Collová, 2010), diskurzí operátory (*discourse operators*, Redekerová, 1990, 1991), fatické konektory (*phatic connectives*, Bazanella, 1990), pragmatické konektory (*pragmatic connectives*, van Dijk, 1979; Stubbs, 1983), pragmatické ukazatele (*pragmatic markers*, Fraser, 1988, 1990, 1996), pragmatické operátory (*pragmatic operators*, Ariel, 1994), sémantické konjunkty (*semantic conjuncts*, Quirk a kol., 1985) aj.

Diskurzí konektory bývají u většiny autorů definovány z několika hledisek – např. z hlediska slovnědruhové příslušnosti, syntaktické povahy, významu, pravdivostních podmínek, pragmatiky, lexikální ustálenosti, fonologického chování, pozice ve větě apod.

a) Slovnědruhová příslušnost¹⁰

Diskurzvní konektory bývají nejčastěji chápány jako výrazy patřící mezi podřadicí a souřadící spojky, předložkové fráze a adverbia (srov. např. Fraser, 1999; Prasadová a kol., 2008, 2010). Někteří autoři (Schiffrinová, 1987; Hansen, 1998; Aijmerová, 2002) mezi konektory řadí i částice či jmenné fráze. Podle Fräsera (1999) jsou diskurzvními konektory i některé idiomy (typu *all things considered*, v češtině *celkem vzato*), které podle něj ovšem tvoří mezi konektory značnou menšinu.

Někteří autoři kombinují kritérium slovnědruhové příslušnosti daného výrazu s jeho pozicí ve větě – srov. např. podle Urgellesové-Collové (2010) mohou být diskurzvními konektory některá příslovce (jako příklady uvádí *still – nicméně, yet – přesto, rather – spíše, even – dokonce*), pokud se ve větě objeví na první pozici.

Slovnědruhová příslušnost a syntaktická povaha by podle našeho názoru ovšem neměly být hlavním kritériem pro vymezení diskurzvních konektorů, a to vzhledem k překrývání hranic mezi některými slovními druhy (srov. např. obtížně definovatelné hranice mezi částicemi a spojkami – srov. kapitola 3.7).

b) Pojmový vs. procedurální význam

Někteří autoři vymezují diskurzvní konektory na základě toho, zda nesou pojmový či procedurální význam (srov. kapitola 3.3.2), tj. zda přispívají k vlastnímu obsahu výpovědi nebo zda fungují pouze jako vodítka, která obsahují instrukce, jak nakládat s pojmovými složkami výpovědi, a která tak pomáhají čtenáři/posluchači rozkrýt komunikační záměr autora.

Rozlišení pojmového a procedurálního významu se odráží již ve filozofických pracích J. L. Austina (1962; srov. rozdíly mezi popisováním a naznačováním) a P. H. Grice (1975, 1989; rozdíly mezi říkáním a implikováním¹¹). Zatímco někteří autoři (srov. např. Rouchota, 1996; Fraser, 1999) vymezují diskurzvní konektory jako výrazy s významem procedurálním, jiní (srov. např. Urgellesová-Collová, 2010) argumentují, že některé konektory obsahují i výrazy s významem pojmovým.

¹⁰ V anglickém prostředí se používá jednotný termín syntaktické třídy, v originále *syntactic classes*, který slučuje slovní druhy a syntaktické struktury. Hodnocení slovnědruhové příslušnosti zde užívaných konektivních výrazů uvádíme ve shodě s autory daných studií.

¹¹ H. P. Grice tento rozdíl rozebírá např. na větě *Bill je filozof, a je proto statečný*. (přeloženo z originálu; 1989: 21). Podle něj sémantická funkce slova *proto* dovoluje mluvčímu **naznačit** či **implikovat**, ovšem ne **říci**, že mezi danými výpověďmi je důsledkový vztah.

Jednou z nejvýznamnějších autorek zabývajících se konektory z hlediska pojmového a procedurálního významu je Diane Blakemoreová, která se dané problematice věnuje v několika svých dílech, a její názory se v tomto ohledu významně vyvíjejí. Zatímco ve svých dřívějších pracích (srov. např. 1987, 1992) definuje konektory jako jazykové prostředky s procedurálním významem (a odlišuje je tak od ostatních lexikálních výrazů nesoucích význam pojmový), ve svých novějších dílech (2002, 2007) dochází k závěru, že některé konektory mohou nést i význam pojmový. Jako příklad uvádí následující větu (přeloženo z originálu, 2007: 328):

- (45) *Myslím, že to, co potřebujeme, je místnost se stolem, to znamená stolem, okolo kterého by studenti mohli sedět.*

Podle Blakemoreové je v daném příkladu konektor *to znamená*, který se podílí na celkovém pojmovém obsahu dané výpovědi, protože pomáhá čtenáři upřesnit pojmový význam slova *stůl*.

Podle našeho názoru je v tomto ohledu stěžejní především to, jaké výrazy jednotliví autoři považují za konektory. Zatímto Fraser (1999), Rouchota (1996) či Blakemoreová ve svých dřívějších dílech (např. 1987) mají pro konektory vymezení užší (tj. jako konektory označují především spojky, příslovce a předložkové fráze), v novějších pracích Blakemoreové (2002, 2007) a v knize Urgellesové-Collové (2010) jsou konektory chápány z širšího hlediska (tj. jako konektory označují i slovní spojení typu *to znamená, jinými slovy* či *výsledkem je*).

c) Pragmatické hledisko – pravdivostní podmínky

Někteří autoři hodnotí diskurzní konektory také podle toho, zda přispívají či nepřispívají k pravdivostním podmínkám daných výpovědí. Např. Blakemoreová (2002) vymezuje diskurzní konektory jako výrazy, které mají pragmatický význam, a tudíž nepřispívají k pravdivostním podmínkám výpovědí. Své tvrzení dokládá na následujícím příkladu (přeloženo z anglického originálu):

- (46) *Ben je z Nového Zélandu. A tak / i přesto / nicméně miluje ragby.*
(Blakemoreová, 2002: 33)

V daném příkladu mají podle Blakemoreové obě výpovědi stále stejné pravdivostní podmínky, nicméně pragmatický význam se mění v závislosti na užití diskurzního konektoru (*a tak, i přesto, nicméně* vyjadřují jiné typy diskurzních vztahů). Z toho je podle ní patrné, že se diskurzní konektory podílejí na utváření především významu pragmatického.

Jako výrazy nepřispívající k pravdivostním podmínkám výpovědí hodnotí diskurzní konektory i další autoři – srov. např. Fraser (1990, 1996) či Rieber (1997). Podle Urgellesové-Collové (2010) se na pravdivostních podmínkách nepodílí většina konektorů, některé ovšem ano (např. *therefore – proto*).

Další autoři nepovažují přispívání k pravdivostním podmínkám za zásadní kritérium pro vymezení diskurzních konektorů – tj. za konektory považují výrazy, které se na pravdivostních podmínkách podílejí (např. *because – protože, and – a, then – pak*) i nepodílejí (např. *but – ale, furthermore – navíc, hence – tudíž*) – srov. např. Schiffrinová (1987) či Knott a Dale (1994).

d) Další kritéria pro vymezení diskurzních konektorů – lexikální povaha, fonologické chování, pozice ve větě aj.

Dalšími kritérii, podle kterých někteří autoři vymezují diskurzní konektory, jsou například lexikální ustálenost daných výrazů, častá redukovaná výslovnost, prozodická „nezávislost“ či obvyklá první pozice ve větě (srov. např. Zwicky, 1985; Schiffrinová, 1987; Fischerová, 2006 aj.).

Většina autorů zmíněná kritéria kombinuje (aplikují je přitom na prototypickou skupinu konektorů – typu *a, ale, proto, však* apod. – s tím, že existují výjimky, pro které daná kritéria neplatí). V následující části práce uvedeme komplexní charakteristiku diskurzních konektorů od vybraných autorů.

Urgellesová-Collová (2010) například charakterizuje diskurzní konektory jako fonologicky krátká a redukovaná slova syntakticky nezapojená do větné struktury, která mohou být z věty vypuštěna bez jakýchkoli gramatických zásahů. Ze sémantického hlediska podle ní konektory obvykle nepřispívají k pravdivostním podmínkám daných výpovědí.

Podle Zwickyho (1985) se diskurzní konektory obvykle objevují na první pozici ve větě, jsou prozodicky nezávislé, tj. od okolního textu oddělené pauzami, jsou obvykle monomorfemické a syntakticky oddělené od zbytku věty. Všechny konektory mají podle něj v textu pragmatickou funkci.

Fischerová (2006) vymezuje konektory jako „malá“, syntakticky, sémanticky a často i prozodicky nezapojená, neohebná slova. Dodává ovšem, že se jedná o charakteristiku typických konektorů a že je zapotřebí studovat i ostatní lexikalizované výrazy.

Schiffrinová (1987) definuje diskurzní ukazatele značně široce – podle ní se může jednat i o paralingvistické jevy a neverbální gesta. Pro lexikální prostředky (v našem pojetí konektory) pak uvádí následující kritéria. Daný výraz či spojení musí být z věty syntakticky vydělitelný, obvykle se vyskytuje na první pozici ve větě.

Prasadová a kol. (2008, 2010) vymezuje dvě kategorie výrazů vyjadřujících diskurzní vztahy. Jako diskurzní konektory označuje pouze výrazy patřící mezi podřadící a souřadící spojky, předložkové fráze a adverbia. Pro všechny ostatní jazykové prostředky s diskurzní spojovací funkcí (typu *příkladem je* atd.) používá termín alternativní vyjádření diskurzních konektorů.

Většina autorů se tedy shoduje v základní charakteristice typických diskurzních konektorů (tj. výrazů jako *a, ale, proto, však* apod.). Liší se ovšem v tom, kde je možné stanovit hranici konektorů – někteří autoři vymezují konektory v užším smyslu (srov. např. Shloushová (1998), Hakulinen (1998), Maschler (2000) omezují konektory pouze na synsémantická slova), jiní v širším smyslu (tj. jako konektory definují např. i slovesné, jmenné fráze apod. – srov. např. Schiffrinová (1987), Hansen (1998); Aijmerová (2002)). Na základě toho se pak jejich komplexnější charakteristika konektorů liší (viz úvahy o procedurálním a pojmovém významu, pravdivostních podmínkách apod.).

Většina autorů se ve svých studiích zmiňuje o tom, že v rámci kategorie diskurzních konektorů existují jejich typičtí zástupci. V mnohých pracích se pak autoři zaměřují na charakteristiku právě těchto typických představitelů (srov. např. Fischerová (2006) analyzuje „the most typical members“ – nejtypičtější zástupce konektorů, Urgellesová-Collová (2010) popisuje „prototypické“ chování konektorů apod., Prasadová a kol. (2008, 2010) dokonce vymezuje dvě skupiny konektivních prostředků – konektory a jejich alternativní vyjádření). Tato skutečnost nás vede k zamyšlení, že v rámci množiny diskurzních konektorů existují dvě různé podskupiny – typické a „méně typické“ konektory. Jinými slovy, množina konektorů má své jádro (centrum) a okraj (periferii). Většina autorů se přitom shoduje na charakteristice centra (tj. typických konektorů, jako je *ale, nebo, protože, a* apod.) a neshoduje na charakteristice periferie – a to i v tom ohledu, zda „méně typické“ konektory vůbec jako konektory označovat.

Podle našeho názoru je proto v dalším výzkumu diskurzních konektorů užitečné zaměřit se právě na vymezení hranic těchto výrazů, a to jak v rámci samotné množiny konektorů (tj. na hranici mezi typickými a „méně typickými“ konektory, tedy mezi centrem a periferií), tak mimo ni – tj. na hranici mezi konektory a výrazy, které funkci konektorů již v textu nemají (viz kapitola 3.7).

3.6.2 Diskurzní konektory v české literatuře

V českém prostředí se diskurzními konektory (a výrazy podílejícími se na strukturaci textu obecně) zabývali především Karel Hausenblas (1962), Jiří Kraus (1981), Jana Hoffmannová (1983, 1984), František Daneš (1985), Josef Hrbáček (1994), Přemysl Adamec (1995), Marek Nekula (1996), Jarmila Tárnyiková (1998a, 1998b) či Ondřej Pešek (2011), ze slovenských lingvistů například Jozef Mistrík (1968, 1975, 1975/6, 1978) či Eugénia Bajzíkova (1975, 1979).

Jeden z nejsystematičtějších a nejucelenějších popisů diskurzních konektorů v češtině přináší ve svých studiích Jana Hoffmannová (srov. např. *Sémantické a pragmatické aspekty koherence textu*, 1983; *Typen der Konektoren und deren Anteil an der Organisierung des Textes*, 1984). Ve svých pracích se podrobně zabývá povahou výrazů či spojení, které mají v textu konektivní funkci. Konektory přitom chápe v širším smyslu, tj. neomezuje se např. na slovnědruhovou příslušnost daných výrazů apod., ale vymezuje je především z funkčního, sémanticko-pragmatického hlediska. Upozorňuje totiž na skutečnost, že konektory nesou jak prvek sémantický (jedná se o vztahovou sémantiku v rámci textu), tak pragmatický (tj. zejména postoj autora k textu, autora k adresátovi).

Ve svých studiích Hoffmannová dospívá k podrobné typologii diskurzních konektorů. Dělí je na: 1. vlastní konektory (vyjadřují se jimi základní vztahy vyskytující se mezi větami v souvětí i mezi většími jednotkami textu – např. *a, neboť, také*); 2. implikativní konektory (mají funkci uspořádání a hierarchizace obsahu, mohou se vztahovat i k obsahovému celku, který není v textu vyjádřen, ale je odvoditelný z kontextu – např. *teprve, jen, ještě*); 3. kompoziční konektory (tj. makrostrukturní – signalizují např. vzájemné vztahy mezi jednotkami v různých částech textu, začátek a konec apod. – např. *na závěr, zároveň, jednou – jindy*); 4. textové orientátory (pomáhají čtenáři orientovat se především v časových a prostorových vztazích v textu – např. *ted', zde, dosud*); 5. kontaktní prostředky (slouží k navazování kontaktu mezi účastníky komunikačního aktu – např. *poslyšte, řekl bych*);

6. metatextové komentáře (vyjadřují postoje mluvčího – např. *upřímně řečeno, lze říci*). Od první k šesté skupině přitom u konektorů roste jejich pragmatický aspekt.

Je tedy patrné, že Hoffmannová označuje jako konektory poměrně široké spektrum jazykových prostředků. Podle ní je jejich užívání „odrazem autorova kompozičního a formulačního úsilí, a rovněž jeho snahy usnadnit recipientovi dekodování makrostruktury a smyslu textu“ (Hoffmannová, 1983: 108) a jejich funkcí může být například „explicitní vymezení tématu textu, motivace komunikačního cíle, východiska myšlenkového procesu na počátku textu; dále odkazy k předchozímu kontextu (levostranné) i poukazy na přechod k následujícímu sdělení (pravostranné); poukazy na posloupnost, pořadí výkladu“ (Hoffmannová, 1983: 112), dále pak zvýraznění, vyjasnění, zpřesnění myšlenky, uvození závěru, shrnutí apod.

Podle Hoffmannové tak mohou mít funkci diskurzních konektorů různé jazykové prostředky lišící se syntaktickou, lexikální i sémantickou povahou – tj. například také slovesa (a to především slovesa říkání, myšlení a vnímání – např. *zdůvodnit, shrnout, vyvodit, popsat* apod.), otevřená skupina sekundárních předložek (často ve spojení s ukazovacím zájmenem – např. *vedle toho, v tomto okruhu, v této souvislosti, v tom případě* apod.) či ustálené fráze a idiomy (*lépe řečeno, přesněji řečeno* apod.) (Hoffmannová, 1983: 113, 114, 121).

Myšlenka, že diskurzní konektory nemusejí patřit pouze do kategorie spojek či částic, ale že funkci konektorů mohou v textu nabývat také slovesa, různé fráze nebo idiomy tedy není v české lingvistice nová. Zatím ovšem tyto jazykové prostředky nebyly v češtině zkoumány na velkých korpusových datech ani neexistoval korpus, který by obsahoval jejich anotaci (takto byly prvně zpracovány a zkoumány pro angličtinu – viz kapitola 3.7.4.3). Korpusové zpracování těchto prostředků v češtině je proto jedním z cílů naší práce.

Jiné, užší pojetí diskurzních konektorů přináší ve svých studiích Josef Hrbáček (*Nárys textové syntaxe*, 1994). Konektory vymezuje především na základě dvou aspektů: 1. konektory vyjadřují vztahy mezi větami (pro výrazy signalizující vztahy vnitrovětné používá termín junktory) a 2. konektory neplní ve větě funkci větných členů. Platnost konektorů podle něj mohou mít v textu spojky, příslovce, částice, předložkové, zájmenné i jiné výrazy (typu *ve skutečnosti, kromě toho*), ovšem kritériem zde není příslušnost k určitým slovním druhům, ale to, že se jedná o prostředky bez větněčlenské platnosti (srov. Hrbáček, 1994: 57).

Hrbáček však připouští, že explicitní spojovací funkci mohou mít i výrazy, které funkci větných členů plní (srov. např. *důsledkem je*). Podle něj mají tyto jazykové prostředky

v textu dvě funkce: „jsou větnými členy svých vět a současně vyjadřují svým lexikálním významem vztah mezivýpovědní (tím, že ho pojmenovávají)“ (Hrbáček, 1994: 54). Tyto prostředky ovšem Hrbáček mezi konektory neřadí právě kvůli jejich větněčlenské platnosti. Toto kritérium je tedy podle něj při vymezení diskurzních konektorů stěžejní.

Diskurzními konektory se ve svých pracích zabývá také František Daneš (1985). Pro nás je zajímavá zejména kapitola nazvaná „Některá funkční slova ve výstavbě věty a textu“. Daneš zde obecně definuje konektory na základě jejich funkce v sémantické výstavbě souvětí a textu (Daneš, 1985: 171), nikoli vzhledem k jejich slovnědruhové příslušnosti. Ve svých úvahách o konektorech přitom navazuje například na slovenského lingvistu Jozefa Mistríka (1968), a to především v myšlence, že konektory „nejen připojují, ale zároveň i modifikují význam podobně jako příslovce“ (Daneš, 1985: 172), tj. obsahují „jisté momenty povahy pragmatické a sémantické“ (Daneš, 1985: 172). Konektory pak mají povahu relátorů (sémantických predikátů), jejichž argumenty jsou větné propozice (tj. jednotlivé věty, souvětí či větné bloky). Daneš dále zdůrazňuje funkční komplexnost konektorů – tj. kombinují se v nich složky logicko-sémantické, pragmaticko-komunikativní, textové apod.

Daneš dále upozorňuje na skutečnost, že termín konektory bývá v lingvistice užíván značně odlišně a často vágně a připojuje drobnou, pro nás ovšem důležitou poznámku, že „široce pojaté pole konektorů (jakožto jazykových výrazů) má zřetelně oblasti centrální a periferní, obsahuje výrazy slovnědruhově i funkčně dosti různé, v jejichž funkčním potenciálu nalézáme vedle funkce spojovací v nestejně míře a v různé distribuci a dominanci mnohé momenty další“ (Daneš, 1985: 171). V této poznámce tedy jasně vyjadřuje myšlenku, že kategorie diskurzních konektorů zahrnuje konektory „centrální“ a „periferní“. Daneš tuto úvahu ovšem dále nerozvíjí a k oběma skupinám neuvádí žádné příklady. Není tedy jasné, podle jakých kritérií tyto skupiny vymezuje a dělení konektorů na „centrální“ a „periferní“ tak u něj zůstává pouze v této poznámce.

Mluvnice češtiny 3 se diskurzním konektorům věnuje v části nazvané „Prostředky koherence textových jednotek“ (autorkou tematického celku o konektorech a dalších prostředcích textové koherence je Jana Hoffmannová). Podle *Mluvnice češtiny 3* mohou být konektory chápány v nejširším pojetí jako „prakticky všechny prostředky koherence textových jednotek“ (1987: 686). Tj. konektory mohou být prostředky jazykové (gramatické i lexikálně-sémantické) i mimojazykové a „každá základní textová jednotka koherentního textu

obsahuje minimálně jeden prvek s navazovací funkcí, který ji spojuje s textovým okolím – tzv. konektor“ (1987: 685).

V mluvnici najdeme opět zmínku o tom, že konektory (jsou zde nazývány konektory připojovací) mají své centrum, které „tvoří skupina výrazů formálně totožných se spojovacími prostředky užívanými uvnitř souvětí a vět“ (1987: 705). Jedná se přitom zejména o téměř všechny souřadící spojky. Dále se mluvnice obecnému vymezení centrálních a periferních konektorů nevěnuje. Přináší ovšem podrobný rozbor několika odlišných skupin výrazů, které v textu nabývají funkce konektorů. Upozorňuje přitom na problematiku určování slovních druhů u některých konektorů – srov.: „Ve funkci připojovacích konektorů se přitom velmi často uplatňují i jiné slovní druhy (příslovce jako *rovněž*, *naopak*, *tady*, *tak*, částice *ať*, *vždyť*, *dokonce* aj.). Mezi funkčním využitím a tradiční, jednoznačnou slovnědruhovou klasifikací některých výrazů vzniká tak značné napětí“ (*Mluvnice češtiny 3*, 1987: 705).

Jako diskurzivní konektory jsou zde označovány (vedle početných skupin spojek, částic a příslovců) také víceslovné spojovací výrazy, „jejichž součástí je obecné substantivum kategoriálního významu a ukazovací zájmeno (*v tom případě*, *z tohoto důvodu*, *hlediska*, *v tomto ohledu*, *smyslu*, *bodě*; *v této souvislosti*; *tato skutečnost* aj.)“ (*Mluvnice češtiny 3*, 1987: 706), dále časové a prostorové relátory (např. *nejdříve*, *za prvé*) či konektory „kompoziční“ (např. *pokud jde o...*, *pokud se týká...*). Podrobně jsou v mluvnici rozebírány také tzv. textové orientátory (které zde nejsou řazeny mezi konektory, ale jsou vedle nich považovány za další prostředky pomáhající zajistit spojitost textu). Kategorie textových orientátorů je přitom podle mluvnice zcela otevřená, často se jedná o spojení, v nichž se objevují slovesa pravení, myšlení a vnímání, převládá tvar 1. osoby plurálu apod. – srov. např. *musíme se zamyslet*, *chceme zdůraznit*, *musíme podtrhnout* aj.

Z analýzy vybraných předních studií o diskurzivních konektorech v češtině je zřejmé, že ani v české lingvistice nejsou konektory chápány jednotně. Někteří autoři preferují užší pojetí (založené především na syntaktickém chování konektorů, tj. výrazů bez větněčlenské platnosti – srov. zejm. Hrbáček, 1994), jiní pojetí širší (vymezující konektory především ze sémanticko-pragmatického hlediska – srov. zejm. Hoffmannová, 1983, 1984).

V naší práci se proto pokusíme na tyto významné studie navázat a přispět k diskusi o diskurzivních konektorech podrobnou analýzou rozsáhlých jazykových dat. Teoretické závěry budou tedy vycházet z velkého množství reálných textů. Na základě jejich podrobné analýzy se pokusíme rozvést především myšlenku o centrálních a periferních konektivních prostředcích podílejících se na celkové koherenci textu.

3.7 Definice a dělení diskurzních konektorů

Jak jsme popsali v předchozí kapitole, pojetí diskurzních konektorů se u jednotlivých autorů liší. Autoři se většinou shodují na základní charakteristice typických konektorů (typu *a*, *ale*, *nebo* atd.), odlišují se ovšem ve vymezení jejich hranic, tj. jaké výrazy ještě můžeme mezi konektory řadit a jaké nikoli. Někteří autoři tak například vyčleňují dvě skupiny konektivních prostředků – 1. „klasické“ konektory (např. *proto*, *však*, *naopak*) a 2. tzv. alternativní lexikální vyjádření konektorů (zkráceně altlexy – srov. např. *příkladem je*, *z tohoto důvodu* atd. – viz Prasadová a kol., 2010), jiní (srov. např. Hoffmannová, 1983; Schiffrinová, 1987) definují konektory široce, tj. jako konektory označují i zmíněná „alternativní vyjádření“, někteří (srov. např. Hrbáček, 1994) vymezují konektory v užším smyslu (tj. „alternativní vyjádření konektorů“ mezi konektory vůbec neřadí) apod.

V této práci se proto zaměříme především na vymezování hranic diskurzních konektorů. Soustředíme se přitom na dvě otázky: 1. kde je možné stanovit hranici mezi diskurzními konektory a výrazy, které za konektory již označovat nemůžeme; 2. kde je možné stanovit hranici v rámci vlastní množiny diskurzních konektorů (jak jsme popsali v předchozí kapitole, nabízí se zde dělení na typické a „méně typické“ konektory, tj. na centrum a periferii diskurzních konektorů).

Jak již bylo naznačeno výše, schopnost vyjadřovat významové vztahy mezi dvěma úseky textu mají jazykové prostředky různých slovních druhů, různé lexikální i syntaktické povahy – srov. již dříve uvedené příklady 47–58:

- (47) *Petr je nemocný a nemůže jet. Proto se výlet ruší.*
- (48) *Petr je nemocný a nemůže jet. Výlet se tedy ruší.*
- (49) *Petr je nemocný a nemůže jet. Výlet se tak ruší.*
- (50) *Petr je nemocný a nemůže jet. Výlet se tím ruší.*
- (51) *Petr je nemocný a nemůže jet. Výlet se tímto ruší.*
- (52) *Petr je nemocný a nemůže jet. Na základě toho se výlet ruší.*
- (53) *Petr je nemocný a nemůže jet. Z tohoto důvodu se výlet ruší.*
- (54) *Petr je nemocný a nemůže jet. Kvůli tomu se výlet ruší.*
- (55) *Petr je nemocný a nemůže jet. Kvůli této skutečnosti se výlet ruší.*
- (56) *Petr je nemocný a nemůže jet. Kvůli jeho nemoci se výlet ruší.*
- (57) *Petr je nemocný a nemůže jet. Kvůli němu se výlet ruší.*
- (58) *Petr je nemocný a nemůže jet. Důsledkem je zrušení výletu.*

V uvedených příkladech je mezi danými diskurzními argumenty (*Petr je nemocný... a výlet se ruší*) vztah příčiny a důsledku. Ve všech případech je tento vztah také explicitně vyjádřen – tj. pomocí konkrétních jazykových prostředků (*proto, tedy, tak, tím, tímto, na základě toho, z tohoto důvodu, kvůli tomu, kvůli této skutečnosti, kvůli jeho nemoci, kvůli němu, důsledkem je*). Otázkou ovšem zůstává, zda všechny dané prostředky můžeme označit jako diskurzní konektory.

V následující podkapitole se pokusíme přispět k diskuzi o diskurzních konektorech, především k jejich obecnému vymezení a k otázce terminologie. Naše poznatky a obecné závěry vycházejí z analýzy rozsáhlých jazykových dat Pražského závislostního korpusu.

3.7.1 Syntaktická povaha diskurzních argumentů

Někteří autoři (srov. např. Poláková a kol., 2012b; Anders a Søggaard, 2013) vymezují diskurzní konektory mimo jiné na základě syntaktické povahy argumentů, které daný konektor spojuje. V těchto pojetích je diskurzní konektor chápán jako jazykový výraz spojující dva slovesné argumenty (tj. úseky textu obsahující sloveso v určitém tvaru) a současně jako výraz patřící mezi určité slovní druhy (nejčastěji spojky, některé druhy částic, některá příslovce). V první části definice (o větné povaze argumentů, mezi nimiž konektor stojí) tak autoři odlišují konektivní funkci daných výrazů v rámci textu (*Dělám ráda ruční práce. A ráda vařím. / Dělám ráda ruční práce a ráda vařím.*) a v rámci nominální fráze (*borůvky a maliny*). Zároveň nejde například říci **a proto boj* nebo **protože napadení*.

Takovéto vymezení konektorů je ovšem problematické, pokud do diskurzní analýzy zahrneme i konektivní struktury typu *příčinou je, odůvodnil to...* atd. Zatímco některé z těchto struktur se v textu chovají stejně jako spojky či strukturující částice (viz následující bod a)), jiné se výrazně odlišují, a to právě s ohledem na povahu diskurzních argumentů, mezi kterými stojí (viz následující body b) a c)).

a) Diskurzní vztah mezi verbálními argumenty

Konektivní struktury, které vykazují podobné chování jako např. spojky či částice, mají schopnost **spojuvat dva úseky textu realizované verbálními argumenty** (a jsou tedy s příslušnými spojkami nebo částicemi vzájemně nahraditelné). Srov. příklad 59, ve kterém je spojení *z tohoto důvodu* zaměnitelné za konektor *proto* a význam zůstane zachován:

- (59) *Jak vyplynulo z vyšetřování, oba muži si přepadení vymysleli.
Z tohoto důvodu byli v těchto dnech z ČR vypovězeni.*

b) Nominalizace druhého diskurzního argumentu

Jiné konektivní struktury **dovolují nominalizaci druhého argumentu** (viz kapitola 3.4.2). Jinými slovy, za konektivní strukturou v těchto případech nenásleduje celá věta, ale nominální fráze – srov. příklad 60:

- (60) *Privatizované mlékárny se však zatím mezi sebou nedokázaly domluvit.
Důsledkem je nekompromisní konkurenční **boj**, který tlačí ceny výrobků až takřka
k nulové rentabilitě zpracovatelů.*

V daném příkladě se vyskytuje konektivní struktura *důsledkem je*, která vyjadřuje vztah příčiny a důsledku mezi verbálním argumentem *privatizované mlékárny se však zatím mezi sebou nedokázaly domluvit* a (rozvitým) nominálním argumentem *nekompromisní konkurenční boj, který tlačí...* Nominalizaci druhého argumentu dovolují především konektivní spojení s podobnou strukturou, jakou má fráze *důsledkem je* – srov. např. *důvodem je, příčinou je, podmínkou je, výsledkem je* atd., tj. konektivní spojení obsahující predikaci. Nominalizace druhého argumentu je v těchto případech alternativou k vyjádření argumentu vedlejší větou – srov. *Důsledkem je, že mezi sebou nekompromisně a konkurenčně bojují.*

Obdobně můžeme analyzovat další příklad:

- (61) *Trenér brazilského fotbalového celku FC Santos Serginho nesmí na lavičku svého
týmu 180 dní.
Důvodem je jeho **napadení** ředitele soupeřova týmu o přestávce zápasu Santos –
Guarani (0:4).*

I v tomto případě je druhý diskurzní argument realizován nominální frází (*napadení*), kterou je možné vyjádřit vedlejší větou – tj. *Trenér brazilského fotbalového celku ... nesmí na lavičku. Důvodem je, že **napadl** ředitele soupeřova týmu...*

Rozdíl mezi nominálním a verbálním vyjádřením druhého argumentu je podle našeho názoru syntaktický, nikoli sémantický. Z hlediska diskurzu jsou tedy obě varianty

rovnocenné¹². Z tohoto důvodu jsme se rozhodli neaplikovat na dané struktury výše uvedenou definici konektoru, podle které konektory spojují vždy dva verbální argumenty. Povahu druhého diskurzního argumentu (je-li nominalizace umožněna strukturou daného konektivního prostředku) tedy nezahrnujeme mezi principy přispívající k obecnému vymezení a definici diskurzních konektorů.

Z tohoto hlediska považujeme za konektor takové jazykové vyjádření, které spojuje dva diskurzní argumenty, ať už je druhý argument (vlivem struktury konektivního prostředku) na povrchu realizován slovesnou či nominální frází (jako je tomu v příkladu 58 – *Petr je nemocný a nemůže jet. Důsledkem je, že se výlet ruší / zrušení výletu.*).

Více k výskytům nominalizace u druhého argumentu v Pražském závislostním korpusu viz kapitola 5.1.2.2.

c) Nominalizace prvního diskurzního argumentu

Nominalizace se může vyskytovat i u prvního diskurzního argumentu, ačkoli v tomto případě není „vynucována“ strukturou diskurzního konektoru. Skutečnost, zda bude první argument slovesný či jmenný, je tedy čistě na výběru autora. Přesto jsme se rozhodli případy s nominalizovaným prvním argumentem z diskurzní analýzy nevykloučovat (a tedy nenásledovat teorie, podle kterých konektory spojují dva slovesné argumenty) – podle našeho názoru je i v tomto případě rozdíl mezi jmenným a slovesným vyjádřením argumentu formální – srov. příklad 62:

(62) *Dalším tématem našich zpráv je zvýšení cen benzínu.*

Kvůli tomu si některé rodiny nebudou moct dovolit jezdit autem.

V daném příkladě¹³ je diskurzní vztah příčiny a důsledku mezi prvním nominálním argumentem (*zvýšení cen benzínu*) a druhým verbálním argumentem (*některé rodiny si nebudou moct dovolit jezdit autem*) – příklad můžeme parafrázovat následujícím způsobem: *Zvýšily se ceny benzínu, a proto si některé rodiny nebudou moct dovolit jezdit autem.*

¹² Opomineme-li případné rozdíly v tom, že mezi oběma variantami může být rozdíl v jejich stylové platnosti (nominalizované, kondenzované vyjádření může být považováno za stylově vyšší, formálnější).

¹³ Teoreticky můžeme uvažovat i o případech, ve kterých by byly nominální oba diskurzní argumenty – srov. *Dalším tématem našich zpráv je zvýšení cen benzínu. Důvodem je světová krize.* V reálných textech PDT se nicméně podobný příklad neobjevil.

Vzhledem k tomu, že jsou obě varianty, tj. jmenná i slovesná realizace prvního argumentu, významově rovnocenné (*Dalším tématem našich zpráv je zvýšení cen benzínu.* vs. *Dalším tématem našich zpráv je, že se zvýšily ceny benzínu.*), zahrnujeme do diskurzívní analýzy i argumenty nominalizované a syntaktickou povahu diskurzívních argumentů tedy nezohledňujeme při vymezování diskurzívních konektorů (jako tomu bylo v první fázi zpracování diskurzívních vztahů v PDT, během které byly anotovány pouze vztahy mezi argumenty obsahujícími sloveso v určitém tvaru – srov. Poláková a kol., 2012a).

K určování rozsahu prvního argumentu pak můžeme využít skutečnost, že některé konektivní struktury se v povrchové rovině vyskytují ve spojení s anaforickým vyjádřením – pro některé z nich je tato kombinace dokonce obligatorní, tj. u sloves je dána valencí (např. *kontrastovat – to kontrastuje s tím; odůvodnit – odůvodnil to tím, že...*), u sekundárních předložek tím, jakou vazbu vyžadují (např. *díky – díky tomu*). Anaforický výraz má v těchto případech vždy koreferenční vztah k předchozímu argumentu – v příkladě 62 je tedy koreferenční vztah mezi *zvýšení cen benzínu* a *tomu*, v příkladě 63 je první diskurzívní argument slovesný, čili výraz *tomu* má koreferenční vztah k celé předchozí propozici:

- (63) *Máme určité kontaktní možnosti ve všech státech, odkud pocházeli zahraniční studenti v bývalém Československu.*
Díky tomu bychom tam mohli hledat uplatnění pro naše lidi, a naše licence je proto pojata dosti široce.

Tato skutečnost je využitelná pro automatické vyhledávání některých konektivních struktur (tj. především předložkových spojení typu *kvůli tomu, díky tomu*) v korpusech s anotovanými koreferenčními vztahy (jako je Pražský závislostní korpus) – tj. aby mělo dané spojení platnost sekundárního konektoru, musí splňovat dvě podmínky: **a) musí se pojit s anaforickým výrazem¹⁴, b) daný výraz musí odkazovat k celému předchozímu diskurzívnímu argumentu¹⁵**. Více k automatickému vyhledávání sekundárních konektorů typu *díky tomu* viz Rysová a Mírovský (2014).

¹⁴ Za diskurzívní konektor nepovažujeme např. spojení *kvůli kašli* v následujícím příkladu: *Marodila jsem celý měsíc. V noci jsem nemohla spát kvůli kašli. Kvůli se zde pojí s neanaforickým výrazem kašli.*

¹⁵ Za diskurzívní konektor nepovažujeme např. spojení *kvůli nimž* v příkladu *Anglická linka si vede už několik let statistiku problémů, kvůli nimž děti volají. Kvůli se v tomto případě pojí s anaforickým výrazem nimž, které ovšem neodkazuje k celému předchozímu argumentu, ale pouze k jeho části (k výrazu problémů). Spojení kvůli nimž zde tedy nevyjadřuje diskurzívní vztah příčiny a důsledku (srov. *Anglická linka si vede už několik let statistiku problémů, a proto děti volají.). Jinými slovy, děti volají na linku kvůli svým problémům a nikoli kvůli tomu, že si anglická linka vede už několik let statistiku problémů.*

Jak již bylo řečeno, diskurzí argument může mít podobu nominální fráze, u které ovšem nesmí dojít ke ztrátě žádné významové složky daného argumentu – srov. např. argumenty *zvýšení cen benzínu a zvýšily se ceny benzínu* jsou významově rovnocenné. V příkladě 57 z úvodu kapitoly (*Petr je nemocný a nemůže jet. Kvůli němu se výlet ruší.*) tomu tak ovšem není. Anaforická část konektivní fráze (*němu*) je v koreferenčním vztahu k výrazu *Petr*, tj. pouze k jedné složce celého diskurzího argumentu. Výrazy *Petr* a *němu* (které mezi sebou v textu mají koreferenční vztah) tedy nevyjadřují stav ani událost atd. (srov. Asher, 1993), jedná se pouze o jednu složku diskurzího argumentu. Z tohoto důvodu nebudeme podobné příklady do diskurzí analýzy zahrnovat (i když jsme si vědomi toho, že se koreferenční vztahy tohoto typu rovněž podílejí na celkové koherenci textu).

Diskurzí argumenty tedy v naší práci vymezujeme především sémanticky – diskurzí argument je abstraktní objekt, který může vyjadřovat stavy, události, situace atd. (viz Asher, 1993). Sémantika a syntax zde spolu ovšem úzce souvisejí – dané sémantické kategorie bývají realizovány především jednotlivými větami (či posloupností několika vět i celým odstavcem) nebo právě nominalizovanými frázemi. **Z hlediska syntaxe tak mají diskurzí argumenty nejčastěji podobu vět (jejichž základem je sloveso v určitém tvaru, a proto je nazýváme slovesnými argumenty) a jmenných frází (jejichž základem bývá nejčastěji slovesné podstatné jméno; tyto argumenty nazýváme nominálními).**

V některých případech mohou být diskurzími argumenty i objekty vyjádřené situací, k níž mohou konektory odkazovat (např. v počátečních výpovědích typu *Tak jsem doma. konektory zasazují výpověď do situačního kontextu*), viz kapitola 3.3.3.

3.7.2 Princip univerzálnosti – diskurzí konektory vs. volné konektivní fráze

Vymezení toho, co je a není diskurzí konektor, komplikuje fakt, že se jednotlivé konektivní struktury vyskytují ve více podobách či realizacích – srov. např. různé varianty předložkových frází obsahující sekundární předložku *díky* a anaforický výraz odkazující k předchozímu diskurzímu argumentu: *díky tomu, díky jeho chorobě, díky této skulině, díky všem těmto opatřením, díky pomoci příbuzných a přátel z USA, díky této preciznosti* aj. či různé realizace konektivních struktur obsahující slovo *důvod*: *jako důvod uvedla, jako důvod odstoupení komuniké ČSOB uvádí, důvodem je, důvodem poklesu je, důvodem tohoto návrhu je* (uvedené příklady pocházejí z PDT).

Z hlediska konektorové platnosti je podle našeho názoru rozdíl mezi spojeními *díky tomu* a *díky jeho chorobě* či mezi *důvodem je* a *důvodem poklesu je*, byť mají uvedené dvojice podobnou strukturu a všechny (anaforicky) odkazují k předchozímu diskurzivnímu argumentu a zároveň se všechny podílejí na celkové koherenci textu – srov. příklady 64 a 65:

(64) *Z ní [z krve, M. R.] se rodí jejich nesmrtelnost až do doby, než John začne neuvěřitelně rychle stárnout.*

Díky jeho chorobě je do osudového víru krvavých vášní a tajemství života i smrti neodvratně vtažena také mladá lékařka Sarah...

(65) *Pokud nebyly všechny k dispozici, některé programy vytvořila Giga sama, anebo podnítila jejich vývoj u jiných firem.*

Díky tomu má dnes k dispozici doslova podnikatelský komplet, který nabízí pod názvem Asistent podnikatele.

V obou příkladech je mezi danými diskurzivními argumenty vztah příčiny a důsledku (*John začne neuvěřitelně rychle stárnout, a tak je do osudového víru krvavých vášní ... vtažena také mladá lékařka Sarah; ... některé programy vytvořila Giga sama..., a tak má dnes k dispozici...*). Příklady se nicméně liší v povaze konektivního prostředku, který oba úseky textu spojuje.

Daná spojení se podle našeho názoru liší tím, zda jsou kontextově závislá (např. *díky jeho chorobě*) **nebo nezávislá** (*díky tomu*) – jinými slovy, zda je můžeme použít (do značné míry) **univerzálně pro vyjádření daného sémantického typu diskurzivního vztahu** (v tomto případě vztahu příčiny a důsledku). **Z tohoto hlediska mohou být spojení *díky jeho chorobě* či *důvodem poklesu je* použita pouze v omezeném množství kontextů** (tj. v prvním případě v textech o chorobách, ve druhém v textech o poklesu), **zatímco *díky tomu* či *důvodem je* nejsou kontextem limitována**. Ve spojení *důvodem je* není anaforický výraz na povrchu vyjádřen (oproti *důvodem poklesu* / *tohoto návrhu* aj.), a může proto být použit k vyjádření vztahu příčiny a důsledku (téměř) nezávisle na kontextu¹⁶. Obdobně ve spojení *díky tomu* si *tomu* jako deiktický výraz vyhledá své sémantické vztahy po zasazení do jakéhokoli kontextu. Slova jako *pokles*, *návrh*, *choroba*, *skulinka*, *odstoupení* aj. tuto schopnost nemají, tj. nemohou svůj význam zcela naplnit až po dosazení do konkrétního textu.

¹⁶ V úvahu nyní nebereme vhodnost daných konektivních prostředků v různých funkčních stylech, jde nám pouze o odlišení spojení typu *důvodem je* a *důvodem poklesu je* apod. ve funkci diskurzivních konektorů.

Z tohoto hlediska je tedy možné použít spojení *díky tomu* jako ukazatel diskurzního vztahu i v kontextu příkladu 66, zatímco *díky jeho chorobě* užít v příkladu 67 nemůžeme:

(66) *Z ní [z krve, M.R.] se rodí jejich nesmrtelnost až do doby, než John začne neuvěřitelně rychle stárnout.*

Díky tomu je do osudového víru krvavých vášní a tajemství života i smrti neodvratně vtažena také mladá lékařka Sarah...

(67) *Pokud nebyly všechny k dispozici, některé programy vytvořila Giga sama, anebo podnítila jejich vývoj u jiných firem.*

**Díky jeho chorobě má dnes k dispozici doslova podnikatelský komplet, který nabízí pod názvem Asistent podnikatele.*

Tato skutečnost je podle našeho názoru pro diskurzní konektory zásadní. Pro závislost či nezávislost konektivních spojení na kontextu budeme nadále používat označení **princip univerzálnosti**.

Princip univerzálnosti tedy hodnotí konektivní struktury podle míry jejich vhodnosti použití pro různé kontexty, tj. zda fungují v textu jako ukazatele daného diskurzního vztahu spíše univerzálně nebo příležitostně. Jinými slovy tento princip určuje – pokud máme několik odlišných kontextů, ve kterých se vyskytuje určitý sémantický typ diskurzního vztahu, např. příčiny a důsledku – do kolika těchto kontextů se daná konektivní struktura (se schopností vyjádřit tento konkrétní typ vztahu) hodí, tj. do jaké míry je „univerzální“.

Z tohoto hlediska chápeme například spojení *díky tomu* jako obecnější (obecněji v textech využitelné) než *díky jeho chorobě*. (K testování univerzálnosti vybraných konektivních výrazů a spojení viz kapitola 6.9.)

Univerzální platnost¹⁷ (či vhodnost použití ve více kontextech) chápeme jako klíčovou vlastnost diskurzních konektorů. Proto jsme se rozhodli stanovit hranici těchto výrazů právě zde. Termín **diskurzní konektor** tedy budeme nadále používat pro **výrazy či spojení, které slouží jako univerzální indikátory příslušných diskurzních vztahů (nezávisle na jejich slovnědruhové příslušnosti)**, tj. např. výrazy jako *ale; nebo; naopak; to je důvod, proč; kvůli tomu; příčinou je; to znamená aj*. Ostatní spojení, která fungují jako konektivní

¹⁷ Univerzálnost diskurzních konektorů ale zpravidla nikdy nedosahuje úplných 100 % – viz kapitola 6.9.

prostředky v závislosti na kontextu (tj. neuniverzálně, pouze příležitostně) budeme nazývat **volné konektivní fráze** (srov. např. *díky jeho chorobě; kvůli tomuto zvýšení; díky této skulince; díky této preciznosti; důvodem poklesu je; důvodem tohoto návrhu je*).

Nyní se vrátíme k příkladům 47–58 uvedeným v úvodu této kapitoly (tj. *Petr je nemocný a nemůže jet. Proto / Tedy / Na základě toho / Z tohoto důvodu / Kvůli tomu / Kvůli této skutečnosti / Kvůli jeho nemoci / Kvůli němu se výlet ruší. / Výlet se tak / tím / tímto ruší. / Důsledkem je zrušení výletu.*).

Podle výše popsaných kritérií můžeme jako univerzální konektory vyhodnotit výrazy *proto, tedy, tak, tím, tímto, na základě toho, z tohoto důvodu, kvůli tomu, kvůli této skutečnosti* a *důsledkem je* (většina vyjadřuje vztah příčiny a důsledku mezi dvěma verbálními argumenty, konektor *důsledkem je* umožňuje nominalizaci druhého argumentu).

Spojení *kvůli jeho nemoci* a *kvůli němu* za konektory nepovažujeme. Ani jedno z těchto spojení nemá univerzální platnost (tj. daná spojení jsou kontextově závislá a nemohou být univerzálně použita v textech různého tematického zaměření pro vyjádření diskurzního vztahu příčiny a důsledku). Spojení *kvůli němu* navíc neodkazuje k celému předchozímu argumentu, ale pouze k jeho části (tvar *němu* odkazuje k výrazu *Petrovi*).

3.7.3 Primární a sekundární konektory

Pro účely naší práce jsme se rozhodli (na základě předchozí sondy provedené na datech Pražského závislostního korpusu) **vymezit diskurzní konektory jako jazykové prostředky, které mají schopnost fungovat v textu jako (do značné míry) univerzální¹⁸ ukazatele diskurzních vztahů mezi dvěma diskurzními argumenty** (bez ohledu na syntaktickou realizaci těchto argumentů, která může být např. ovlivněna strukturou daného konektoru). Vzhledem k tomu, že daná kritéria splňují značně různorodé jazykové prostředky (srov. např. jednoslovné výrazy *a, také, totiž* a víceslovná spojení typu *to je důvod, proč; příčinou bylo* atd.), vyčleňujeme v rámci kategorie diskurzních konektorů dvě skupiny, pro které zavádíme označení **primární a sekundární konektory**.

¹⁸ Jako univerzální konektory chápeme výrazy, které jsou vhodnými ukazateli diskurzních vztahů pro velké množství různých kontextů – více viz kapitola 6.9.

3.7.3.1 Primární konektory

Jako primární konektory označujeme především **syntaktika s primární spojovací funkcí**, tj. výrazy patřící především mezi **spojky** (např. *a, ale, však, i*) a **částice strukturující text** nebo jeho části (vymezení slovních druhů přebíráme podle *Mluvnice češtiny 2* – srov. příklady strukturujících částic *totiž, přitom, navíc, pak, potom, nakonec, například* atd.), dále **některá vztažná zájmena** (např. *což*) a **vztažná zájmenná příslovce** (např. *proto, přičemž*)¹⁹.

Mluvnice češtiny 2 definuje spojky a vztažné spojovací výrazy jako jazykové prostředky, které „mají funkci spojovat výrazy nebo věty a explicitně tak vyjadřovat vnitrovětné a mezivětné vztahy koordinace a dominance (subordinace)“ (*Mluvnice češtiny 2*, 1986: 214). Částice strukturující text nebo jeho části (zkráceně strukturující částice) „fungují jako prostředky připojování, navazování, a to uvnitř vět a souvětí i mezi textovými jednotkami“ (*Mluvnice češtiny 2*, 1986: 237).

U zmíněných příslovcí a zájmen se spojovací funkcí dochází často k přechodu ve spojky či strukturující částice (srov. např. *proto, potom, pak, což*), dané výrazy tedy ztrácejí větněčlenskou platnost. Proces přechodu v jiný slovní druh ovšem nemusí být plně dokončen, může tak docházet k nesouladu při určování zmíněných slovních druhů. V *Mluvnici češtiny 2* se tak například o výrazu *proto* hovoří na některých místech jako o spojce (srov. „pro koordinační vztah důsledkový je základní spojka *proto*“ (*Mluvnice češtiny 2*, 1986: 215)), na jiných jako o vztažném příslovci, srov. „Spojovací funkci sdílejí se spojkami tzv. výrazy vztažné, relativní. Jsou to vztažná zájmena (*co, cokoli(v) [...]*), vztažná zájmenná příslovce (*co, dokdy [...], proto [...]*).“ (*Mluvnice češtiny 2*, 1986: 219).

Vzhledem ke zmíněnému překrývání v rámci slovních druhů (především spojek, částic a příslovcí) a ke skutečnosti, že u mnohých příslovcí se spojovací funkcí dochází k přechodu k jinému slovnímu druhu, budeme chápat primární konektory **především jako syntaktika, tj. výrazy bez větněčlenské platnosti**.

Vzhledem k tomu, že primární konektory nezasahují do syntaxe dané věty (v tom smyslu, že neplní funkci větného členu), mohou být většinou z textu **vypuštěny, aniž by se věta stala negramatickou nebo (v daném kontextu) ztratila některý svůj významový rys** – srov. příklad 68:

¹⁹ Za primární konektory současně považujeme také některé grafické značky (dvojtečku a pomlčku), číselné a písmenné výrazy sloužící ke strukturaci textu, nejčastěji k vyjádření výčtu (např. *I., 2., 3.; a, b, c; i, ii, iii*) a větnou negaci (pokud slouží k vyjádření diskurzivního vztahu).

- (68) *V dnešní době je zpravidla nepoužitelné všechno, co je starší než čtrnáct dní.*
Proto je dobré vracet se ke zdrojům, jejichž kvalitu potvrdil čas.

V daném příkladě se vyskytuje primární konektor *proto*, který explicitně signalizuje vztah příčiny a důsledku. Konektor může být v tomto případě vypuštěn bez jakýchkoli syntaktických zásahů do dané věty a diskurzní vztah zůstane zachován (tj. zůstane implicitní).

Primární konektory jsou v našem pojetí většinou **jednoslovné, gramatikalizované výrazy** (některé mohou tvořit korelativní páry), **z lexikálního hlediska ustálené, ve většině případů nedovolující modifikaci** (nelze například říci **obecně a, *jednoduše ale*, výjimku tvoří některé výrazy typu *hlavně proto*).

Primární konektory bývají v různých studiích (srov. např. Rouchota, 1996; Fischerová, 2006; Urgellesová-Collová, 2010 aj.) užívány jako typičtí zástupci diskurzních konektorů, jejich frekvence ve vyjadřování diskurzních vztahů je mnohem vyšší než u konektorů sekundárních (srov. kapitola 6.9). Z tohoto důvodu si můžeme **primární konektory představit jako centrum množiny diskurzních konektorů a sekundární konektory jako její periferii.**

3.7.3.2 Sekundární konektory

Sekundární konektory jsou jazykové struktury, které v textu plní spojovací funkci (stejně jako konektory primární). Na rozdíl od primárních konektorů jsou sekundární konektory zpravidla méně ustálené: srov. např. *důvod je, důvodem je* (sekundární konektory) vs. *proto* (primární konektor). Sekundární konektory jsou na rozdíl od primárních zároveň většinou víceslovné a v některých případech bývá obtížné určit jejich rozsah: tj. např. to, zda je konektorem struktura *to je důvod* nebo *to je důvod, proč* nebo se v tomto případě jedná o dva samostatné konektory.

Vzhledem k tomu, že sekundární konektory v textu fungují jako celé jazykové struktury, nemohou být jako konektory chápána jednotlivá slova, která je tvoří. Např. tedy slovo *důvod* není samo o sobě konektorem, platnost konektoru nabývá až celá struktura jako např. *z tohoto důvodu* či *důvodem je*. Slovo *důvod* navíc nemusí být nutně součástí konektivní

struktury, a v některých případech se tedy na strukturaci diskurzu jako univerzální konektivní prostředek nepodílí – srov. příklad 69:

- (69) *Zaměstnanci mají právo stávkovat pouze v případě, že se důvod stávky týká uzavření kolektivní smlouvy.*

V daném příkladě slovo *důvod* není součástí konektivní struktury, která by v textu mohla nabývat funkce konektoru.

Druhou skupinu konektorů tedy tvoří **zpravidla víceslovné struktury či fráze s funkcí univerzálních ukazatelů diskurzních vztahů**. Znamená to, že jednotlivá slova, která dané struktury tvoří, nemívají spojovací funkci ze své podstaty (jako spojky či strukturující částice), ale platnost konektoru nabývají až jako celá struktura. Z tohoto důvodu užíváme pro daná konektivní spojení označení sekundární konektory (srov. např. *podmínkou je, z tohoto důvodu, díky tomu, důvodem je* atd.).

Analýza rozsáhlých jazykových dat ukázala, že sekundární konektory tvoří širokou a rozmanitou škálu, a to zejména po stránce lexikální, syntaktické i sémantické (viz kapitola 6.1).

3.7.3.3 Prostupnost hranic mezi primárními a sekundárními konektory

Jak jsme popsali v předchozí podkapitole, v rámci diskurzních konektorů rozlišujeme dvě skupiny, tj. konektory primární a sekundární. Hranice mezi nimi je přitom podle našeho názoru prostupná.

Mezi sekundárními konektory můžeme vysledovat různé podskupiny, které jsou svou podstatou a chováním blízké konektorům primárním. Některé z nich mohou tuto hranici překročit a stát se konektory primárními.

Velkou skupinu sekundárních konektorů tvoří předložkové fráze, tj. struktury obsahující sekundární předložku (např. *díky, kvůli, vinou* atd.) a anaforický výraz odkazující k prvnímu diskurznímu argumentu. Jako sekundární konektory se tedy v textu objevují

spojení typu *díky tomu*²⁰, *kvůli tomu*, *vinou toho* atd. Tyto sekundární konektory jsou podle našeho názoru blízké konektorům primárním.

Podobnou strukturu má z historického hlediska totiž např. primární konektor *proto* vzniklý z předložky *pro* a zájmena *to*. Původně byl tedy tento konektor dvouslovný, posléze prošel gramatikalizací a ustálil se na jednoslovné formě *proto* s primární konektivní funkcí (tento vývojový proces je v češtině produktivní – viz analogické příklady typu *přesto*, *potom*, *předtím* apod.). Produktivnost tohoto tvoření dokládají protějšky českého *proto*, které vznikly podobným způsobem – srov. např. anglický konektor *therefore* (vzniklý spojením slova *there* a předložky *fore*, staroanglické a středoanglické formy předložky *for*; používaný ve významu ‚v důsledku toho‘) či podobný vývoj u konektorů *daarvoor* v holandštině, *dafür* v němčině či *derfor* v dánštině. Vývoj tohoto konektoru tedy proběhl v několika jazycích podobně, a není tudíž vázaný pouze na jeden konkrétní jazyk. V tomto ohledu by tedy obdobný vývojový proces mohl nastat i u dalších konektivních struktur, především u těch, které mají podobnou strukturu, jako měl původně konektor *proto*.

Z tohoto důvodu nechápeme primární a sekundární konektory jako dvě striktně oddělené kategorie, ale jako kategorie s prostupnou hranicí. **Sekundární konektory mohou postupně procházet procesem gramatikalizace, po jehož dovršení se z nich mohou stát konektory primární.**

3.7.3.4 Úvahy ke vzniku primárních konektorů

Jak jsme uvedli v úvodu kapitoly o diskurzních konektorech, primárními konektory rozumíme výrazy patřící především mezi spojky a strukturující částice (a některá jednotlivá příslovce a vztavná zájmena, která jsou ovšem na přechodu právě ke spojkám či částicím). Na základě předchozích úvah o prostupnosti hranic mezi primárními a sekundárními konektory a postupném procesu ustalování konektorů jsme se dále zabývali otázkou vzniku primárních konektorů obecně.

Přechodem mezi jednotlivými slovními druhy a vznikem spojek a částic se ve svých pracích zabýval např. Josef Zubatý (1909; 1910), soustavněji pak Jaroslav Bauer (1958; 1962a; 1962b; 1963; 1964).

²⁰ Jako sekundární konektor chápeme celou předložkovou frázi *díky tomu*, ne pouze *díky*, protože věta se samotným *díky* (srov. **díky jsem vyhrál* vs. *díky tomu jsem vyhrál*) by byla negramatická. *Díky* (stejně jako další sekundární předložky) se navíc musí pojit s anaforickým výrazem odkazujícím k předchozímu argumentu, aby výsledná předložková fráze získala diskurzní konektivní funkci. Zároveň některé současné primární konektory vznikly z podobných předložkových frází (srov. *přesto* z předložky *přes* a zájmena *to* apod.).

Podle Bauera vznikaly spojky z jiných slovních druhů, tj. dané výrazy či spojení nabývaly spojovací funkci postupně. Zatímco u některých spojek je tento proces ukončen, některé zůstaly na přechodu mezi dvěma slovními druhy. V těchto případech je pak obtížné slovnědruhovou příslušnost daného výrazu určit – srov.:

„Při klasifikaci neohebných slov podle slovních druhů se setkáváme s mnoha přechodnými případy. To platí zejména o spojkách, které vznikaly postupně v různých obdobích jazykového vývoje ze slov patřících k jiným slovním druhům: ze vztažných zájmen a příslovcí, z jiných příslovcí, z částic, někdy dokonce z různých tvarů ohebných slov, z několikaslovných výrazů nebo z celých vět. Původní funkce těchto slov se přitom často zachovala, a tak je někde těžko určit, jde-li o užití jiného slova ve spojovací funkci nebo naopak o užití spojky ve funkci nespojkové, anebo jde-li už jen o dvě stejně znějící slova s různými funkcemi a s různou slovnědruhovou příslušností, tedy o homonyma.“ (Bauer, 1964: 131)

Obtížně definovatelná je tak slovnědruhová příslušnost konektorů, které jsou na přechodu mezi příslovci a spojkami (či částicemi). V kapitole 3.7.3.1 jsme primární konektory definovali jako výrazy patřící mezi spojky a strukturující částice a několik dalších výrazů, které bývají obvykle řazeny mezi příslovce, ale které se postupně ustalují jako spojky či strukturující částice (tj. jako výrazy s čistě konektivní funkcí). Jaroslav Bauer hovoří o problematice určování slovních druhů (hlavně spojek a příslovcí) ve svém článku „Spojky a příslovce“²¹:

„V historickém vývoji češtiny – stejně jako jiných slovanských jazyků – docházelo často k tomu, že různá příslovce nabývala spojovací funkce a měnila se ve spojky. Někde se tento proces dokončil, jinde ještě probíhá nebo zůstal nedokončen. V těchto případech jsme pak často na rozpacích, jak takové spojovací slovo zařadit — je to spojka nebo příslovce? Naše mluvnice a slovníky se tu rozcházejí nebo se rozhodnutí vyhnou tím, že prostě mluví o „spojovacích výrazech“ bez bližší specifikace. Chybí zde jasné kritérium, které by umožnilo stanovit slovnědruhovou funkci spojovacího výrazu.“ (Bauer, 1962: 29)

²¹ Hranicemi mezi spojkami, příslovci a částicemi se ve svých pracích dále zabývá M. Vondráček (1998, 1999).

V člancích Jaroslava Bauera můžeme najít rovněž analýzu původu některých (v současné době typických) primárních konektorů – např. *a*, *ale*, *však*, *avšak*, *proto*, *přesto*, *tedy*, *tudíž*, *přece* aj. *A* bylo původně deiktické citoslovce s významem blížícím se novočeskému *hle* (Bauer, 1963: 25). *Ale* se z historického hlediska skládá ze dvou komponentů: *a* (s citoslovečným původem) a *-le* (s významem omezovacího příslovce *jen*), podobně se *a* spojovalo s *však* ve spojku *avšak* (Bauer, 1963: 26), přičemž samotné *však* mělo původně příslovečný význam podobný německému *jedenfalls* s významem ‚vždy‘ (Bauer, 1963: 25). Konektory *proto*, *přesto* jsou ustrnulé předložkové pády, jejichž spojovací funkce se vyvinula z funkce anaforické (tyto výrazy jsou podle Bauera na přechodu mezi příslovci a spojkami; Bauer, 1962: 30). *Tedy* má původ v časovém *tehdy*; *tudíž* v místním (a přeneseně způsobovém) *tudy*. *Přece* vzniklo ze spojení *před se* s původním významem ‚dále kupředu‘ (Bauer, 1962: 31).

Je tedy zřejmé, že dané primární konektory (dnes plně gramatikalizované) nebyly konektory od svého původu, ale staly se jimi postupným ustalováním formy a funkce (tj. nabýváním funkce primárně spojovací).

V procesu vzniku primárních konektorů tedy docházelo ke ztrátě či oslabování původního lexikálního významu příslušného slova (či jednotlivých složek daného spojení). (K oslabování lexikálního významu u sekundárních konektorů viz kapitola 3.7.3.5.)

Podstatou vzniku primárních konektorů je tedy postupné ustalování formy i funkce dané přechodem těchto výrazů či spojení (patřících zpravidla mezi jiné slovní druhy) ke spojkám či strukturujícím částicím a jejich následná gramatikalizace, přičemž tento proces není u některých z nich zatím dokončen.

Z tohoto hlediska můžeme chápat sekundární konektory (tj. výrazy typu *to je důvod*, *proč*) jako jazykové prostředky, u kterých můžeme sledovat proces postupného ustalování, který ovšem zatím není v takové fázi, ve které by mohlo dojít k jejich úplné gramatikalizaci. Jinými slovy bychom se mohli dívat na sekundární konektory jako na potenciální konektory primární (vzhledem k předchozímu historickému vývoji současných primárních konektorů).

Současné primární konektory tedy pravděpodobně vznikaly z podobných víceslovných frází či slovních druhů jako současné sekundární konektory. Tento jev můžeme sledovat i v jiných jazycích – srov. např. anglický konektor *because* (*protože*) vznikl ze spojení „by case“ (doslovně ‚podle / vzhledem k případu‘), konektor *nevertheless* (*nicméně*) z víceslovné fráze „never the less“ (ve významu ‚o nic méně‘, analogické k frázi „never the worse“ (ve významu ‚o nic horší/hůře‘)) či primární konektor *for* (*nebot*) vznikl pravděpodobně ze

staroanglických frází typu „for þon þy“ (doslovně „for the (reason) that“ – ‚z toho (důvodu), že‘) apod. (viz Harper, 2001). **V tomto ohledu nechápeme diskurzní konektory jako uzavřenou množinu výrazů, ale jako škálu jazykových prostředků v různém stupni gramatikalizace.**

3.7.3.5 K otázce plnovýznamovosti

V předchozí podkapitole jsme se zmínili o tom, že v procesu ustalování jistého výrazu či spojení ve funkci primárního konektoru dochází k oslabování či úplné ztrátě jeho původního lexikálního významu (např. *přece* ztratilo význam ‚dále kupředu‘, *a* význam ‚hle‘ apod.). Postupným procesem gramatikalizace tedy dané výrazy či spojení ztrácejí svoji plnovýznamovost (jak jsme popsali výše, spojky a částice, které považujeme za primární konektory, vznikly z jiných slovních druhů – především příslovcí a citoslovcí, z předložkových frází apod.). Původně se tedy jednalo o plnovýznamová slova či o spojení s plnovýznamovou složkou. Lexikální význam byl ovšem v těchto případech oslaben a dané výrazy se v jazyce ustálily jako synsémantika (příp. jsou stále v procesu ustalování, viz některá příslovce či relativní zájmena).

Tato skutečnost nás vedla k zamyšlení nad tím, co se děje s plnovýznamovostí u sekundárních konektorů, tj. spojovacích prostředků, které se v jazyce ještě neustálily natolik, aby mohly být považovány za primární konektory. V kapitole 3.7.3 jsme jako jeden z rozdílů mezi primárními a sekundárními konektory uvedli skutečnost, že primární konektory jsou synsémantika, zatímco konektory sekundární obsahují autosémantickou složku. Vzhledem k tomu, že sekundární konektory fungují v textu jako univerzální ukazatele diskurzních vztahů, dochází u nich tedy postupně k ustalování formy i funkce i k vyprazdňování původního lexikálního významu (blíže viz kapitola 6.1).

Dichotomii plnovýznamovosti vs. neplnovýznamovosti problematizuje ve svých studiích např. Jaroslav Bauer. V článku „Spojky a příslovce“ (1962) upozorňuje na problematicky zařaditelné výrazy, které jsou od původu příslovce, ale které nemají plný lexikální význam – srov.:

„Potíže působí zvláště zájmenná příslovce, která sama plný lexikální význam nemají a pouze odkazují na různé okolnosti místní, časové, příčinné a jiné. Proto nebývá zřetelná ani jejich členská platnost ve větě, a mohou být chápána i jako pouhé vyjádření vztahu mezi větami

(nebo větnými členy). Zvláště u původních zájmených příslovci vztahných je přechod ke spojkové platnosti velmi častý, neboť relativa jsou už ve své původní funkci slova spojovací.“ (Bauer, 1962: 29)

Oslabování vlastního věcného významu je vysvětlitelné právě přechodem daných výrazů ke spojkám či strukturujícím částicím, tj. k funkci primárních konektorů. Tento jev ale nenastává pouze na přechodu mezi příslovci a spojkami, ale u všech výrazů či spojení, které jsou v procesu ustalování a nabývání konektivní funkce – tímto vývojem tedy procházejí také další typy sekundárních konektorů.

U sekundárních konektorů tak dochází k oslabování lexikálního významu jednotlivých složek dané struktury a nový (tj. konektivní) význam nabývá celé spojení. Jinými slovy, u příslušných autosémantických složek je oslabován jejich původní význam ve prospěch fungování struktury jako celku. Některé z nich se v jazyce již ustálily natolik, že jsou významově téměř zcela nahraditelné primárními konektory – srov. např. sekundární konektory *z toho důvodu*, *z té příčiny*, *pro tu příčinu*, které *Slovník spisovného jazyka českého* (1989) uvádí jako významové ekvivalenty k primárnímu konektoru *proto*. Příslušné autosémantické složky samy o sobě konektivní funkci nemají a v textu plní roli samostatných, plnovýznamových slov. Plnovýznamovost je u nich oslabena právě v ustalujících se strukturách, které konektivní funkci získávají jako celek.

Pro ilustraci jsme ve *Slovníku spisovného jazyka českého* (1989) a *Slovníku spisovné češtiny* (2005) dále vyhledali slovo *směr*, které se objevuje v konektivní struktuře *v tom/tomto směru*. Dané heslo má v obou slovnících čtyři významy – tři se týkají samotného slova *směr* (tj. zkráceně ‚myšlená linie pohybu k cíli‘; ‚zaměření; souhrn názorů a zásad ve vědě‘, ‚umění‘), jeden význam se ovšem nevztahuje pouze k danému slovu, ale k celému spojení *v tom směru* – jako významové ekvivalenty jsou pro něj uváděna spojení typu *co se toho týče*, *v té věci* apod., tj. spojení, která podobně jako *v tom/tomto směru* přispívají ke strukturaci textu²². V obou slovnících je tedy konektivní funkce struktury *v tom/tomto směru* zachycena.

V kategorii sekundárních konektorů tedy dochází k ustalování formy, fixování konektivní funkce a oslabování vlastního věcného významu jednotlivých složek dané struktury. Tyto struktury pak mají jako celek v textu platnost sekundárních diskurzních konektorů.

²² Dané výrazy vyjadřují diskurzní vztahy ovšem pouze v případě, že odkazují k celému předchozímu argumentu.

3.7.3.6 Vzájemná nahraditelnost primárních a sekundárních konektorů

Jak už jsme naznačili výše, primární a sekundární konektory jsou v některých případech v textu téměř zcela zaměnitelné. Jedná se o příklady, ve kterých se vyskytuje diskurzni vztah mezi dvěma verbálními argumenty – srov. příklad z PDT:

- (70) *S ohledem na toto ustanovení by se hrubé chování muselo týkat vaší osoby a nestačí pouze nevhodné zacházení s předmětem darovací smlouvy, to je darem.*
Z tohoto důvodu by byla vaše žaloba na vrácení daru u soudu zamítnuta.

V daném příkladě je sekundární konektor *z tohoto důvodu* nahraditelný primárním konektorem *proto*.

V případě, že je druhý diskurzni argument nominalizovaný, náhrada sekundárního konektoru za primární si vynucuje syntaktické změny příslušného argumentu. Srov. příklad z PDT:

- (71) *V oblasti Kyjovska, v jihovýchodní části Slovácka – kraje krásných žen, folkloru a vína – stagnuje trh s nemovitostmi jako všude.*
*Příčinou je neexistence reálně splatných půjček obyvatelstvu, zejména hypoték na nákup nemovitostí pro bydlení, **nerealizace** bankovních zástav, dále pak **tvrdší úvěrová politika** bankovních domů a **plošný pokles** kupní síly obyvatelstva.*

V uvedeném příkladě struktura sekundárního konektoru *příčinou je* umožňuje nominalizaci druhého diskurzniho argumentu (tj. *neexistence reálně splatných půjček obyvatelstvu...*). Při náhradě sekundárního konektoru za primární konektor musí tedy v tomto případě dojít ke změně syntaktické povahy daného argumentu (protože primární konektory podobnou schopnost, tj. nominalizovat diskurzni argument, nemají) – srov.:

- (72) *V oblasti Kyjovska, v jihovýchodní části Slovácka – kraje krásných žen, folkloru a vína – stagnuje trh s nemovitostmi jako všude, protože **neexistují** reálně splatné půjčky obyvatelstvu, zejména hypotéky na nákup nemovitostí pro bydlení, **nejsou realizovány** bankovní zástavy, dále pak **panuje tvrdší úvěrová politika** bankovních domů a **plošně klesá** kupní síla obyvatelstva.*

Ačkoli diskurzivní vztah i význam jsou v obou příkladech zachovány, změna diskurzivního konektoru v podobných případech vyžaduje zároveň změny syntaktické.

V případě, že je sekundární diskurzivní konektor ve formě samostatné věty, jsou možnosti náhrady daného konektoru za konektor primární omezené. Plná náhrada je možná pouze tehdy, pokud příslušný primární konektor může být modifikovaný stejnými rozvitími, která se objevují ve struktuře daného sekundárního konektoru – srov. např. sekundární konektor *Důvod je jednoduchý*. může být nahrazen rozvitým primárním konektorem *jednoduše proto, že*, přičemž lexikální význam zůstane zachován. V případě komplexnějšího rozvití je náhrada primárním konektorem pouze částečná – některé lexikální složky nemohou být při transformaci zachovány – srov. sekundární konektor z PDT *Další důvod je složitější*.²³ nemůže být plně nahrazen rozvitým primárním konektorem **složitěji proto, že*.

Zde opět narážíme na problém určování hranic sekundárních konektorů – tj. například na to, zda můžeme za sekundární konektory považovat všechny tři následující příklady z PDT: *Důvod je jednoduchý; Důvod je složitý; Důvod je filozofický*. V tomto ohledu se nabízí otázka, zda v podobných případech nemůže být dalším kritériem (vedle principu univerzálnosti) právě nahraditelnost daných struktur primárním konektorem – tj. můžeme např. říci *jednoduše proto, že*, ale nikoli **složitě/filozoficky proto, že*. Proti tomu ovšem stojí argument, že u primárního konektoru *proto, že* může stát pouze omezená množina slov (srov. např. *jednoduše/pochopitelně/zřejmě/pravděpodobně/hlavně proto, že*) – tedy objevuje se např. se slovem *jednoduše*, ale ne se slovem *složitě*. Dále můžeme říci např. *Hlavním/zásadním/základním/podstatným důvodem bylo, že...* ale jen *hlavně proto, že*, nikoli **zásadně/podstatně proto, že* apod. (ačkoli přídavná jména *hlavní, zásadní, základní, podstatný* jsou významová synonyma).

Z tohoto důvodu nebereme nahraditelnost primárním konektorem jako zásadní kritérium při určování hranic sekundárních konektorů (je to pouze kritérium částečné, pomocné), držíme se především principu univerzálnosti. Tedy čím je rozvití v dané struktuře lexikálně obecnější (tj. univerzálnější), tím má větší šanci stát se v textu součástí univerzálního diskurzivního konektoru. Z tohoto hlediska strukturu *Důvod je filozofický*.

²³ Zajímavé je, že některá přídavná jména jsou v rámci sekundárního konektoru užívána ve formě shodného přívlastku, jiná v predikativní funkci. Můžeme tak např. říci *hlavním důvodem je*, ale ne **důvod je hlavní*, a naopak *důvod je složitý*, ale ne **složitým důvodem je*. Ve strukturách typu *důvod je složitý* má přídavné jméno hodnotící funkci (tj. hodnotí či charakterizuje důvod), ve strukturách typu *hlavním důvodem je* má funkci výběru (z možných důvodů vybíráme ten hlavní). Obdobně můžeme užít *udělal to z jednoduchých důvodů*, nikoli **udělal to z hlavních důvodů*. Přívlastek *hlavní* můžeme ovšem použít např. ve struktuře *jedním z hlavních důvodů*, tj. opět ve funkci výběru z několika možností.

nevyhodnocujeme jako sekundární konektor, zatímco struktury *Důvod je jednoduchý*. a *Důvod je složitý*. ano (i když *Důvod je složitý*. není plně nahraditelný primárním konektorem).

Náhrada sekundárního konektoru konektorem primárním může být ovšem podle našeho názoru vhodným kritériem při určování rozsahu sekundárního konektoru, a to v tom smyslu, jaké složky patří ještě ke konektoru a jaké již k argumentu – tj. při řešení otázky, zda je sekundárním konektorem *to je důvod* nebo *to je důvod, proč, příčinou je* nebo *příčinou je skutečnost, že, zdůvodnil to* nebo *zdůvodnil to tím, že* apod. – srov. příklady z PDT:

(73) *Jako téměř vždy, když jde o velmi lukrativní zakázku, ozvaly se po schválení výsledků městskou radou námitky těch, kdo nezvítězili a cítí se poškozeni.*

To je důvod, proč ještě nebyla podepsána smlouva s dodavatelem, či přesněji organizátorem rekonstrukčních prací už třetí měsíc uzavřeného objektu.

(74) *Gyula Horn se vyslovil pro možné zavedení majetkové daně.*

Zdůvodnil to tím, že utahování opasků se nemůže vztahovat pouze na lidi žijící ze mzdy.

V obou případech je mezi danými argumenty diskurzivní vztah příčiny a důsledku vyjádřený sekundárními konektory *to je důvod, proč* a *zdůvodnil to tím, že*. Při ověřování rozsahu těchto konektorů (tedy hranice mezi sekundárním konektorem a diskurzivním argumentem) nám může pomoci náhrada konektorem primárním – srov. příklady 75 a 76:

(75) *Jako téměř vždy, když jde o velmi lukrativní zakázku, ozvaly se po schválení výsledků městskou radou námitky těch, kdo nezvítězili a cítí se poškozeni.*

Proto ještě nebyla podepsána smlouva s dodavatelem, či přesněji organizátorem rekonstrukčních prací už třetí měsíc uzavřeného objektu.

(76) *Gyula Horn se vyslovil pro možné zavedení majetkové daně, protože utahování opasků se nemůže vztahovat pouze na lidi žijící ze mzdy.*

Za primárními konektory (v uvedených příkladech *proto* a *protože*) následují v textu diskurzivní argumenty daného vztahu (tj. v příkladě 75 *ještě nebyla podepsána smlouva s dodavatelem*, v příkladu 76 *utahování opasků se nemůže vztahovat pouze na lidi žijící ze mzdy*). Všechny

ostatní složky, které se za primárními konektory v daných příkladech neobjevují, ale vyskytují se v příkladech 73 a 74, patří k sekundárnímu konektoru. Sekundárními konektory jsou tedy v uvedených příkladech spojení *to je důvod, proč* a *zdůvodnil to tím, že*.

Náhradu sekundárního konektoru konektorem primárním tedy budeme v naší práci používat jako obecné kritérium k ověřování hranice mezi příslušným konektorem a diskurzním argumentem.

Problematickými částmi budou přitom pravděpodobně výrazy typu *že, proč* uvozující vedlejší věty (viz struktury *zdůvodnil to tím, že; to je důvod, proč* aj.). Vedle náhrady primárním konektorem můžeme v těchto případech zohlednit i hledisko historické, tj. historický vývoj některých primárních konektorů typu *jestliže, když, až, poněvadž, jelikož* apod., jejichž součástí je zesilovací příklonka *-ž(e)* (srov. Holub a Kopečný, 1952; Rejzek, 2001).

3.7.3.7 Souvýskyt primárních a sekundárních konektorů – složené konektory

Předběžná analýza dat z Pražského závislostního korpusu ukázala, že se sekundární konektory mohou objevovat ve spojení s konektory primárními, dané spojení pak v textu (ve většině případů) funguje jako jeden složený konektor (vyjadřující jeden diskurzní vztah) – srov. příklad 77:

(77) *Téměř každý vytěžený diamant má kvalitu drahokamu, a to je důvod, proč se tak nákladná těžba z moře firmě vyplácí.*

V uvedeném příkladě je diskurzní vztah příčiny a důsledku vyjádřený spojením primárního konektoru *a* a sekundárního konektoru *to je důvod, proč*. *To je důvod, proč* je v tomto případě nahraditelné primárním konektorem *proto* (po náhradě pak vznikne obvyklý konektivní prostředek *a proto*).

Ke spojování primárního a sekundárního konektoru dochází většinou u vnitrovětných vztahů, nicméně může se objevit i u vztahů mezivětných – srov. příklad 78, ve kterém se vyskytuje primární konektor *a* se sekundárním konektorem *to znamená, že* (dané spojení je možné nahradit konektory *a proto, a tak*):

(78) *Býval šéfem tajné služby.*

A to znamená, že na své soky leccos ví.

Je tedy zjevné, že sekundární konektory vykazují v daných příkladech podobné chování jako konektory primární, tj. mohou se v textu objevovat jako složené výrazy, přičemž nejčastějšími složkami jsou spojky *a*, *ale*, *nebo* (srov. primární konektory *a proto*, *ale také*, *nebo dokonce*, *a tudíž* atd.).

Mluvnice češtiny 2 používá pro daná spojení termín spojkové výrazy. Kombinovat se přitom mohou: 1. dvě, příp. více spojek (*ale i, přesto však* aj.), 2. spojka s výrazem patřícím k jinému slovnímu druhu (zpravidla k příslovcím a částicím), přičemž daný výraz bývá často na přechodu ke spojkám (srov. např. *a dokonce*, *a přece* aj.), 3. dva, příp. více výrazů, které nepatří mezi spojky, ale zpravidla mezi příslovce nebo částice (*ba právě naopak* apod.) (*Mluvnice češtiny 2*, 1986: 221, 222).

Jak jsme popsali výše, s primárními konektory se mohou kombinovat i konektory sekundární, přičemž vznikají analogická spojení typu *a z tohoto důvodu*; *a to je důvod*, *proč* apod.

Ve spojování sekundárních konektorů s konektory primárními můžeme vidět také analogii k historickému vývoji některých primárních konektorů – viz konektory *ale*, *avšak*, které vznikly ze dvou komponentů – z *a* (které v době sloučení bylo pravděpodobně již spojkou či na přechodu ke spojce) a z příslovce *-le* (s významem omezovacího *jen*) a ve druhém případě opět z *a* a ze *však* (s významem *vždy*) (viz Bauer, 1963). Se spojkou *a* se pojí i další výrazy typu *a proto*, *a tak*, *a tudíž* aj., které s *a* nejsou zatím spojeny těsně (tj. píší se zvlášť), ale které se v textu vyskytují s vysokou frekvencí a jsou v současné době v PDT vnímány jako jeden konektor.

Konektivní struktury typu *z tohoto důvodu*, *to znamená*, *to je důvod*, *proč* aj. vykazují v tomto ohledu podobné chování jako *-le*, *však*, *proto* atd. Tato skutečnost podle našeho názoru podporuje myšlenku, že se jedná o spojení s platností diskurzních konektorů, u kterých postupně probíhá proces ustalování formy a funkce.

V rámci anotace diskurzních vztahů v PDT jsou tyto případy, ve kterých byly příslušné vztahy signalizovány primárním a sekundárním konektorem (tj. příklady typu *a z tohoto důvodu*) anotovány jako složený konektor vyjadřující jeden diskurzní vztah.

V naší práci tedy kromě primárních (gramatikalizovaných) a sekundárních (dosud ne zcela gramatikalizovaných) konektorů rozlišujeme ještě konektory **složené** a **nesložené**.

Nesložený konektor v sobě zahrnuje (ze synchronního hlediska) konektor pouze jeden (ať už primární nebo sekundární), např. *a*; *ale*; *avšak*; *z tohoto důvodu*; *za těchto podmínek*. Z příkladů je patrné, že nesložený konektor může být jednoslovný i víceslovný. **Složený konektor** v sobě obsahuje více než jeden konektor (ať už jde o kombinaci dvou primárních konektorů nebo kombinaci primárního a sekundárního konektoru), např. *a tak*; *i když*; *a proto*; *a z tohoto důvodu*. Složené konektory jsou vždy víceslovné.

V některých případech může vyvstat otázka, zda je dané slovní spojení složeným konektorem nebo není. U složených konektorů (typu *a proto*, *i když*) dochází podle našeho názoru k oslabení jedné významové složky (často se jedná o konektory *a*, *i* s konjunktivním významem), tj. složený konektor vnímáme jako jazykový prostředek sloužící k vyjádření primárně jednoho typu významového vztahu (např. tedy složený konektor *a proto* vyjadřuje „silněji“ či zřetelněji vztah příčiny a důsledku než vztah konjunkce, *i když* vztah přípustky apod.). U daných konektorů tedy můžeme pozorovat postupný proces gramatikalizace – *a*, *i* v těchto případech postupně ztrácejí svůj původní význam a stávají se součástí konektoru vyjadřujícího jiný typ diskurzního vztahu. Z historického hlediska můžeme podobný jev pozorovat u již plně gramatikalizovaných, jednoslovných konektorů, jako jsou *ale*, *avšak*, *anebo*, které se původně skládaly také z více složek, přičemž u konjunktivního *a* došlo postupně k oslabení původního významu a k jeho spojení s následujícím konektorem. V současné době tedy tyto konektory vyjadřují v textu zcela zřetelně vždy pouze jeden vztah – *ale*, *avšak* vztah opozice, *anebo* vztah disjunktivní či konjunktivní alternativy. Významový vztah konjunkce byl v těchto případech zcela potlačen. Jako složený konektor tedy chápeme výrazy skládající se z více konektorů, u kterých dominuje jeden významový vztah, viz příklad 79.

Pokud si každý ze společně se vyskytujících konektorů zcela zřetelně ponechává svůj původní význam (tj. na daném místě textu silně vnímáme dva různé významové vztahy), považujeme je za navzájem různé diskurzní konektory a u každého z nich označujeme daný významový typ a dosah diskurzního vztahu zvlášť, viz příklad 80.

(79) *Valentýn je za rohem, a proto jsme pro Vás připravili pár tipů na akční ceny hotelů v Praze na valentýnský víkend.*

(80) *Kdysi jsem si život užíval, ale potom jsem se připojil na internet.*

V příkladu 79 je mezi dvěma slovesnými argumenty (*Valentýn je za rohem a připravili jsme pro Vás pár tipů na akční ceny hotelů v Praze na valentýnský víkend*) zřetelnější diskurzí vztah příčiny a důsledku (než vztah konjunkce, kterou typicky vyjadřuje konektor *a*), signalizovaný složeným konektorem *a proto*.

Naopak v příkladu 80 jsou mezi slovesnými argumenty (*kdysi jsem si život užíval a připojil jsem se na internet*) zřetelné (či rovnocenné) dva různé významové vztahy – vztah opozice vyjádřený konektorem *ale* a vztah nesoučasnosti (předčasnosti a následnosti) vyjádřený konektorem *potom*. Spojení konektorů *ale potom* tedy v tomto případě nepovažujeme za jeden složený konektor, ale za dva konektory samostatné.

Některé konektory mohou být zároveň **modifikované**²⁴ – srov. např. *hlavně proto*; *právě kvůli tomu*; *zjednoduchých důvodů*; *vlastně proto, že*; *především pak*; *jediným důvodem bylo*; *s jasným cílem*, *možná díky tomu* apod. Modifikace se přitom (vzhledem k jejich lexikálně-syntaktické povaze) vyskytuje častěji u sekundárních konektorů.

V některých případech pak může být obtížnější určit, zda se u víceslovného konektoru jedná o konektor složený či modifikovaný. Rozlišovacím kritériem je tedy to, že složené konektory se skládají z více konektorů (tj. všechny části složeného konektoru jsou samy o sobě konektorem – srov. *a tak*, *a proto*, *ale naopak*, *a dokonce*), modifikované konektory naopak obsahují výraz, jehož hlavní funkce není spojovací, ale modifikační, tj. svým významem upřesňuje příslušný diskurzí vztah (říká například, že je daná podmínka hlavní, příčina jednoduchá, cíl jasný apod.). Význam těchto výrazů je tedy především hodnotící (srov. *hlavně*, *jednoduchý*, *jediný*, *jasný*...) či modální (např. *možná*, *vlastně*...).

V naší práci tedy vymezujeme **tři základní opozice**, podle kterých je možné diskurzí konektory členit. Konektory mohou být:

a) primární (*však*, *a*, *nebo*, *ale*, *proto*...) **či sekundární** (*důvodem je*, *kvůli tomu*, *s jasným cílem*) – hlavním rozlišovacím kritériem je v tomto případě míra gramatikalizace daného výrazu;

b) složené (*a tudíž*, *a kvůli tomu*, *ale naopak*) **či nesložené** (*nebo*, *ale*, *z tohoto důvodu*) – složené konektory se skládají z více konektorů, nesložené se skládají (ze synchronního hlediska) z jednoho konektoru (nesložené konektory mohou být jednoslovné (*ale*))

²⁴ Modifikované konektory jsou méně univerzální pro užití v dalších kontextech než konektory nemodifikované.

i víceslovné (*z tohoto důvodu*)); složené konektory se v textu mohou vyskytovat vedle sebe (*a proto*), nebo odděleně (*bud' nebo, když tak, v případě, že pak*), v tomto případě tvoří tzv. korelativní páry;

c) modifikované (*hlavně proto, s jasným cílem, jedině kvůli tomu, především vinou toho*) **či nemodifikované** (*proto, s cílem, kvůli tomu, vinou toho*) – modifikované konektory obsahují hodnotící či modální výraz, který upřesňuje význam příslušného diskurzního vztahu.

Uvedené aspekty se mohou vzájemně kombinovat – tj. např. složený konektor může být zároveň modifikovaný apod. (srov. *a hlavním důvodem bylo, a právě proto*). Ve složený konektor se mohou spojovat konektory primární (*a proto*) i primární a sekundární (*a z tohoto důvodu*).

3.7.4 Zpracování diskurzních vztahů a „sekundárních konektorů“²⁵ ve velkých korpusech

Hlavním cílem této práce je detailně analyzovat sekundární konektory v češtině na datech Pražského závislostního korpusu. V následující části nejprve popíšeme, jak se sekundárními konektory pracují jiné velké korpusy s diskurzni anotací. Posléze představíme vlastní anotaci jazykového materiálu na celých datech korpusu PDT (tj. provedenou přibližně na 50 000 vět) a poté vyvodíme o sekundárních konektorech další závěry a podrobně popíšeme jejich chování v reálných textech.

Pražský závislostní korpus je zatím jediným korpusem češtiny, který obsahuje anotaci diskurzních vztahů (v současné době pouze vyjádřených explicitně, tj. primárními a sekundárními konektory). Pro ostatní jazyky ovšem podobných korpusů existuje více. V následující kapitole představíme hlavní z nich. Zaměříme se přitom na různé způsoby, jakými je možné diskurzni vztahy ve velkých korpusech zpracovávat. Jednotlivé přístupy či pojetí se liší především v tom, zda se pokoušejí o zachycení textu („shora“) v jedné spojitě struktuře (a neanotují jednotlivé jazykové prostředky, které vyjadřují vztahy v rámci této struktury, tj. hlavně konektory – srov. např. teorii RST) či zda jejich zpracování vztahů v textu vychází právě („zdola“) od identifikace a anotace diskurzních konektorů jako signalizátorů textových vztahů (srov. např. Penn Discourse Treebank). Mezi další rozdíly patří například to, zda jsou diskurzni vztahy anotovány na lineárních textech či ve stromových strukturách, jaké typy vztahů jsou značeny (tj. implicitní, explicitní apod.) či jak jsou definovány diskurzni konektory.

3.7.4.1 RST Discourse Treebank

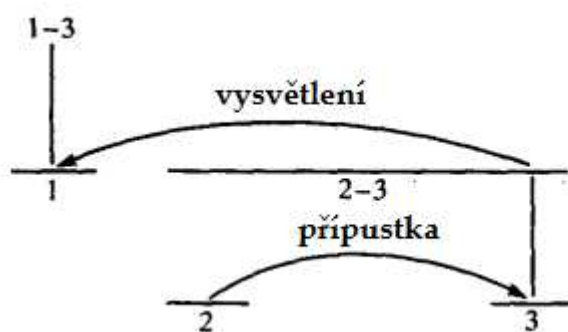
Jedním z prvních velkých korpusů, který obsahuje anotaci diskurzních vztahů, je RST Discourse Treebank (Carlson a kol., 2001). Tento korpus obsahuje 385 textů (anglických článků z časopisu Wall Street Journal). Diskurzni vztahy jsou zde anotovány podle teorie Rhetorical Structure Theory (RST, srov. Mann a Thompsonová, 1988) chápající text jako síť vztahů mezi jednotlivými textovými jednotkami, které se spojují ve vyšší celky (tj. text je chápán jako hierarchická struktura). Zpracování diskurzních vztahů v tomto pojetí (na rozdíl od Pražského závislostního korpusu) tedy spočívá ve vytváření spojitě struktury zachycující

²⁵ Termín sekundární konektory zavádíme nově v naší práci. V jiných korpusech budou tedy podobné výrazy nazývány jinak.

příslušné úseky textu a vztahy mezi nimi (které mohou být vyjádřeny explicitně či implicitně – konkrétní jazykové prostředky signalizující příslušné vztahy ovšem v tomto pojetí značeny nejsou) – srov. následující úsek textu (přeloženo z originálu viz Mann a Thompsonová, 1988: 252):

1. *Další den hudby je naplánován na sobotu 21. července (poledne – půlnoc).*
2. *Více informací vám zašleme později.*
3. *Ale bylo by dobré, abyste si tento den označili v kalendáři.*

Struktura daného textu by byla podle teorie RST zaznačena následujícím způsobem – srov. obrázek 2. Mezi větami 2 a 3 je podle této teorie vztah přípustky, tyto věty se spojí ve vyšší celek, který svým obsahem vysvětluje obsah věty první.



Obrázek 2: Ukázka anotace textů podle teorie RST (viz Mann, Thompsonová, 1988: 252)

Z teorie RST vycházejí další přístupy ke zpracování diskurzu, např. teorie Segmented Discourse Representation Theory (SDRT, Asher, 1993, později Asher a Lascarides, 2003), která usiluje o vytvoření formálně precizního teoretického rámce pro zachycení struktury diskurzu. Na základě teorie SDRT vznikly např. anotované korpusy a anotační projekty DISCOR (Reese a kol., 2007), ANNODIS (Afantenos a kol., 2012) či STAC (Asher a kol., 2012).

3.7.4.2 Discourse Graphbank

Korpus Discourse Graphbank (Wolf a kol., 2004) obsahuje 135 anglických textů z časopisu Wall Street Journal a AP Newswire. Autoři tohoto projektu (Wolf a Gibson, 2005) kritizují zachycování diskurzivních vztahů ve stromových strukturách (jako je tomu např.

v korpusu RST Discourse Treebank). Podle nich stromové struktury neumí zachytit případy, ve kterých jsou mezi dvěma úseky textu dva různé významové vztahy. Projekty typu RST Discourse Treebank nutí anotátory rozhodnout se pouze pro jeden z těchto typů, a tím anotaci diskurzu ochuzují – srov. příklad 81:

(81) *Chtěli jsme jít do kina, ale pak jsme zůstali doma.*

V daném příkladu jsou zároveň dva diskurzí vztahy – vztah opozice a nesoučasnosti (předčasnosti a následnosti) vyjádřené konektory *ale* a *pak*. V projektech, jako je RST Discourse Treebank, by se tedy anotátor musel rozhodnout pouze pro jeden z nich. Pojetí diskurzí vztahů v korpusu Discourse Graphbank naopak překrývání určitých typů významových vztahů dovoluje (podobně jako PDT). Textové vztahy v Discourse Graphbanku proto také nejsou zachycovány ve stromových strukturách, ale ve „volnějších“ grafech (srov. Wolf a Gibson, 2005).

3.7.4.3 Penn Discourse Treebank a navazující projekty

Dalším významným projektem (představujícím odlišné pojetí zpracování diskurzí vztahů než v korpusech RST Discourse Treebank a Discourse Graphbank) je pensylvánský korpus Penn Discourse Treebank (Prasadová a kol., 2008). Penn Discourse Treebank je projekt navazující na syntaktický korpus Penn Treebank. Obsahuje přibližně 49 000 vět, texty pocházejí z časopisu Wall Street Journal. V korpusu Penn Discourse Treebank nejsou diskurzí vztahy zachyceny ve stromových strukturách či grafech, ale na lineárních textech. Zpracovány jsou přitom diskurzí vztahy vyjádřené explicitními konektory, alternativními lexikálními vyjádřeními diskurzí konektorů (zkráceně altlexy), diskurzí vztahy vyjádřené implicitně, dále tzv. nulové vztahy (tj. případy, ve kterých není mezi dvěma sousedícími jednotkami žádný diskurzí vztah, tedy např. místa, kde začíná nový odstavec apod.) a vztahy mezi jmennými entitami. Diskurzí konektory jsou zde chápány jako výrazy otevírající pozice pro dva diskurzí argumenty (diskurzí vztah je tedy veden mezi dvěma úseky textu nazývanými argumenty).

Penn Discourse Treebank je významný mimo jiné proto, že se jím inspirovala celá řada dalších projektů zaměřených na korpusové zpracování diskurzí vztahů v různých jazycích – srov. diskurzí projekty pro hindštinu (Oza a kol., 2009), arabštinu (Al-Saifová a Markertová, 2010), turečtinu (Zeyreková a kol., 2010), francouzštinu (Danlosová a kol.,

2012), čínštinu (Zhou a Xue, 2012) či němčinu (Stede a Neumann, 2014). Způsob anotace korpusu Penn Discourse Treebank byl použit také při zpracování italských mluvených dialogů (Tonellová a kol., 2010) či na anglických biomedicínských textech (Prasadová a kol., 2011) a navazuje na něj také zpracování diskurzních vztahů v Pražském závislostním korpusu.

3.7.4.4 Další významné diskurzní korpusy

V posledních letech vznikají další diskurzní projekty, které se snaží zpracovat diskurzní vztahy v textech různých žánrů, stylů apod., často ve spojení s anotací dalších diskurzních jevů (jako je např. anafora a koreference), v některých případech dokonce ve více jazycích zároveň. Vznikají tak projekty, jako je např. GECCo (Lapshinova-Koltunski a kol., 2012) – jedná se o korpus anglických a německých textů (psaných i mluvených) obsahující anotaci diskurzních vztahů a koreference. Projekt DisFrEn (Crible, 2014) obsahuje anotaci diskurzních vztahů na mluvených textech francouzštiny a angličtiny, projekt LOCAS-F (Degandová a kol., 2014) na mluvených textech francouzštiny, CSTNews Corpus (Cardoso a kol., 2011) na psaných textech brazilštiny a portugalštiny, TüBa-D/Z treebank (Gastelová a kol., 2011) na psaných německých textech. V současné době se připravuje korpus LSFb (French Belgian Sign Language; Meurant a Sinteová, 2013), tj. francouzsko-belgický korpus znakového jazyka, který bude obsahovat (mimo jiné) anotaci některých diskurzních jevů – mezivětných a vnitrovětných konektivních prostředků.

Je tedy zřejmé, že zpracování diskurzních jevů se v současné době věnuje celá řada zahraničních projektů. Diskurzní vztahy jsou přitom zpracovávány v psaných i mluvených textech, v různých jazycích i funkčních stylech.

3.7.4.5 Zpracování „sekundárních konektorů“ ve velkých korpusech

Všechny zmíněné korpusy obsahují zpracování diskurzních vztahů – jak jsme již ale naznačili, jednotlivá pojetí se poněkud liší, a to například ve způsobu zachycování diskurzních vztahů, v klasifikaci jednotlivých významových vztahů či ve vymezení pojmu diskurzní konektor.

Vzhledem k tomu, že cílem této práce je především výzkum sekundárních konektorů, zabývali jsme se dále tím, zda zmíněné korpusy zahrnují také anotaci těchto výrazů. Problematická je ovšem v tomto případě jejich definice – u většiny korpusů se uvádí, že korpus obsahuje zpracování diskurzních konektorů, v mnoha případech ovšem chybí jejich

přesné vymezení. Není tedy (mimo jiné) jasné, zda termín diskurzí konektory zahrnuje v daných pojetích také konektory sekundární (tak, jak je definujeme my v kapitole 3.7) či nikoli. Patrně nejbližším termínem k sekundárním konektorům je pojem alternativní lexikální vyjádření diskurzí konektorů (zkráceně altlexy) používaný v pensylvánském korpusu Penn Discourse Treebank. Vzhledem k tomu, že je tento termín ve světové literatuře poměrně rozšířený (projekt Penn Discourse Treebank má mnoho pokračovatelů), pojem diskurzí konektory bude v mnohých případech spíše zahrnovat pouze konektory primární.

Zatímco tedy v této práci používáme termíny primární a sekundární konektory, u většiny zmíněných projektů by odpovídaly nejspíše termínům diskurzí konektory a jejich alternativní lexikální vyjádření, tj. altlexy (vymezení konektorů a altlexů podle Penn Discourse Treebank je ale v některých ohledech jiné než naše vymezení primárních a sekundárních konektorů, viz níže kapitola 3.7.4.5).

Termín alternativní lexikální vyjádření diskurzí konektorů tedy není s termínem sekundární konektory totožný. V projektu Penn Discourse Treebank (zkráceně PDTB) jsou altlexy pojímány jako výrazy s konektivní diskurzí funkcí, které nejsou na seznamu diskurzí konektorů sestaveném pro anotaci diskurzí vztahů v PDTB (srov. Prasadová a kol., 2010: 1025–1026). Vymezení diskurzí konektorů a jejich alternativních vyjádření zde tedy nemá hlubší teoretické opodstatnění, ale vychází spíše z potřeb anotací. Altlexy začaly být v PDTB značeny díky anotaci implicitních vztahů – anotátoři dostali pokyn ke každému implicitnímu diskurzímu vztahu (definovanému jako vztah, který není signalizován žádným konektorem z předem sestaveného seznamu) přiřadit konektor, který by byl pro daný kontext nejvhodnější. Anotátoři ovšem nacházeli místa, kde by doplnění případného konektoru bylo nadbytečné (srov. např. *to je proto důvod, proč* – vztah příčiny a důsledku je zde signalizován dvakrát). Tato místa, ve kterých by bylo doplnění konektoru z předem sestaveného seznamu konektorů nadbytečné, byla v PDTB anotována jako altlexy.

Pojem alternativní lexikální vyjádření konektorů zahrnuje v PDTB širší množinu výrazů než pojem sekundární konektory – v rámci altlexů se například nerozlišují univerzální a neuniverzální konektivní struktury (tj. jako altlex označují (v našem pojetí) sekundární konektory, např. *that is why (to je, proč)*, i volné konektivní fráze, např. *the increase was due mainly to (zvýšení bylo hlavně kvůli tomu)*, viz Prasadová a kol., 2010: 1026).

Jako altlex je označen jakýkoli výraz signalizující příslušný diskurzí vztah, který není na původním seznamu konektorů. Jak autoři sami přiznávají (srov. Prasadová a kol., 2010: 1026; Prasadová a kol., 2014: 926–927), může se tak zároveň stát, že je jako altlex označen také výraz, který by díky své lexikálně-syntaktické povaze měl být považován za konektor,

ale vzhledem k tomu, že se nevyskytuje na původním seznamu konektorů, je v konečné anotaci zachycen jako altlex. Podle autorů projektu se tato skutečnost týká dokonce 15 % všech anotovaných altlexů (srov. Prasadová a kol., 2014: 926–927). Jako altlexy tak byly v PDTB například označeny výrazy *too (také)*, *thereafter (poté)*, *even (dokonce)*, *still (přesto)*, *only (jen)* atd., tj. výrazy, které bychom v našem pojetí řadili mezi primární konektory. Hranice mezi konektory a altlexy tak v PDTB není dána hlubším teoretickým výzkumem, ale je určována předem stanoveným seznamem konektorů.

Anotace alternativních lexikálních vyjádření konektorů není v PDTB plně systematická a konzistentní. Jako altlex bylo označeno celkem 624 výrazů (tj. 1,5 % ze všech anotovaných diskurzních vztahů v PDTB). Jak autoři sami připouštějí (srov. Prasadová a kol., 2010: 1028–1029), jednotlivé typy konektivních struktur ovšem nejsou anotovány konzistentně (autoři projektu vybrali pět typů výrazů označených jako altlexy a vyhledali jejich výskyty v celém korpusu PDTB – zjistili přitom, že dané výrazy byly označeny jako altlexy ve 47 případech, v 82 ovšem zůstaly bez anotace, tj. skutečný počet vybraných výrazů ve funkci altlexu je dvakrát, u některých až třikrát vyšší, než zachycuje stávající anotace v PDTB). Celkový počet altlexů je tedy podle autorů projektu ve skutečnosti mnohem vyšší,²⁶ než současná anotace ukazuje (ke srovnání výsledků anotace altlexů v angličtině v PDTB a sekundárních konektorů v češtině v PDT viz kapitola 6.2).

Částečná anotace altlexů se objevuje i v některých diskurzních projektech navazujících na Penn Discourse Treebank – např. v projektu Hindi Discourse Relation Bank pro hindštinu. Altlexy zde dosáhly 6,14 % z celkového počtu diskurzních vztahů (tj. v rámci všech typů vztahů anotovaných v daném korpusu, vyjádřených tedy např. explicitně i implicitně), nicméně anotace proběhla na poměrně malém vzorku dat (celkový počet anotovaných diskurzních vztahů v roce 2009 činil 602 výskytů, 6,14 % altlexů zde tedy znamená pouze 37 výskytů – srov. Oza a kol., 2009: 160). Vymezení diskurzních konektorů a jejich alternativních lexikálních vyjádření je zde plně převzata z projektu Penn Discourse Treebank.

Další významné diskurzní projekty – např. Potsdam Commentary Corpus (srov. Stede a Neumann, 2014), GECCo (Lapshinova-Koltunski a kol., 2012) aj. anotaci sekundárních konektorů (či altlexů) neobsahují. Diskurzní konektory většinou vymezují jako výrazy patřící mezi neohebné slovní druhy, a proto spojení typu *z tohoto důvodu* (které se v textu mohou

²⁶ Srov. „AltLex expressions are under-annotated in the PDTB because they were only annotated when annotators found it redundant to insert an implicit connective between adjacent sentences.“ (Prasadová a kol., 2014: 927)

objevit např. v množném čísle, v různých obměnách apod.) do anotace diskurzních vztahů nezahrnují.

Je tedy zřejmé, že přední zahraniční projekty zaměřené na zpracování diskurzních vztahů ve velkých korpusech systematickou anotací sekundárních konektorů neobsahují. Nejvíce prostoru věnuje těmto výrazům (nazývaným alternativní lexikální vyjádření diskurzních konektorů) projekt Penn Discourse Treebank, ovšem ani zde není jejich anotace zcela konzistentní a nevychází z hlubšího teoretického základu.

V této práci se tedy pokusíme o první systematické zachycení sekundárních konektorů ve velkých korpusových datech (a o jejich následnou analýzu).

4 Jazykový materiál

Výzkum sekundárních diskurzních konektorů jsme prováděli na datech Pražského závislostního korpusu, okrajově na datech Českého národního korpusu (SYN2013PUB). V další části práce jsme pak využili i data korpusu DIALOG (pro srovnání sekundárních konektorů v psané a mluvené češtině). Tyto tři korpusy v následující části práce blíže představíme.

4.1 Pražský závislostní korpus

Pražský závislostní korpus (PDT) obsahuje velké množství anotovaných jazykových dat. Jedná se o publicistické texty z různých novin a časopisů (Lidové noviny, Mladá fronta Dnes, Českomoravský profit, Vesmír).

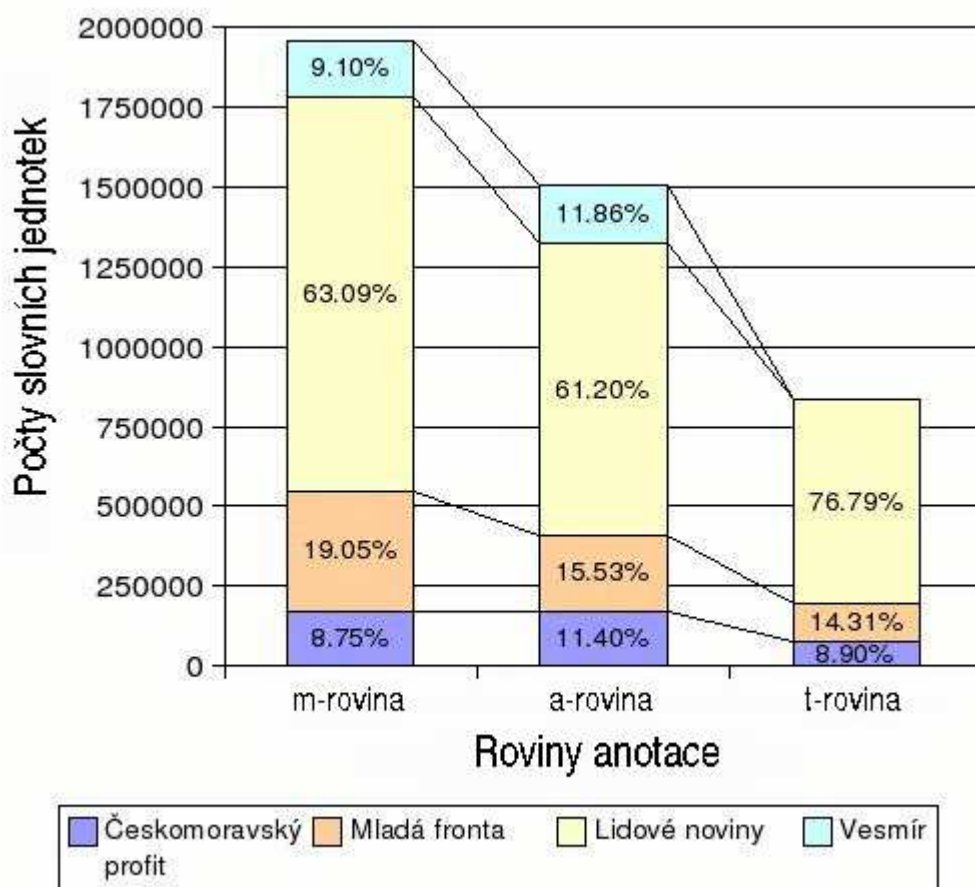
Pražský závislostní korpus zpracovává věty na několika jazykových úrovních zároveň – jedná se o rovinu morfologickou, analytickou (tj. povrchovou syntax) a tektogramatickou (tj. rovinu hloubkovou, syntakticko-sémantickou). Dále existuje ještě jedna neanotovaná rovina (tzv. slovní), ve které je text rozdělen do dokumentů a odstavců. PDT obsahuje na každé ze tří anotovaných rovin různé množství textů – srov. obrázek 3 zobrazující rozvržení počtu textů (resp. počtu slovních jednotek) z příslušných novin a časopisů na jednotlivých rovinách. Obrázek je převzatý z dokumentace k Pražskému závislostnímu korpusu 2.0 (Hajič a kol., 2006).

Pro účely naší práce budeme používat data z roviny tektogramatické, protože na této rovině obsahuje PDT jako jediný korpus češtiny (konkrétně ve verzi PDT 3.0) také anotaci nadvětných jevů, tj. koreference a asociační anafory a diskurzních vztahů. Zároveň obsahuje tektogramatická rovina i anotaci aktuálního členění větného.

Jednotlivé věty jsou v PDT zpracovávány ve stromových strukturách. Výhodou takovéto anotace je především to, že ve stromových strukturách můžeme vidět anotaci všech zmíněných rovin najednou.

Z tohoto hlediska je tedy korpus (mimo jiné) vhodný pro výzkum lingvistických jevů přesahujících několik jazykových plánů zároveň. V rámci sekundárních diskurzních konektorů tedy můžeme zkoumat například to, jaká je jejich syntaktická a sémantická role ve větě, jak korelují diskurzní vztahy s koreferencí a anaforou (srov. např. konektory obsahující

anaforické výrazy typu *kvůli tomu*, *důsledkem toho je* apod.). Díky vícevrstvé anotaci jazykových dat je možné podobné jevy v korpusu vyhledávat i automaticky.



Obrázek 3: Rozvržení počtu slovních jednotek na jednotlivých anotačních rovinách v Pražském závislostním korpusu (viz Hajič a kol., 2006)

Tektogramatická rovina Pražského závislostního korpusu obsahuje celkem 3 165 dokumentů, tj. 49 431 vět a 833 195 slovních jednotek. Data jsou přitom rozdělena do několika množin – na data trénovací (*train*), vývojová testovací (*dtest*) a evaluační testovací (*etest*). Trénovací data tvoří přibližně 80 % celkového obsahu korpusu (ta mohou uživatelé volně využívat), vývojová testovací 10 % a evaluační 10 % – srov. tabulka 1 zobrazující konkrétní počet dokumentů, vět a slovních jednotek v jednotlivých množinách tektogramatické roviny Pražského závislostního korpusu (převzata ze zadávací dokumentace k PDT 2.0 – srov. Hajič a kol., 2006).

tamw	train	dtest	etest	total
Umístění na CD-ROM v data /full/	tamw/train-1/ ... tamw/train-8/	tamw/dtest/	tamw/etest/	tamw/*/
# dokumentů	2 533 (80,0 %)	316 (10,0 %)	316 (10,0 %)	3 165 (100,0 %)
# vět	38 727 (78,3 %)	5 228 (10,6 %)	5 476 (11,1 %)	49 431 (100,0 %)
# slovních jednotek	652 544 (78,3 %)	87 988 (10,6 %)	92 663 (11,1 %)	833 195 (100,0 %)

Tabulka 1: Množství dat v jednotlivých vrstvách Pražského závislostního korpusu (viz Hajič a kol., 2006)

Anotace jazykových dat Pražského závislostního korpusu je neustále rozšiřována a obohacována, průběžně tedy vycházejí nové, aktualizované verze korpusu. Pro účely naší práce byla použita data PDT 3.0 (Bejček a kol., 2013) obohacená o dosud nepublikovanou anotaci sekundárních konektorů, zároveň v nich během anotace sekundárních konektorů byla mírně upravena anotace primárních konektorů (tyto opravy rovněž zatím nebyly publikovány).

4.1.1 Anotace primárních konektorů v Pražském závislostním korpusu

První fáze anotace diskurzních vztahů v češtině proběhla na datech Pražského závislostního korpusu 2.5 (Bejček a kol., 2011). Anotace byla dokončena v roce 2012 a byla zpřístupněna jako Pražský diskurzní korpus 1.0 (srov. Poláková a kol., 2012b). Anotace byla přitom omezena na diskurzní vztahy (vnitrovětné i mezivětné) vyjádřené explicitními primárními konektory, tj. výrazy patřícími mezi vybrané slovní druhy (především spojky typu *ale*, *a*, *proto* a strukturující částice typu *také*, *totiž*, *vždyt'* atd.). Podmínkou pro anotaci diskurzního vztahu byla zároveň syntaktická povaha obou argumentů, tj. diskurzní vztah byl značen pouze tehdy, pokud byl mezi dvěma slovesnými argumenty – srov. příklad 82 z PDT s konektorem *a tak* signalizujícím vnitrovětný vztah příčiny a důsledku a příklad 83 z PDT s konektorem *navíc* vyjadřujícím mezivětný vztah gradace.

(82) *Tyto potraviny a bavlna se v České republice nepěstují, a tak jejich dovoz naše zemědělce neohrozí.*

(83) *Elektrina v bytech funguje pouze několik hodin denně.*

Poprvé od počátku války minulý týden srbská strana navíc zablokovala přívod plynu.

Příklad 83 je zároveň zachycený na obrázku 4. Obrázek ukazuje dané argumenty ve stromové struktuře PDT (argument 1 je *Elektrina v bytech funguje pouze několik hodin denně*; argument 2 je *Poprvé od počátku války minulý týden srbská strana zablokovala přívod plynu*; vztah gradace mezi argumentem 1 a argumentem 2 je vyjádřen explicitně – primárním konektorem *navíc*).

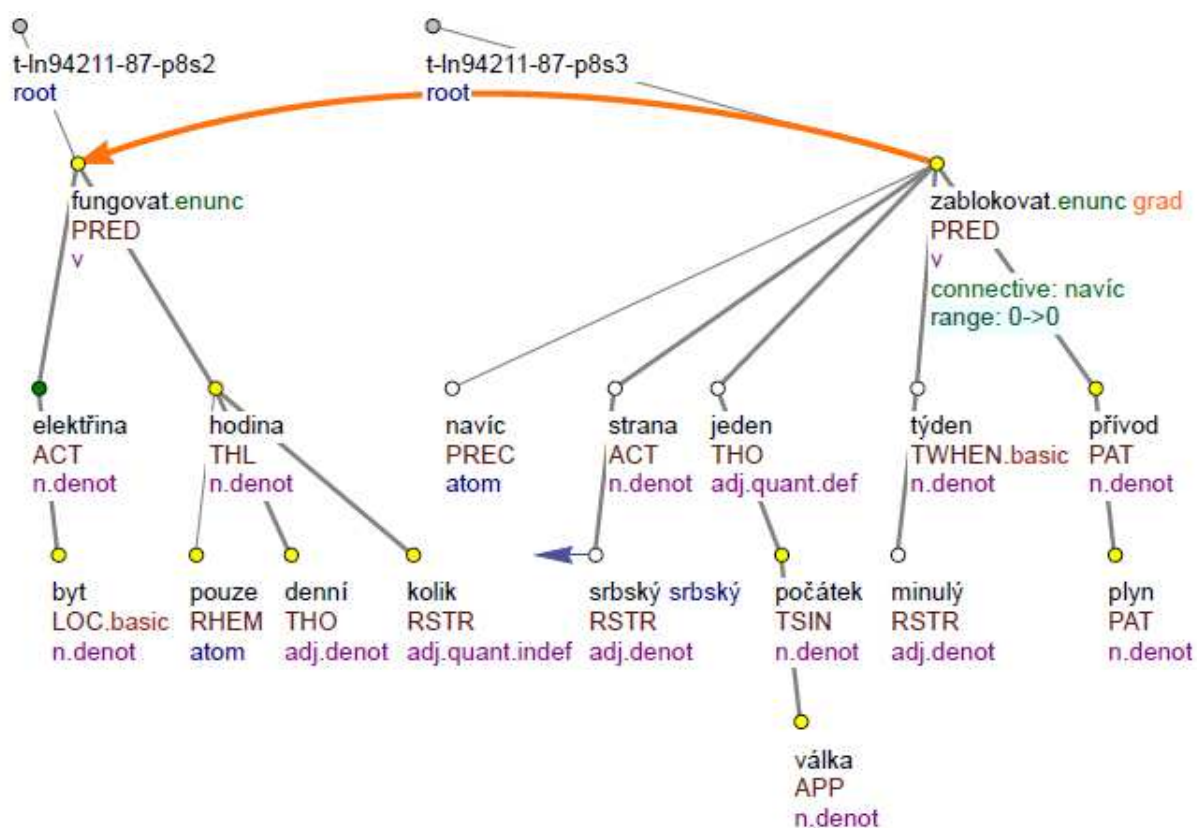
Diskurzní vztahy jsou v PDT značeny oranžovými šipkami, které vedou mezi dvěma řídicími slovesy (v tomto případě *fungovat* a *zablokovat*). Daná slovesa reprezentují jako hierarchicky nejvyšší uzly celou příslušnou větu, tj. celé argumenty. Ačkoli tedy technicky vede diskurzní vztah (znázorněný šipkou) mezi dvěma slovesy (uzly), ve skutečnosti tvoří diskurzní argument celá věta (tj. do argumentu patří všechny uzly daného stromu závislé na příslušném řídicím slovese).

Značka „connective: navíc“ znamená, že je daný diskurzní vztah vyjádřen primárním konektorem *navíc*. Značka „range: 0→0“ zachycuje rozsah diskurzních argumentů. V tomto případě je diskurzní vztah pouze mezi zobrazenými dvěma větami (tj. mezi větou, od které

vede diskurzí šípka, a větou, ke které šípka směřuje). Diskurzí vztahy ovšem mohou být i mezi většími úseky textu, např. mezi větou a odstavcem nebo mezi dvěma odstavci.²⁷

Diskurzí argumenty jsou v PDT definovány podle Asherovy koncepce abstraktních objektů (1993) a principu minimality (Prasadová a kol., 2007), tj. do diskurzího argumentu patří pouze výpovědi (nebo jejich části), které nesou informace nezbytně nutné pro interpretaci příslušného diskurzího vztahu (viz kapitola 3.4). V reálných textech mohou být diskurzí argumenty realizovány například větou, souvětím, samostatnou vedlejší větou, posloupností několika vět i celým odstavcem.

Diskurzí vztahy jsou v PDT značeny mezi argumenty, které jsou v textu bezprostředně vedle sebe (viz příklad 83), ale i mezi argumenty, které spolu přímo nesousedí, tj. mezi dva argumenty může být „vložen“ text (věta či posloupnost více vět), který není součástí ani jednoho z argumentů.



Obrázek 4: Ukázka anotace primárních konektorů v Pražském závislostním korpusu, viz příklad 83: *Elektřina v bytech funguje pouze několik hodin denně. Poprvé od počátku války minulý týden srbská strana navíc zablokovala přívod plynu.*

²⁷ Například značka „range: 2→6“ by znamenala, že do argumentu, od kterého vede diskurzí vztah, patří celkem tři věty (věta, kde diskurzí šípka začíná, plus dvě věty po ní následující), a do argumentu, ke kterému diskurzí vztah vede, patří sedm vět (věta, kde diskurzí šípka končí, plus šest vět po ní následujících).

V průběhu anotace neměli anotátoři k dispozici předem sestavený seznam primárních konektorů, podle kterého by určovali, který výraz mezi primárními konektory patří a který ne (tímto způsobem byla anotace diskurzivních vztahů prováděna např. na datech pensylvánského korpusu Penn Discourse Treebank – srov. Prasadová a kol., 2008; tj. výraz, který není na sestaveném seznamu konektorů, není vyhodnocen jako konektor). Anotátoři byli na začátku seznámeni s obecnou definicí primárních konektorů a sami je posléze v textech hledali a rozpoznávali.

Prvotní anotace diskurzivních vztahů v PDT, jak již bylo zmíněno výše, zahrnuje zpracování mezivětných i vnitrovětných vztahů. Mezivětné vztahy byly anotovány manuálně, vnitrovětné částečně automaticky (resp. automatická anotace vnitrovětných vztahů byla následně zkontrolována a manuálně upravena anotátory, viz Jínová a kol., 2012).

4.2 Korpus DIALOG

Korpus DIALOG je multimedialní korpus mluvené češtiny. Obsahuje jazykové projevy dialogického typu – jedná se o nahrávky a přepisy diskusních pořadů vysílaných v televizi. Korpus je určen především k výzkumu současné mluvené češtiny. K vyhledaným přepsaným pasážím textů je zároveň možné přehrát si odpovídající videozáznam, k dispozici je také zobrazení akustické křivky F0.

Pro účely naší práce jsme použili nejnovější verzi tohoto korpusu, tj. DIALOG 1.1 (2012). Celková velikost korpusu je 932 373 textových slov (tj. 1 264 902 textových slov a technických značek). Korpus zachycuje 758 mluvčích (245 z nich vysloví více než 1 000 slov) a obsahuje 150 nahrávek a přepisů.

DIALOG 1.1 je tedy svojí velikostí srovnatelný s Pražským závislostním korpusem 3.0 (932 373 textových slov v DIALOGu vs. 833 195 v PDT). Značná část pořadů, ze kterých nahrávky pocházejí, je zároveň z oblasti publicistiky (srov. např. Otázky Václava Moravce, Aréna, 7 čili sedm dní, 21 apod.) – texty obou korpusů jsou tedy podobné i stylově.

Korpus DIALOG 1.1 je morfologicky anotovaný a lemmatizovaný (anotace proběhla automaticky). Anotace diskurzivních vztahů není v korpusu obsažena.

Data z korpusu DIALOG 1.1 jsme použili pro jednu dílčí studii v rámci našeho výzkumu, a to pro porovnání vybraných sekundárních konektorů v psané a mluvené češtině (viz kapitola 6.8).

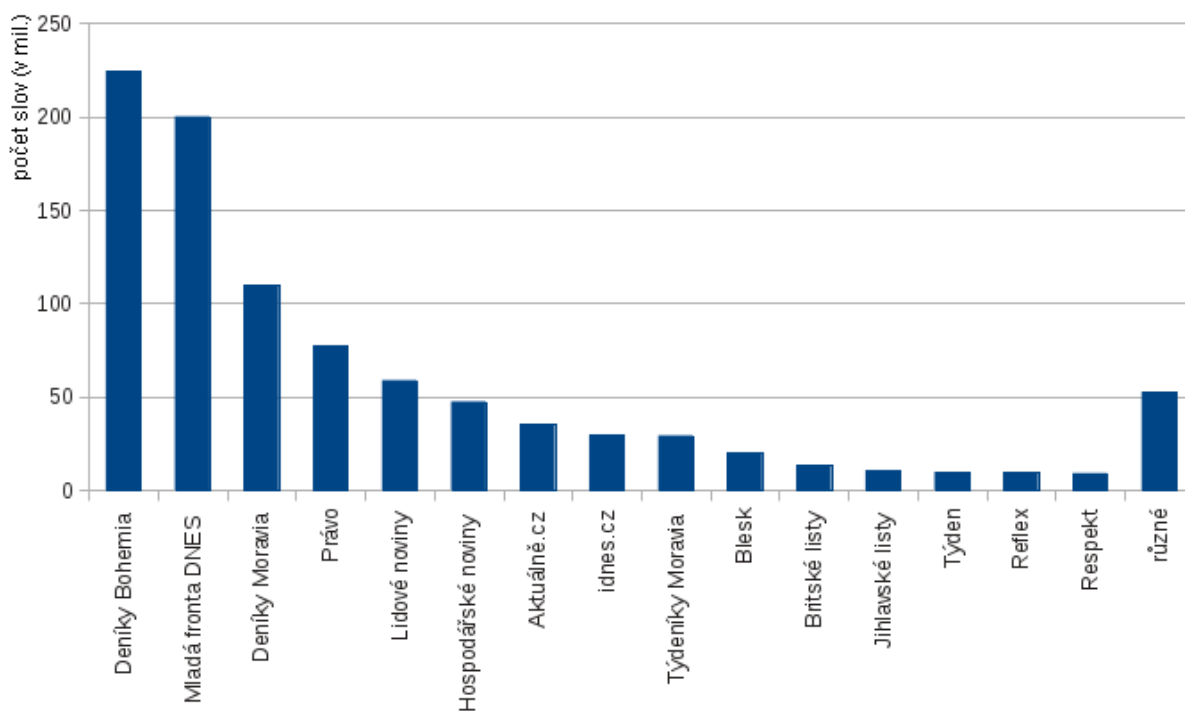
4.3 Český národní korpus

Český národní korpus (ČNK) je akademický projekt, jehož hlavním cílem je vybudování rozsáhlého souboru počítačově uložených a zpracovaných textů především psané češtiny. ČNK slouží primárně k jazykovému výzkumu. Pomocí vyhledávacího programu KonText je možné v korpusu vyhledávat klíčová slova, příp. spojení slov v kontextu.

V rámci projektu ČNK byla vytvořena řada korpusů, jedná se především o korpusy psané češtiny (např. SYNT2010, SYN2013PUB), korpusy mluvené češtiny (např. ORAL2013), diachronní korpusy (např. DIAKORP), cizojazyčné korpusy (např. Aranea), paralelní korpus (InterCorp) aj.

V naší práci jsme využili data korpusu SYN2013PUB pro analýzu frekvence vybraných sekundárních konektorů a volných konektivních frází.

Korpus SYN2013PUB obsahuje 935 milionů slov z výhradně publicistických textů, složení korpusu podle titulů viz obrázek 5 (srov. <https://ucnk.ff.cuni.cz/syn2013pub.php>).



Obrázek 5: Složení korpusu SYN2013PUB podle titulů (viz <https://ucnk.ff.cuni.cz/syn2013pub.php>)

5 Metody a postup práce

Vzhledem k tomu, že žádný český korpus neobsahuje systematické zpracování sekundárních konektorů, museli jsme příslušnou anotaci nejprve sami provést (tj. stanovit anotační principy a poté je aplikovat na konkrétní data), abychom získali jazykový materiál pro vlastní výzkum.

Ten jsme prováděli primárně na datech Pražského závislostního korpusu (PDT). PDT je pro naše účely vhodný především proto, že jako jediný korpus češtiny obsahuje anotaci primárních konektorů. Jeho předností je tedy zejména to, že po zpracování sekundárních konektorů bude možné obě anotace vzájemně porovnat (a srovnat tak primární a sekundární konektory z hlediska jejich výskytu a chování v reálných psaných textech).

Při vytváření anotačních principů jsme vycházeli z anotace primárních konektorů. Při práci s konkrétními texty jsme ale zjistili, že dané principy nemůžeme aplikovat bez výhrady. Již během prvotní, zkušební anotace sekundárních konektorů se ukazovaly obecné jazykové odlišnosti mezi oběma skupinami konektorů. Tyto odlišnosti jsme pak zohlednili při vytváření anotačních principů, které byly využity pro anotace konečné.

Po anotaci zkušební vzorku dat (při které jsme řešili různé teoretické problémy a otázky týkající se např. rozsahu sekundárních konektorů, tj. zda je konektor např. celé spojení *to je důvod, proč* nebo jenom část *to je důvod* atd., stanovení hranic mezi konektory a výrazy, které už mezi konektory řadit nemůžeme apod.) byly sestaveny anotační principy pro sekundární konektory, které v následující části práce ve stručnosti představíme. Některé teoretické otázky a návrh jejich řešení jsme popsali již v kapitole 3.7. Popis anotace primárních konektorů v PDT, ze které jsme vycházeli při sestavování anotačních principů pro sekundární konektory, byl představen v kapitole 4.1.1.

Jednotlivé dílčí metody použité během našeho dalšího výzkumu (tj. např. jak jsme vyhodnocovali dokončenou anotaci, jak a s jakými výsledky jsme testovali univerzálnost sekundárních konektorů apod.) budou pro lepší přehlednost představeny v jednotlivých kapitolách v rámci popisu výsledků práce.

5.1 Anotace sekundárních konektorů v Pražském závislostním korpusu

Práce na anotaci sekundárních konektorů začaly v roce 2012. Podnět k jejich zpracování vzešel z prvotní anotace primárních konektorů, v průběhu které anotátoři dostali pokyn (kromě anotace primárních konektorů) označovat v poznámkách i takové jazykové prostředky, které podle nich v textu slouží jako sekundární konektory (v té době nazývané altlexy – alternativní vyjádření konektorů – podle terminologie převzaté od Prasadové a kol., 2008, 2010).

V rámci prvotní diskurzí anotace tedy celkově nešlo o systematické zpracování sekundárních konektorů, ale především o sběr jazykového materiálu k dalšímu výzkumu (dané poznámky byly pracovní, určené výhradně pro budoucí výzkum). Celkem se v rámci této anotace objevilo 306 anotátorských poznámek upozorňujících v textu na výskyt sekundárního konektoru, jejich zpracování přitom nebylo zcela konzistentní (např. určitý typ sekundárního konektoru nebyl zachycen ve všech svých výskytech apod., srov. Rysová, 2012, 2013b).

Přesto byly tyto anotátorské poznámky velmi přínosné – na jejich základě bylo možné sestavit prvotní seznam sekundárních konektorů (resp. spíše jejich jádrových slov) vyskytujících se v Pražském závislostním korpusu – srov. tabulka 2.

Prvotní seznam sekundárních konektorů či jejich jádrových slov v PDT	
<i>Argumentovat</i>	<i>Přeloženo</i>
<i>Díky</i>	<i>Přes</i>
<i>Dlužno dodat</i>	<i>Přesněji</i>
<i>Do třetice</i>	<i>Přičíst</i>
<i>Dodat/dodávat</i>	<i>Příčina: být příčinou</i>
<i>Doplnit/doplňovat</i>	<i>Příklad: být příkladem; příklad pochází z; jako příklad slouží; uvést jako příklad</i>
<i>Důsledek: důsledkem je; to je důsledek</i>	<i>Případ: být opačným případem; neplatí to v případě; v (tomto) případě</i>
<i>Důvod: z tohoto důvodu; z těchto důvodů; z uvedených důvodů</i>	<i>Přispívat</i>
<i>Jak je vidět</i>	<i>Přístupovat</i>
<i>Jako poslední</i>	<i>Rozdíl: být rozdílem</i>
<i>Kontrastovat</i>	<i>Rozpor: to je v rozporu</i>
<i>Kvůli</i>	<i>Řečeno: stručně řečeno; jinak řečeno; zjednodušeně řečeno; jednoduše řečeno</i>
<i>Na prvním místě</i>	<i>Souviset</i>
<i>Na rozdíl od</i>	<i>Stejným dechem</i>
<i>Na základě</i>	<i>Tím spíš</i>
<i>Následovat</i>	<i>Týkat se</i>
<i>Navzdory</i>	<i>Účel: za tím/tímto účelem; pro tento účel</i>
<i>Ne dosti na tom</i>	<i>Upřesnit</i>
<i>Nehledě na</i>	<i>Směr: v tomto směru</i>
<i>Nemluvě o</i>	<i>V (této) souvislosti</i>
<i>O to více</i>	<i>Ve skutečnosti</i>
<i>Odhadnout</i>	<i>Vést</i>
<i>Odůvodnění: s odůvodněním, že</i>	<i>Vinou</i>
<i>Odůvodnit/odůvodňovat</i>	<i>Výjimka: být výjimkou; tvořit výjimku</i>
<i>Plynout</i>	<i>Vyplývat</i>
<i>Podmínka: být podmínkou; s podmínkou, že</i>	<i>Výsledek: být výsledkem</i>
<i>Podobně</i>	<i>Vysvětlit</i>
<i>Pokračovat</i>	<i>Vzhledem k</i>
<i>Pravděpodobnější</i>	<i>Závěr: závěr je; závěr zní</i>
<i>Právě tak</i>	<i>Zdůvodnit/zdůvodňovat</i>
<i>Předcházet</i>	<i>Znamenat</i>
<i>Případně</i>	<i>Způsobit</i>

Tabulka 2: Prvotní seznam sekundárních konektorů, resp. jejich jádrových slov v Pražském závislostním korpusu

Tabulka zachycuje především jádrová slova jednotlivých sekundárních konektorů (srov. např. podstatná jména *důvod*, *příčina*, *výjimka*, slovesa *způsobit*, *předcházet*, sekundární předložky *vinou*, *kvůli* apod.). U některých z nich jsou za dvojtečkou uvedeny konkrétní realizace v PDT (označené prvními anotátory diskurzních vztahů) – srov. např. *Výjimka: být výjimkou; tvořit výjimku*. Ustálené fráze jsou uvedeny celé, resp. v takové podobě, v jaké mají funkci sekundárních konektorů (srov. např. *přeloženo*). Jsme si zároveň vědomi toho, že daný seznam sekundárních konektorů není konečný – vzhledem k rozmanitosti sekundárních konektorů můžeme předpokládat, že se v reálných textech budou objevovat další struktury s konektivní funkcí, které v naší anotaci nebyly zachyceny. Příslušný seznam sekundárních konektorů tak bude možné dále doplňovat a rozšiřovat.

Na základě uvedeného seznamu jsme v PDT automaticky vyhledávali jednotlivá jádrová slova (příp. celé ustálené fráze) – např. všechny výskyty slova *důvod*. Pro tato vyhledávání byla použita celá data PDT, tj. přibližně 50 000 vět. Nalezené výskyty (celkem se jednalo o téměř 6 000 výskytů) byly posléze zpracovávány ručně anotátory. Jejich úkolem bylo rozhodnout, zda se v nalezených případech jedná o výskyt ve funkci sekundárního konektoru (pak byl daný výskyt anotátorem označen jako sekundární konektor) nebo o výskyt struktury sice shodné s určitým sekundárním konektorem, ale v nekonektivní funkci (pak daný výskyt jako sekundární konektor anotován nebyl).

Anotátoři zároveň určovali, jaký typ vztahu daný sekundární konektor vyjadřuje (k dispozici přitom měli seznam významových vztahů použitý při předchozí anotaci primárních konektorů). Pokud anotátor určil, že se jedná o vztah, který se na daném seznamu nenachází, zapsal nový typ vztahu do anotátorské poznámky.

Anotátoři zároveň označovali, mezi kterými a jak rozsáhlými argumenty (úseky textu) vztah vyjádřený sekundárním konektorem je. Anotátoři přitom nebyli omezováni tím, jakou syntaktickou povahu by měl argument mít (na rozdíl od anotace primárních konektorů, při níž se označovaly vztahy pouze mezi slovesnými argumenty).

Také možný rozsah argumentů nebyl předem nijak omezován, stejně jako to, v jaké vzájemné vzdálenosti se argumenty v textu mohou nacházet (např. bylo teoreticky možné, aby byl označen významový vztah mezi prvním a posledním odstavcem textu, aniž by do tohoto vztahu zasahovaly odstavce, které se nacházely mezi těmito dvěma krajními body).

Kromě sekundárních konektorů na uvedeném seznamu byly do anotace zahrnuty také výrazy, které byly jako konektory označeny již v první fázi anotace diskurzních vztahů v PDT, tj. v době, kdy ještě nebyly důsledně rozlišovány primární a sekundární konektory. Některé z těchto výrazů tedy byly nově vyhodnoceny jako konektory sekundární a byly v anotaci

manuálně přeznačeny (tj. byla zrevidována prvotní anotace diskurzních vztahů a upravena podle našeho dělení konektorů na primární a sekundární). Jednalo se především o některá příslovce typu *konkrétně* a předložkové fráze typu *k tomu, s čímž* apod. Kompletní abecední a frekvenční seznam primárních a sekundárních konektorů zachycených v současné (dosud nepublikované) anotaci v PDT je uveden v příloze 1.

Manuální anotaci sekundárních konektorů prováděli tři různí anotátoři (všichni se zároveň podíleli na předchozí anotaci primárních konektorů, měli tedy s anotací diskurzu již zkušenosti²⁸). Po dokončení těchto prací byla celá anotace znovu zkontrolována²⁹. Ukázka textu z Pražského závislostního korpusu s anotací významových diskurzních vztahů vyjádřených konektory je uvedena v příloze 2.

²⁸ Anotace prováděly Pavlína Jínová, Veronika Pavlíková a Magdaléna Rysová. Technickou podporu anotací zajišťoval Jiří Mírovský.

²⁹ Kontroly anotací prováděla Magdaléna Rysová.

5.1.1 Odlišnosti od anotace primárních konektorů a další teoretické otázky a problémy

Při anotaci prvotního vzorku dat (jejímž účelem bylo prozkoumat chování sekundárních konektorů v reálných textech a následně sestavit obecné anotační principy) vyvstávaly různé teoretické otázky či problémy. Při zpracovávání těchto konektivních struktur jsme zjistili, že na ně nelze aplikovat stejné anotační zásady, které byly stanoveny pro primární konektory. Víceslovné konektivní struktury se v mnoha případech chovají odlišně než „klasické“ (tj. primární) konektory, a proto nebylo možné pouze převzít již stanovené anotační principy, ale bylo nezbytné některé z nich modifikovat, případně vytvořit zcela nové.

5.1.1.1 Anotace nominálních argumentů

Pro anotaci primárních konektorů bylo například jasně daným pravidlem, že se zachycují pouze diskurzní vztahy mezi dvěma slovesnými argumenty (jak již bylo zmíněno výše). Při anotaci sekundárních konektorů se ovšem objevovaly příklady typu *Důsledkem je stárnutí obyvatelstva* apod. (srov. kapitola 3.4.2). Vzhledem k tomu, že věty *Důsledkem je stárnutí obyvatelstva* a *Důsledkem je, že obyvatelstvo stárne* se liší pouze syntaktickou povahou diskurzního argumentu (tj. *stárnutí obyvatelstva* vs. *obyvatelstvo stárne*), rozhodli jsme se do anotace sekundárních konektorů zahrnovat i tyto případy, tj. neomezovat diskurzní argumenty jejich povrchovou syntaktickou podobou.

5.1.1.2 Rozsah sekundárních konektorů

Další teoretická otázka se týkala rozsahu sekundárního konektoru. Jak již bylo řečeno v teoretické části práce (srov. především kapitola 3.7.3.6), sekundární konektory jsou většinou víceslovné struktury. V některých případech ovšem není zcela jednoznačné, kde je jejich hranice, resp. jaké složky ke konektoru patří a kde je hranice mezi konektorem a argumentem – srov. příklad 84:

(84) *Divize kamenina akciové společnosti Calofrig Borovany jako první v oboru získala certifikát řízení jakosti podle ISO 9000.*

To byl hlavní důvod, proč jsme položili několik otázek Janu Bakulemu, řediteli této divize, v jejímž výrobním programu jsou především kanalizační roury různých průměrů.

Mezi danými větami je diskurzivní vztah příčiny a důsledku vyjádřený sekundárním konektorem. Otázkou ovšem je, jakými složkami je konektor tvořen. Základní strukturou konektoru (která se v PDT objevuje v různých realizacích) je *to je/byl důvod*. Otázkou v tomto případě je, zda ke konektoru přiřadit i přívlastek *hlavní* a vztahné zájmené příslovce *proč*.

U jednotlivých rozvití (dále např. *jednoduchý, složitý, filozofický, jasný* apod. – viz kapitola 3.7.3.6) zvažujeme především jejich lexikální povahu (obecnější rozvití se hodí do více kontextů než rozvití konkrétnější, lexikálně naplněnější, tj. mají vyšší pravděpodobnost stát se součástí univerzálního konektoru). Jako pomocné (nikoli absolutní) kritérium v těchto případech slouží náhrada příslušného sekundárního konektoru konektorem primárním (bez ztráty významného významového rysu daného textového úseku).

Do příkladu 84 můžeme dosadit modifikovaný primární konektor *hlavně proto*, lexikální význam tedy zůstane zachován. Z tohoto důvodu jsme se rozhodli přívlastek *hlavní* (a obecně obdobná rozvití) chápat jako součást sekundárního konektoru.

Nicméně vzhledem k rozmanitosti struktur sekundárních konektorů zůstává jejich rozsah v některých případech předmětem diskuze.

5.1.1.3 Hranice mezi sekundárním konektorem a argumentem

Při určování hranice mezi konektorem a argumentem (v příkladě 84 se jedná o přiřazení zájmeného příslovce *proč*) uplatňujeme nahraditelnost primárním konektorem jako základní kritérium. Tj. testovaný sekundární konektor nahradíme primárním, všechny ostatní složky, které v takto nově vzniklé větě zůstanou (a nejsou primárním konektorem), jsou součástí argumentu – srov. příklad 85:

(85) *Divize kamenina akciové společnosti Calofrig Borovany jako první v oboru získala certifikát řízení jakosti podle ISO 9000.*

Hlavně proto jsme položili několik otázek Janu Bakulemu, řediteli této divize, v jejímž výrobním programu jsou především kanalizační roury různých průměrů.

V daném příkladě je hranice mezi konektorem a argumentem zcela zřetelná – jako modifikovaný primární konektor můžeme určit *hlavně proto*, jako diskurzní argument *jsme položili několik otázek...* Primární konektor *hlavně proto* tedy v tomto případě nahrazuje celou strukturu *to byl hlavní důvod, proč*, tj. zájmenné příslovce *proč* z příkladu 84 řadíme k sekundárnímu konektoru, nikoli k argumentu.

Zároveň však existují sekundární konektory, které nemají mezi primárními konektory přímý ekvivalent (např. *v tomto ohledu, na závěr* – viz kapitola 6.6).³⁰

Jsme si vědomi toho, že v některých případech může být určování hranice mezi konektorem a argumentem diskutabilní.

5.1.1.4 Dvouúrovňová anotace sekundárních konektorů

Při anotaci zkušebního vzorku dat se ukázalo, že sekundární konektory vykazují vysokou míru variability, tj. jeden typ sekundárního konektoru se objevoval v různých realizacích (srov. např. *to / tato skutečnost je / byl / byla důvod, proč / kvůli kterému* apod.). Z toho je patrné, že sekundární konektory nejsou tak ustálené (či ustrnulé) jako konektory primární a vykazují větší míru heterogenity.

Variabilita je přitom daná především slovnědruhovou příslušností jednotlivých složek konektoru – slovesa se vyskytují například v různých časech (např. *příčinou je* vs. *příčinou bylo*), podstatná jména v různých pádech (*důvod je* vs. *důvodem je*) apod. Proto bylo sestavení jednotných anotačních principů složitější než pro konektory primární.

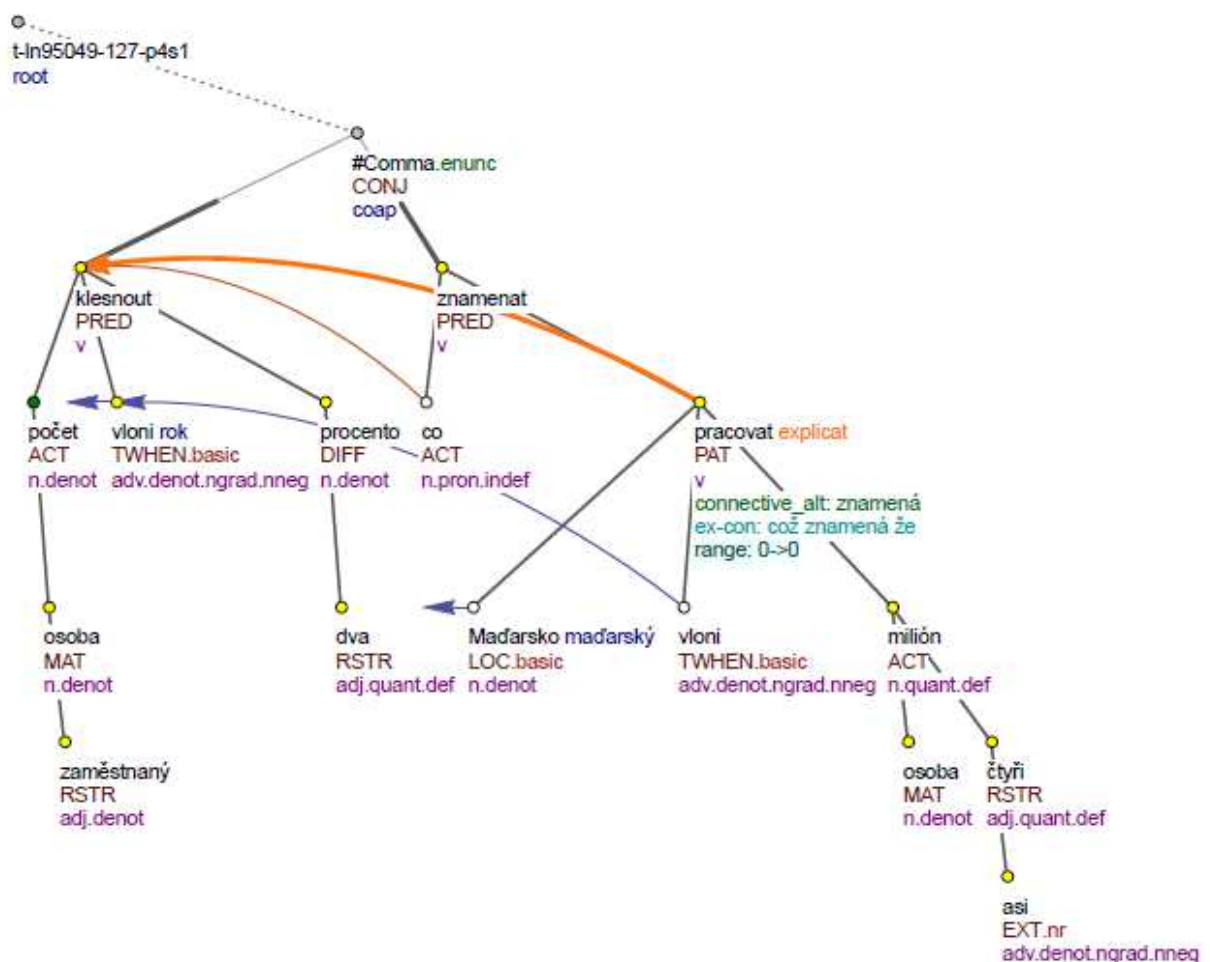
Vzhledem ke zmíněné variabilitě sekundárních konektorů a vzhledem k tomu, že většina z nich obsahuje jakési jádro (často vyjadřující sémantický typ diskurzního vztahu – srov. např. *důvod, příčina, kontrastovat, na rozdíl od, znamenat* apod.), rozhodli jsme se pro dvouvrstvou anotaci těchto konektorů. U každého konektoru proto anotátor jednak označil jeho jádro a jednak zapsal také celou konektivní strukturu, a to v takové podobě, jak se

³⁰ I z tohoto důvodu nepřejímáme označení alternativní lexikální vyjádření konektorů (zkráceně altlexy, viz Prasadová a kol., 2010) – ne vždy je sekundární konektor alternativou k primárnímu.

vyskytovala v textu – srov. obrázek 6 z PDT. Obrázek zachycuje stromovou strukturu následujícího souvětí:

- (86) *Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.*

V daném příkladě je mezi diskurzními argumenty (argumentem 1 je *počet zaměstnaných osob klesl loni o dvě procenta*; argumentem 2 je *v Maďarsku loni pracovaly asi čtyři milióny osob*) diskurzní vztah explikace (vysvětlení) – tj. obsah druhého argumentu vysvětluje obsah prvního argumentu. Diskurzní vztah je explicitně vyjádřen sekundárním konektorem *což znamená, že*. Jádrem konektoru je v tomto případě sloveso *znamenat* (v anotaci je dané jádro zaznamenáno vždy ve tvaru, v kterém se objevilo v textu, tj. na obrázku 6 ve tvaru *znamená*).



Obrázek 6: Ukázka dvouvrstvé anotace sekundárních konektorů v Pražském závislostním korpusu na souvětí: *Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.*

Na obrázku je diskurzí vztah explikace znázorněný oranžovou šipkou mezi uzly *klesnout* a *pracovat* reprezentujícími celé argumenty (ostatní šipky vyjadřují koreferenční vztahy, které nejsou předmětem našeho zkoumání, a proto je nebudeme blíže komentovat). Jako konektor je zde anotované spojení *což znamená, že*. Ve vyšším patře je přitom jádro konektoru – *znamená* (označené jako „connective_alt“, tj. connective_altlex – tuto pracovní zkratku používáme pro jádra sekundárních konektorů, které jsou v zahraniční literatuře nazývané altlexy – srov. Prasadová a kol., 2008, 2010). V nižším patře je potom zachycena celá struktura sekundárního konektoru – *což znamená, že* (pro anotace opatřena zkratkou „ex-con“, tj. extensive connective, v překladu „rozšířený konektor“).

Dvouvrstvá anotace sekundárních konektorů nám posléze umožní (automaticky) vyhledávat jednotlivé konektory podle jejich základů či jader, tj. bude možné např. vyhledat všechny sekundární konektory s jádrem *důvod, výjimka* apod. Budeme tak moci např. vysledovat, v jakých konektivních strukturách či variantách se dané jádro objevuje (srov. např. *příčina: příčinou je, vidět příčinu v, hledat příčinu v* apod.), která varianta převažuje (a začíná se tak ustalovat jako nejčastěji používaný sekundární konektor v rámci příslušných struktur) atd.

5.1.2 Anotační principy pro zpracování sekundárních konektorů (a volných konektivních frází) v PDT

V následující části práce představíme technické informace a pokyny k anotaci sekundárních konektorů v PDT. Cílem této části práce je tedy mimo jiné ukázat, jak je možné zpracovávat diskurzní konektory v korpusu prakticky, po technické stránce³¹. Tuto kapitolu zde uvádíme především proto, že by mohla být inspirací pro zpracování diskurzních vztahů v dalších korpusech.

5.1.2.1 Diskurzní vztah mezi verbálními argumenty

Pokud byl diskurzní vztah vyjádřený sekundárním konektorem mezi dvěma slovesnými argumenty, rozhodli jsme se pro (technicky) stejnou anotaci, která proběhla na primárních konektorech (pouze s rozdílem dvouvrstvé anotace u sekundárních konektorů) – viz obrázek 6. Diskurzní vztah je tedy v těchto případech značen oranžovou šipkou vedoucí mezi danými argumenty. U ní je zároveň uveden také sémantický typ diskurzního vztahu.

Účelem takto zvolené anotace bylo, aby zpracování primárních a sekundárních konektorů (vyjadřujících vztah mezi dvěma slovesnými argumenty) v korpusu bylo pokud možno konzistentní – primární a sekundární konektory jsou v těchto případech většinou navzájem zaměnitelné (při zachování daného významu).

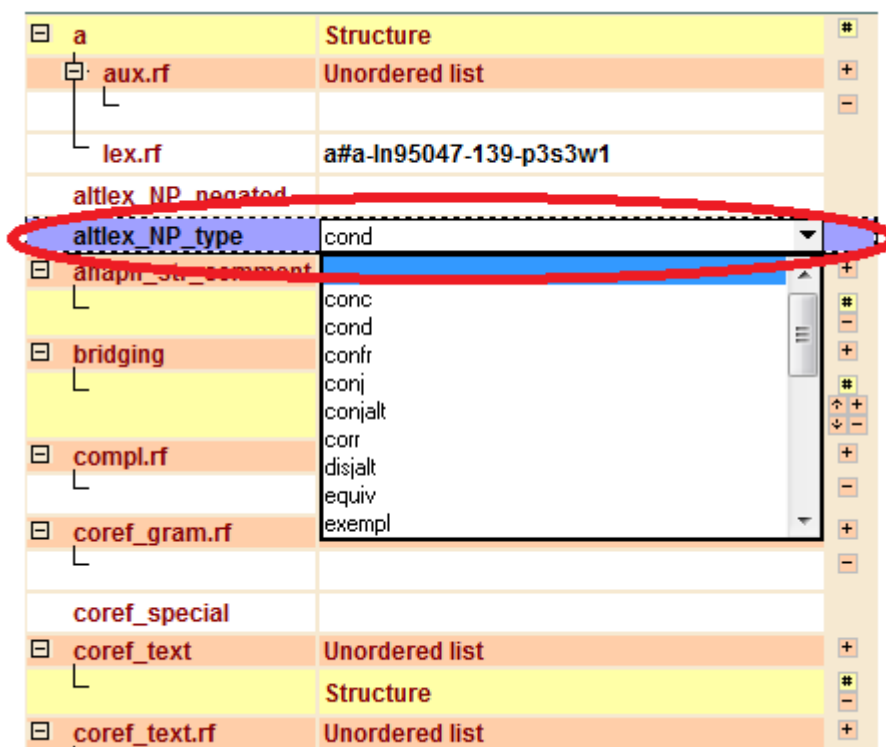
5.1.2.2 Diskurzní vztah mezi argumenty, z nichž je alespoň jeden nominální

Diskurzní vztahy vyjádřené sekundárními konektory mezi dvěma diskurzními argumenty, z nichž je alespoň jeden vyjádřen nominální frází, jsme se rozhodli technicky odlišit od předchozích vztahů (tj. vztahů mezi dvěma slovesnými argumenty). Toto rozlišení nám posléze usnadní automaticky zpracovat celkové výsledky a obě skupiny porovnat po stránce jazykové. Můžeme tak například zjistit, které sekundární konektory upřednostňují nominalizaci diskurzních argumentů, které ji alespoň umožňují (i když u nich převažují argumenty slovesné) apod.

Po technické stránce byla anotace těchto diskurzních vztahů zachycena následujícím způsobem. K uzlům na tektogramatické rovině byl přidán nový atribut (s označením „altlex_NP_type“) umožňující označit příslušný uzel jako jádro sekundárního konektoru

³¹ Za vytvoření technického prostředí pro anotaci sekundárních konektorů a volných konektivních frází vděčíme Jiřímu Mírovskému.

signalizujícího diskurzí vztah alespoň s jedním nominálním argumentem. Pokud byl daný uzel takto označen, musel být zároveň vyplněn sémantický typ daného vztahu (např. opozice, podmínka apod.). Daný vztah tedy v tomto případě není značen šipkou (jako tomu bylo u vztahů mezi slovesnými argumenty), ale přidáním příslušného atributu u jádra sekundárního konektoru – srov. obrázek 7:



Obrázek 7: Atribut „altlex_NP_type“ pro diskurzí vztahy s nominálním argumentem

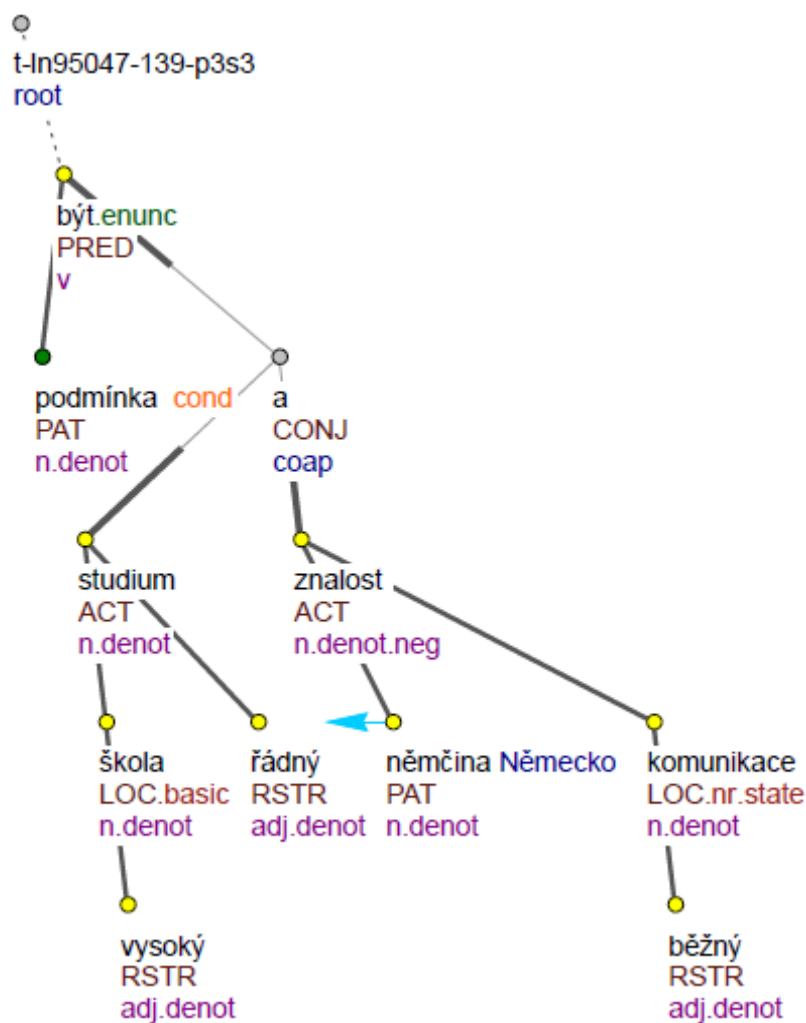
Obrázek 7 zobrazuje jednotlivé atributy konkrétního uzlu, které je možné na tektogramatické rovině vyplnit. Pro nás je důležitý zmíněný atribut „altlex_NP_type“ (na obrázku označený červeně). U daného atributu se zároveň vyplňuje typ diskurzího vztahu, tj. anotátor vybírá z nabídky typů diskurzí vztahů (zobrazených pod zkratkami, jako např. „cond“ pro vztah podmínky, „confr“ pro vztah konfrontace, „conj“ pro vztah konjunkce apod.). Na obrázku 7 anotátor vybral diskurzí vztah podmínky.

Pro ilustraci jsme vybrali z Pražského závislostního korpusu následující příklad:

- (87) *O uplatnění v Německu usilují především studenti ze zemí střední a východní Evropy. Jde o maximálně tříměsíční výpomoci hlavně v zemědělství a gastronomii, a to pouze v době letních prázdnin.*
Podmínkou je řádné studium na vysoké škole a znalost němčiny na úrovni běžné komunikace.

V daném příkladě je mezi diskurzními argumenty vztah podmínky (srov. parafrázi *O uplatnění v Německu mohou usilovat zájemci pouze tehdy, pokud studují na vysoké škole a mají znalost němčiny na úrovni běžné komunikace.*). První diskurzní argument je slovesný, vyjádřený dvěma po sobě následujícími větami (*O uplatnění v Německu usilují především studenti ze zemí střední a východní Evropy. Jde o maximálně tříměsíční výpomoci hlavně v zemědělství a gastronomii, a to pouze v době letních prázdnin.*). Druhý diskurzní argument je vyjádřen dvěma koordinačně spojenými nominálními frázemi (*řádné studium na vysoké škole a znalost němčiny na úrovni běžné komunikace*). Mezi danými argumenty je diskurzní vztah podmínky vyjádřený sekundárním konektorem *podmínkou je*, jehož jádrem je podstatné jméno *podmínka*. U slova *podmínka* (resp. příslušného uzlu na tektogramatické rovině) tedy bude vyplněn atribut „altlex_NP_type“ s diskurzním vztahem podmínky (viz předchozí obrázek 7).

Ve stromové struktuře se vyplnění atributu zobrazí tím způsobem, že se u jádra příslušného sekundárního konektoru (tj. zde u slova *podmínka*) objeví zkratka vybraného diskurzního vztahu (v tomto případě zkratka *cond*), srov. obrázek 8.



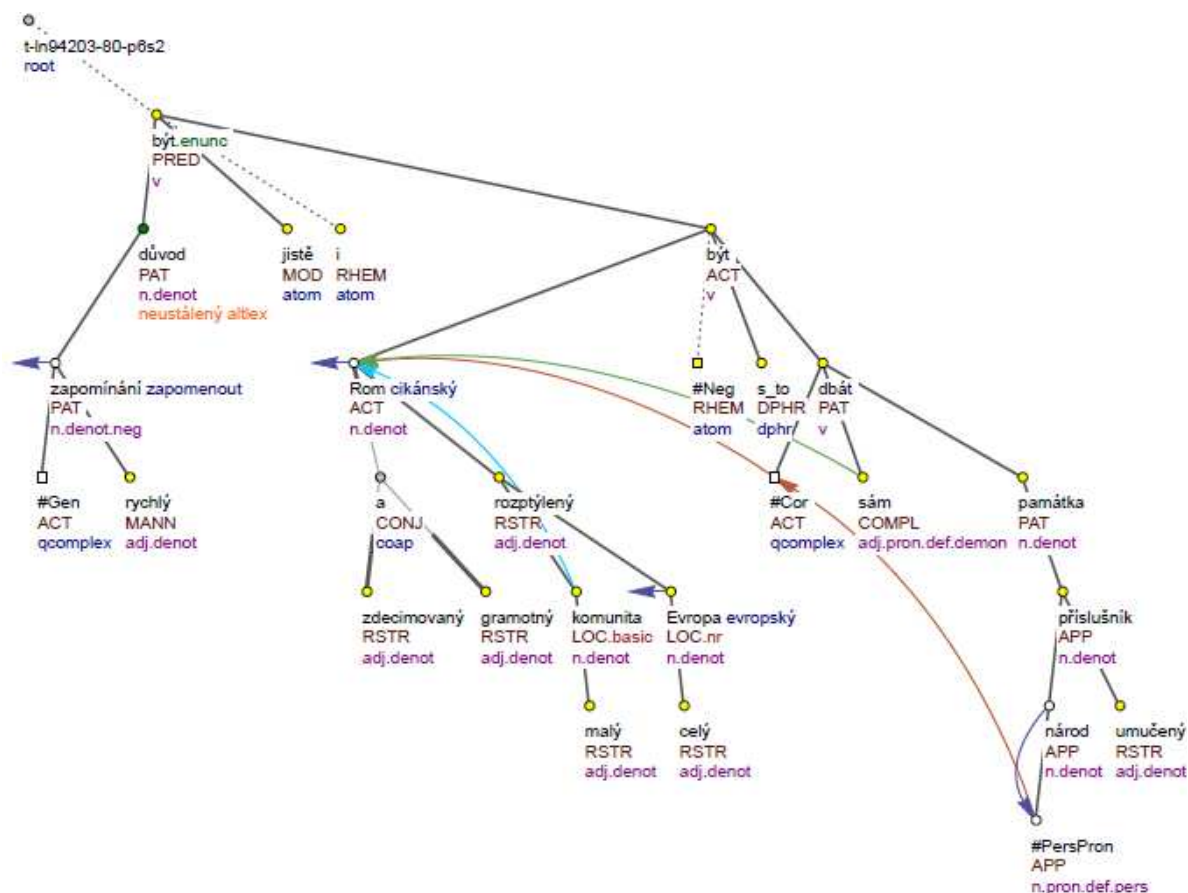
Obrázek 8: Anotace diskurzního vztahu s nominálním argumentem v Pražském závislostním korpusu, zobrazena stromová struktura věty *Podmínkou je řádné studium na vysoké škole a znalost němčiny na úrovni běžné komunikace*, viz předchozí příklad 87.

Z praktických důvodů (tj. aby bylo možné lépe zpracovat a automaticky vyhodnotit výsledky anotace) se tedy anotace diskurzních vztahů s nominálními a vztahů s verbálními argumenty technicky liší. Nicméně po zpracování výsledků anotace bude možné anotaci nominálních argumentů sjednotit s anotací argumentů verbálních, aby byl i v tomto případě explicitně zachycen rozsah diskurzních argumentů.

5.1.2.3 Volné konektivní fráze

Jak jsme uvedli v kapitole 3.7.2, jako diskurzní konektory chápeme takové jazykové prostředky, které mají (téměř) univerzální platnost ukazatelů diskurzních vztahů, tj. které můžeme k vyjádření příslušného diskurzního vztahu použít v různých kontextech (např. *díky tomu, důvodem je, stručně řečeno, to znamená* apod.).

Ostatní jazykové prostředky, které mají schopnost signalizovat diskurzní vztah pouze v omezeném počtu kontextů (srov. např. *díky této skulince, díky tomuto zvýšení* apod.), nazýváme volné konektivní fráze. I ty jsme se rozhodli do anotace zahrnout, abychom mohli např. porovnat jejich počet s počtem sekundárních konektorů v textu apod. Nicméně vzhledem k tomu, že se nejedná o univerzální konektory, anotovali jsme je pouze poznámkou upozorňující na jejich výskyt v korpusu (tj. nezachycovali jsme např. typ diskurzního vztahu, který vyjadřují, nebo rozsah jejich diskurzních argumentů). Pro označení volných konektivních frází jsme zvolili pracovní poznámku „neustálený altlex“ (odrážející skutečnost, že se daná fráze v jazyce neustálila jako univerzální konektor) – srov. obrázek 9 z PDT:



Obrázek 9: Ukázka anotace volné konektivní fráze z Pražského závislostního korpusu. Stromová struktura zachycuje souvětí *Důvodem rychlého zapomínání bylo jistě i to, že zdecimovaní a negramotní Romové, rozptýlení v malých komunitách po celé Evropě, nebyli s to sami dbát o památku umučených příslušníků svého národa*, viz příklad 88.

Stromová struktura na obrázku 9 zachycuje následující souvětí (v závorkách k němu uvádíme i bližší kontext):

- (88) (*Donedávna byla historie cikánského tábora v Osvětimi-Březince takřka zapomenuta.*)
Důvodem rychlého zapomínání bylo jistě i to, že zdecimovaní a negramotní Romové, rozptýlení v malých komunitách po celé Evropě, nebyli s to sami dbát o památku umučených příslušníků svého národa.

V uvedeném příkladu je diskurzivní vztah příčiny a důsledku signalizovaný spojením *důvodem rychlého zapomínání bylo jistě i to, že*. Příklad bychom mohli parafrázovat následujícím způsobem (s rozvitým primárním konektorem *jistě i proto, že* či sekundárním konektorem *důvodem bylo jistě i to, že*):

- (89) *Donedávna byla historie cikánského tábora v Osvětimi-Březince takřka zapomenuta – jistě i proto, že / důvodem bylo jistě i to, že zdecimování a negramotní Romové, rozptýlení v malých komunitách po celé Evropě, nebyli s to sami dbát o památku umučených příslušníků svého národa.*

Spojení *důvodem rychlého zapomínání bylo jistě i to, že* explicitně v daném příkladu signalizuje vztah příčiny a důsledku, nicméně platnost univerzálního konektoru nemá (tj. není jako konektor využitelné ve více kontextech, ale je vázané na tento konkrétní). Textová návaznost je v tomto případě zajištěna (kromě jiného) aktuálním členěním větným – réma předchozí věty se stává tématem věty následující (*zapomenuta – zapomínání*).

Na obrázku 9 je daná věta zachycena ve stromové struktuře, u slova (uzlu) *důvod*, tj. u jádra konektivní fráze, je zobrazena oranžová poznámka „neustálený altlex“. Metoda testování univerzálnosti jednotlivých konektivních struktur bude spolu s výsledky představena v kapitole 6.9.

5.1.2.4 Slovesné sekundární konektory

Při anotaci úvodního vzorku dat se ukázalo, že nejproblematictějšími strukturami (s ohledem na jednotnost anotace) jsou slovesné sekundární konektory. Z tohoto důvodu jejich anotaci představíme blíže.

Základem či jádrem slovesných sekundárních konektorů je sloveso. Dané struktury (resp. ty, které byly zachyceny v rámci anotace v PDT) pak můžeme rozdělit do dvou skupin podle toho, jak jsou u daného slovesa obsazené jednotlivé participanty, resp. jeho valenční pozice. První skupinu tvoří slovesa, která mají zpravidla dva obligatorní aktanty – aktor a patiens. Jeden z nich je přitom vyjádřen anaforickým výrazem odkazujícím k předchozímu diskurzivnímu argumentu, druhý představuje vlastní diskurzivní argument – srov. příklad 90:

- (90) *Na druhé straně vím, že v řadě malých obcí nejsou zaběhnuté politické strany, jednotlivec není schopen stranu vytvořit či založit její odbočku.*
To znamená, že bude kandidovat bez podpory politické strany.

V uvedeném příkladě má spojení *to znamená, že* funkci sekundárního diskurzivního konektoru (vyjadřuje zde diskurzivní vztah příčiny a důsledku a je nahraditelný primárním konektorem *proto* či *a tak*). Jádro konektoru – sloveso *znamenat* – je dvouvalenční, tj. vyžaduje doplnění

aktorem a patientem. Aby mělo výsledné spojení platnost sekundárního konektoru, musí být pozice aktora obsazena anaforickým výrazem (který odkazuje k předchozímu diskurzivnímu argumentu – srov. zde zájmeno *to*), v pozici patientu je pak druhý diskurzivní argument (v tomto případě vyjádřený vedlejší větou *bude kandidovat bez podpory politické strany*).

Obdobně se chová např. sloveso *kontrastovat* – srov. příklad 91:

(91) *Ačkoli produktivita evropských dolů stále stoupá, nejsou s to konkurovat levnému uhlí ze zámoří.*

S tím kontrastuje skutečnost, že poptávka po uhlí prudce stoupá, zvláště v asijských zemích, které se dnes na světovém trhu s uhlím podílejí více než 40 %.

V uvedeném příkladu je sekundárním diskurzivním konektorem spojení *s tím kontrastuje skutečnost, že*, jehož jádrem je sloveso *kontrastovat*. Toto sloveso má ve svém valenčním rámci dva obligatorní aktanty – aktor a patiens. Aby příslušné spojení v textu plnilo funkci sekundárního konektoru, musí být v pozici patientu (u tohoto slovesa) anaforický výraz odkazující k předchozímu argumentu (zde *s tím*), v pozici aktora je potom druhý diskurzivní argument (zde ve formě vedlejší věty, na kterou odkazuje kataforický výraz *skutečnost*).

Druhou skupinu pak tvoří slovesa, která se pojí s životným aktorem (srov. *zdůvodnit/zdůvodňovat, odůvodnit/odůvodňovat* apod.). Většinou se jedná o slovesa, v nichž role dalších participantů obsazují anaforický výraz odkazující k předchozímu argumentu a druhý diskurzivní argument – srov. příklad 92:

(92) *Vláda včera odmítla tři návrhy zásad zákona o referendu poslanců ČMUS, LB a ČSSD. V. Klaus to odůvodnil tím, že návrhy jsou v rozporu s ústavou.*

V uvedeném příkladě je diskurzivní vztah příčiny a důsledku (*Vláda včera odmítla tři návrhy zásad zákona..., protože jsou v rozporu s ústavou.*) vyjádřený spojením *to odůvodnil tím, že*. Jádrem tohoto sekundárního konektoru je sloveso *odůvodnit* s životným aktorem (*V. Klaus*), v pozici patientu je anaforické zájmeno *to* (odkazující k prvnímu diskurzivnímu argumentu). Vedlejší věta (*návrhy jsou v rozporu s ústavou*) představuje druhý diskurzivní argument. Jako sekundární konektor v tomto případě chápeme spojení *to odůvodnil tím, že* (nahraditelné

primárním konektorem *protože*). Životný aktor (který se mění v závislosti na kontextu) u tohoto typu sloves k sekundárnímu konektoru nepočítáme.

V tomto ohledu jsou sekundární konektory se slovesným jádrem primárním konektorům nejbližší – slovesa totiž vyžadují různá doplnění v závislosti na valenci a na kontextu (viz životné aktory u předešlého typu), a je u nich tedy velmi nízká pravděpodobnost ustrnutí v určité formě a následná gramatikalizace.

5.1.2.5 Nové vztahy

Klasifikace sémantických (či přesněji sémanticko-pragmatických) typů diskurzních vztahů byla převzata, jak již bylo řečeno, z anotace primárních konektorů (srov. podkapitola 3.5.3.1). Ovšem vzhledem k tomu, že sekundární konektory tvoří v porovnání s primárními značně heterogenní skupinu, můžeme předpokládat, že i klasifikace typů diskurzních vztahů bude u sekundárních konektorů jiná než u primárních. Z tohoto důvodu dostali anotátoři sekundárních konektorů pokyn, aby případné „nové“ typy vztahů (tj. ty, které nejsou na původním seznamu sestaveném pro primární konektory), označovali poznámkami. Na základě těchto poznámek bude možné posléze klasifikaci typů vztahů rozšířit.

5.2 Evaluace anotace

Anotace sekundárních konektorů v Pražském závislostním korpusu začala v roce 2012 a trvala tři roky (včetně kontrol a měření anotátorské shody). Anotace je zatím nepublikovaná, ale připravuje se k vydání jako součást jedné z budoucích verzí PDT. Jak již bylo zmíněno výše, anotaci prováděly tři anotátorky (vystudované lingvistky, které se podílely již na anotaci primárních konektorů v PDT). Anotace zachycuje (stejně jako u primárních konektorů) diskurzní vztahy mezivětné i vnitrovětné.

V průběhu anotace byla změřena mezianotátorská shoda, která slouží především k vyhodnocení úspěšnosti anotačních principů na konkrétních datech, k hledání případných problémových jevů, odhalení nejčastějších anotátorských chyb apod. Měření anotátorské shody³² proběhlo na 500 větách anotovaných dvěma anotátorkami zároveň, zaměřili jsme se přitom na sedm různých aspektů. 1. Nejprve byla změřena celková shoda v rozpoznání existence vztahu (tj. do jaké míry se anotátorky shodují v tom, že na daném místě v textu existuje nějaký diskurzní vztah); 2. posléze celková shoda v rozpoznání existence i neexistence vztahu (tj. do měření byla započtena i shoda na tom, že na daném místě žádný diskurzní vztah není)³³; dále byla změřena shoda na 3. rozpoznání univerzálnosti daného konektivního prostředku (tj. zda se jedná o sekundární konektor či volnou konektivní frázi) a 4. určení typu argumentu (tj. zda se jedná o sekundární konektor mezi verbálními či nominálními argumenty); 5. dále shoda na sémantických typech diskurzních vztahů vyjádřených sekundárními konektory; 6. shoda na rozpoznání „jádra“ daného konektoru (např. slova *důvod* ve spojení *to je důvod, proč*) a 7. shoda na rozpoznání celkového rozsahu sekundárního konektoru (tj. zda anotátorky označily jako konektor např. celé spojení *to byl hlavní důvod, kvůli kterému* nebo pouze *to byl důvod* apod.).

V počítačové lingvistice se některé hodnoty mezianotátorské shody (např. shoda na ne/existenci vztahu) vypočítávají pomocí míry F1. Ta je harmonickým průměrem dvou hodnot – *recall* (udávající, kolik z existujících vztahů anotátor zachytil) a *precision* (udávající přesnost práce anotátora, tj. kolik vztahů zachytil anotátor správně).

³² Za změření mezianotátorské shody vděčíme Jiřímu Mírovskému.

³³ Tj. máme např. 10 možných výskytů konektorů signalizujících přítomnost diskurzního vztahu. První anotátor označil pět z nich jako konektory a druhý anotátor šest. Do prvního měření (tj. na existenci vztahu) bude v tomto případě započítáno pět výskytů konektorů (tj. takové výskyty, na kterých se oba anotátoři shodli, že se jedná o konektory), do druhého měření budou zároveň započítány i 4 výskyty, u kterých se anotátoři shodli, že se o konektory nejedná. V našem případě se tedy anotátoři neshodli pouze v jednom výskytu – první anotátor ho za konektor neoznačil, druhý ano.

Další hodnoty mezianotátorské shody (např. shoda na typech vztahů apod., srov. tabulka 3) se udávají (vedle prostého procentuálního vyčíslení) pomocí přesnější Cohenovy kappy, která zohledňuje i skutečnost, že některá místa anotátorské shody mohou být pouze náhodná.

Výsledky mezianotátorské shody pro anotaci sekundárních konektorů (a volných konektivních frází) v Pražském závislostním korpusu jsou zobrazeny v tabulce 3.

Typ mezianotátorské shody	F1	%	Cohenova κ
1. Shoda na existenci vztahu	0,70		
2. Shoda na existenci i neexistenci vztahu	0,90		
3. Shoda na rozpoznání univerzálnosti		100	
4. Shoda na rozpoznání typu argumentu (VP vs. NP)		100	
5. Shoda na typech vztahu:	celkem	82	0,78
	pouze pro verbální argumenty	83	0,76
	pouze pro nominální argumenty	80	0,73
6. Shoda na rozpoznání jádra konektoru		78	
7. Shoda na rozsahu konektoru		81	

Tabulka 3: Mezianotátorská shoda pro anotaci sekundárních konektorů v PDT

Nejvyšší shoda (100 %) je u rozpoznání univerzálnosti daného výrazu jako diskurzního konektoru a rozpoznání typu diskurzního argumentu. Ukázalo se tedy, že anotátorky neměly obtíže s rozlišováním sekundárních konektorů a volných konektivních frází (k testování univerzálnosti konektivních frází viz kapitola 6.9) a s odlišnou anotací verbálních a nominálních argumentů (technicky tedy s odlišením tří typů anotace – diskurzních vztahů mezi verbálními argumenty pomocí šipky, diskurzních vztahů s alespoň jedním nominálním argumentem vyplněním atributu „altlex_NP_type“ a volných konektivních frází pomocí poznámky „neustálený altlex“).

Výsledky mezianotátorské shody při anotaci sekundárních konektorů jsme porovnali se shodou na primárních konektorech (převzatou z Poláková a kol., 2013). V rámci anotace primárních konektorů v Pražském závislostním korpusu byla změřena celková shoda na existenci diskurzního vztahu a shoda na určení sémantického typu vztahu – srov. tabulka 4.

Typ mezinotátorské shody	Primární konektory	Sekundární konektory
Shoda na existenci vztahu (F1)	0,83	0,70
Shoda na typech vztahu (%)	77	82
Shoda na typech vztahu (Cohenova κ)	0,71	0,78

Tabulka 4: Srovnání mezinotátorské shody pro anotaci primárních a sekundárních konektorů v Pražském závislostním korpusu

Z tabulky je patrné, že mezinotátorská shoda je pro primární a sekundární konektory srovnatelná. U primárních konektorů je vyšší shoda na existenci diskurzního vztahu vyjádřeného konektorem (F1: 0,83 vs. 0,70). Tato skutečnost není překvapivá vzhledem k značně vyšší heterogenitě a rozvolněnosti sekundárních konektorů. Sekundární konektory jsou na škále implicitně a explicitně vyjádřených diskurzních vztahů blíže implicitnosti než konektory primární – v některých případech totiž není snadné jednoznačně určit, zda je příslušný vztah pouze implicitní, tj. interpretovatelný pouze z obsahů příslušných výpovědí, nebo zda se na jeho signalizaci podílí konkrétní jazykový prostředek.

Anotátorkám dělalo například obtíže určování diskurzního vztahu ve větách se spojením *ve skutečnosti* (tj. to, zda se jedná o vztah implicitní nebo explicitní, tj. zda můžeme *ve skutečnosti* považovat za konektor nebo ne). Přechod mezi implicitností a explicitností diskurzních vztahů je tedy poměrně zamlžený³⁴ a rozhodování o tom, zda některé prostředky v textu působí či nepůsobí jako sekundární konektory, je pro anotátory obtížnější než určení, zda se v textu nachází nebo nenachází primární konektor.

Pokud si tedy představíme škálu mezi explicitními konektivními prostředky a implicitními vztahy, jeden pól budou tvořit primární (tj. ustálené, gramatikalizované) konektory, druhý výrazy bez konektivní funkce (diskurzní vztah je pouze implicitní³⁵). Mezi těmito dvěma póly stojí **sekundární konektory**, tedy (dosud plně) **negramatikalizované**³⁶, **univerzální konektivní struktury**, a **volné konektivní fráze**, tedy **negramatikalizované**, **neuniverzální konektivní struktury** – srov. následující schéma:

(univerzálnost)

implicitnost → **volné konektivní fráze** → **sekundární konektory** → **primární konektory**
(gramatikalizace)

³⁴ M. Taboádová (2009) dokonce zpochybňuje existenci čistě implicitních vztahů, viz kapitola 3.5.1.

³⁵ Přistoupíme-li na to, že implicitní diskurzní vztahy existují.

³⁶ Tj. jsou v (různé fázi) procesu gramatikalizace, který zatím není plně dokončen.

Z výsledků mezianotátorské shody je tedy zřejmé, že hranice mezi sekundárními konektory, volnými konektivními frázemi a implicitností (které leží na škále blízko sebe) je pro anotátory obtížněji definovatelná než hranice mezi implicitností a primárními konektory.

U rozpoznávání sémantických typů diskurzních vztahů je na druhou stranu nepatrně vyšší shoda u sekundárních konektorů – srov. 77 % (0,71 C. κ) pro primární konektory a 82 % (0,78 C. κ) pro sekundární konektory. Vyšší shoda může být v tomto případě dána tím, že většina sekundárních konektorů obsahuje „transparentní“ slovo (např. slovo *podmínka*, *důsledek*, *příčina* apod.), které svým lexikálním významem odkazuje k odpovídajícím diskurzním vztahům, např. vztahům podmínky, příčiny a důsledku apod. (ačkoli tomu tak nemusí být vždy).

Celkově je tedy výsledná mezianotátorská shoda pro anotaci sekundárních konektorů uspokojivá (tj. srovnatelná s jinými anotacemi v rámci nadvětných jevů). Pro anotace na velkých datech přitom obecně platí, že se mezianotátorská shoda mění v závislosti na jazykové rovině, která je zpracovávána. Např. celková shoda pro anotaci na morfologické rovině Pražského závislostního korpusu byla 97 % (srov. Hajič, 2006), pro anotaci textové koreference 72 % a pro anotaci asociační anafory 46 % (Poláková a kol., 2013). Ačkoliv tato čísla nejdou přísně vzato přímo srovnávat, protože se jedná o různé anotační úlohy a různé míry měření, lze z nich vyzorovat celkovou tendenci, tj. mezianotátorská shoda „klesá s tím, jak se zvyšuje rovina jazyka“.

6 Výsledky práce

V následující části práce představíme hlavní výsledky získané analýzou anotace sekundárních konektorů v Pražském závislostním korpusu. Zaměříme se přitom na srovnání získané charakteristiky sekundárních konektorů s obdobnou anotací primárních konektorů (srov. Poláková a kol., 2012b), které by mohlo přispět k obecným teoretickým závěrům o diskurzních konektorech v češtině. Dílčí studií je také popis testování a vyhodnocení univerzálnosti diskurzních konektorů, jejíž výsledky byly při anotaci použity.

V práci dále představíme jednotlivé části výzkumu provedené na výše popsané anotaci, tj. drobné srovnání sekundárních konektorů v psané a mluvené češtině a možnosti automatického vyhledávání vybraných sekundárních konektorů v Pražském závislostním korpusu.

6.1 Charakteristika sekundárních konektorů na základě analýzy nově anotovaných dat PDT

Výsledky analýzy nově provedené anotace sekundárních konektorů v Pražském závislostním korpusu jsou zobrazeny v tabulce 5. Tabulka zachycuje počet výskytů sekundárních konektorů a volných konektivních frází. Zohledňuje zároveň skutečnost, zda se daný konektivní prostředek vyskytl mezi dvěma slovesnými argumenty (v tabulce je tento případ označen jako „VP_VP“) či mezi argumenty, z nichž byl alespoň jeden nominální (v tabulce „VP_NP“).

Typ konektivního prostředku	VP_VP	VP_NP	Celkem
Univerzální sekundární konektory	924	237	1 161
Neuniverzální volné konektivní fráze	79	72	151
Celkem	1 003	309	1 312

Tabulka 5: Výsledky anotace sekundárních konektorů v Pražském závislostním korpusu

Výsledky anotace ukázaly, že Pražský závislostní korpus obsahuje v současné době 1 161 výskytů sekundárních konektorů a 151 výskytů volných konektivních frází.

Z hlediska syntaktické povahy diskurzních argumentů celkově převažují vztahy mezi dvěma slovesnými argumenty – v tomto případě se na výsledku odráží skutečnost, že ne

všechny konektivní struktury dovolují nominalizaci druhého argumentu (srov. např. předložkové fráze typu *díky tomu*, některá ustálená spojení typu *stručně/jednoduše/obecně... řečeno* apod.). K nominalizaci druhého diskurzního argumentu tíhnou především určité struktury (obsahující predikaci) – srov. např. *příčinou je, důsledkem je, výsledkem je, podmínkou je* apod. (srov. kapitola 3.7.1).

Výsledky anotace sekundárních konektorů jsme porovnali s výsledky anotace primárních konektorů v PDT – srov. tabulka 6:

	Výskyt v PDT	%
Primární konektory	20 255	94,6
Sekundární konektory	1 161	5,4
Celkem	21 416	100

Tabulka 6: Srovnání výsledků anotace primárních a sekundárních konektorů v Pražském závislostním korpusu

Z porovnání výsledků obou anotací vyplývá, že z celkového počtu diskurzních vztahů vyjádřených explicitně (tj. pomocí konkrétních jazykových prostředků) tvoří primární konektory 94,6 %, sekundární 5,4 % (viz graf 1). Z výsledků je patrné, že označení primární a sekundární je v tomto ohledu příznačné – k vyjadřování diskurzních vztahů v textu slouží primárně (tj. z 94,6 %) primární konektory.

I když výskyt primárních konektorů v textech převažuje, není počet sekundárních konektorů v PDT zanedbatelný. Anotace sekundárních konektorů obohatila dosavadní diskurzni anotaci PDT o 1 161 nových vztahů, z toho je zřejmé, že anotace postrádající sekundární konektory by nebyla úplná.



Graf 1: Rozložení diskurzních konektorů v Pražském závislostním korpusu

Anotace sekundárních konektorů může zároveň usnadnit výzkum implicitních diskurzních vztahů v textu (tj. vztahů, které nejsou v povrchové rovině vyjádřeny žádným jazykovým prostředkem). Díky dokončené anotaci primárních i sekundárních konektorů bude možné snadněji vyhledávat implicitní vztahy podle míst, ve kterých se nevyskytuje žádný primární ani sekundární konektor, ale u kterých přítomnost diskurzního vztahu předpokládáme (vzhledem k tomu, že text je chápán jako síť vzájemně propojených vztahů, ve kterých žádná jednotka nezůstává osamocena).

6.1.1 Lexikální charakteristika sekundárních konektorů

Z lexikálního hlediska představují sekundární konektory (dosud) negramatikalizované fráze či spojení, které vykazují větší či menší míru ustálenosti. Některé z nich mají své jádro (nejčastěji podstatné jméno – *důvod, příčina, podmínka, vysvětlení* atd., či sloveso – *odůvodnit, zapříčinit, vysvětlit* atd.), které signalizuje určitý typ diskurzního vztahu (srov. např. slovo *podmínka* signalizuje diskurzní vztah podmínky, *příčina* vztah příčiny a důsledku apod.) a které vstupuje do různých kombinací – srov. např. *příčinou je; vidět příčinu v; hledat příčinu v* atd. Tyto sekundární konektory nejsou ustrnulé, v textu se mohou vyskytovat v různých obměnách, např. v různých časech (*důvodem je* vs. *důvodem bylo*), některé slovesné struktury v aktivu i pasivu (*odůvodnil to tím* vs. *je to odůvodněno tím*) atd.

Mezi sekundární konektory patří ovšem i ustálenější (tj. gramaticky i lexikálně omezená – srov. Čermák, 2007) víceslovná spojení, přičemž některá z nich mohou vykazovat nepatrnou míru variability (srov. např. *obecně/krátce/zjednodušeně řečeno*³⁷). Tyto sekundární konektory jsou obvykle struktury, jejichž podstatou je jistá nepravidelnost či anomálie (tj. např. omezená kombinovatelnost, nemožnost paradigmatické záměny apod. – srov. Čermák, 2007: 30–31). Patří sem také ustrnulé jednoslovné výrazy typu *přeloženo* či *rozumějme*, které mají funkci konektoru pouze v jedné ze svých gramatických forem (v ostatních platnost konektoru nemají – srov. *Přeložil knihu do angličtiny* či *Rozumím německy*).

Ustálené sekundární konektory bychom mohli dále dělit na menší podskupiny (podle knihy Františka Čermáka *Frazeologie a idiomatika*, 2007) – např. na gramatické frazémy a idiomy (srov. skupinu prepozičních spojení vzniklých smíšením synsémantik a autosémantik – *z důvodu, s cílem, v důsledku* apod.) či na propoziční frazémy a idiomy (jeden typ těchto frazémů a idiomů nese dokonce označení „konektivní“ – srov. např. *stručně řečeno* – Čermák, 2007: 308, 442).

6.1.2 Syntaktická charakteristika sekundárních konektorů

a) Syntaktické fráze

Ze syntaktického hlediska můžeme sekundární konektory analyzovat podle toho, jaké typy syntaktických frází tvoří. Nejpočetnější jsou v PDT verbální fráze (jejichž hlavou je sloveso, např. *odůvodnil to, to znamená, vysvětlila to* atd.; verbální fráze tvoří celkem 37 % v rámci 1 161 výskytů sekundárních konektorů v PDT), předložkové fráze (obsahující sekundární předložku a anaforický výraz, např. *díky tomu, vzhledem k tomu, na základě toho, v důsledku toho, z tohoto důvodu* atd.; předložkové fráze tvoří 25 % v rámci všech výskytů sekundárních konektorů v PDT) a fráze zakládající větné struktury skládající se především ze spojení slovesa s oslabeným lexikálním významem (např. *být, uvést, sloužit, tvořit*) a podstatného jména, které nese jádro lexikálního významu dané fráze (srov. např. *důvod je jednoduchý, příčinou je, uvést jako příklad, jako příklad slouží, výjimku tvoří* atd.; celkově tyto struktury tvoří 17 % v rámci všech výskytů sekundárních konektorů v PDT).

³⁷ U uvedených spojení výrazně narůstá pragmatický prvek, dané struktury vyjadřují postoj mluvčího k obsahu výpovědi. V tomto ohledu se liší od jiných víceslovných struktur, u kterých převládá prvek sémantický (srov. např. *z tohoto důvodu*). V této práci se sémanticko-pragmatickým popisem konektorů zabývat nebudeme. Více k dané problematice viz např. Hoffmannová (1983).

b) Větněčlenská platnost sekundárních konektorů

Většina daných frází či jejich komponentů (tj. 87 % v rámci všech výskytů³⁸ sekundárních konektorů v PDT) plní ve větě roli větných členů – tj. např. předložkové fráze typu *kvůli tomu, z tohoto důvodu* apod. plní funkci příslovečných určení, jednotlivé složky frází typu *jako příklad uvedl* či *to znamená* fungují jako různé větné členy (*uvedl, znamená* jako přísudek, *to* jako podmět atd.). Ostatní případy (tj. 13 %) tvoří konstrukce typu *stručně řečeno, jak je vidět* aj., které slouží jako větné modifikátory (vztahující se k obsahu celé věty či ke způsobu jeho vyjádření, tj. komentující obsah či styl celé výpovědi).

Tato skutečnost výrazně odlišuje sekundární konektory od konektorů primárních, které jako autosémantika neplní ve větě role větných členů ani větných modifikátorů.

c) Syntaktické chování sekundárních konektorů v textu – zajímavosti

Heterogenita sekundárních konektorů se projevuje i v jejich syntaktickém chování v textu. Je zajímavé, že některé z nich se v textu objevují ve formě hlavních či (nepravých) vedlejších vět – srov. příklady 93 a 94:

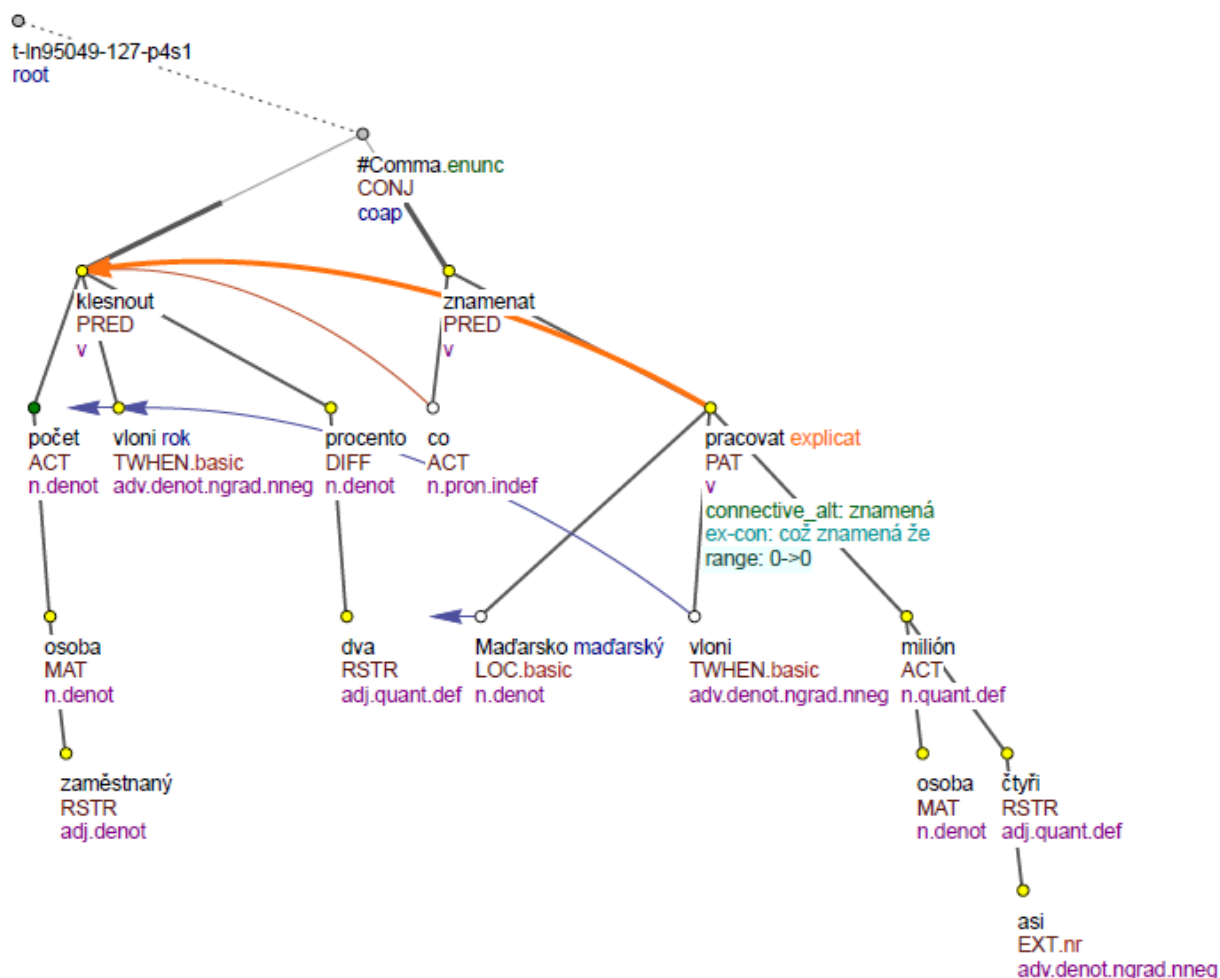
- (93) *Jako téměř vždy, když jde o velmi lukrativní zakázku, ozvaly se po schválení výsledků městskou radou námitky těch, kdo nezvítězili a cítí se poškozeni.*
To je důvod, proč ještě nebyla podepsána smlouva s dodavatelem, či přesněji organizátorem rekonstrukčních prací už třetí měsíc uzavřeného objektu.
- (94) *Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.*

V příkladě 93 se vyskytuje sekundární konektor *to je důvod proč*, který vyjadřuje vztah příčiny a důsledku mezi dvěma argumenty (*jako téměř vždy, když jde o velmi lukrativní zakázku, ozvaly se... a ještě nebyla podepsána smlouva s dodavatelem...*). Sekundární konektor se zde objevuje ve formě hlavní (řídící) věty (s přesahem až do věty vedlejší – zahrnuje v sobě i uvozovací výraz *proč*). Je tedy velice zajímavé, že druhý diskurzí argument se dostává do pozice věty závislé na konektoru. Jinými slovy, sekundární konektor může být dokonce syntakticky výše než diskurzí argument.

³⁸ Do analýzy jsme nezahrnuli pouze případy typu *Důvod je jednoduchý.*, tj. sekundární konektory tvořící samostatné věty, které stojí syntakticky mimo diskurzí argumenty, a neplní tedy funkci větných členů.

Sekundární konektor *což znamená* v příkladě 94 opět syntakticky řídí druhý argument (v *Maďarsku loni pracovaly...*) a zároveň je nepravou vedlejší větou prvního argumentu (*počet zaměstnaných osob klesl...*). Pokud si daný příklad představíme jako závislostní strom, první argument a sekundární konektor budou syntakticky výše než druhý argument. Vidíme tedy, že **diskurzí argumenty (z hlediska diskurzu rovnocenné úseky textu), nemusejí být na stejné úrovni i syntakticky.**

Srov. obrázek 10 z Pražského závislostního korpusu (věty uvozené vztahným zájmenem *což* bývají hodnoceny jako nepravé vedlejší věty vztahné navazovací, tj. tvoří „hypotaktický podtyp souvětneho typu vyjadřujícího sémantickou paralelnost propozic“ – srov. *Mluvnice češtiny 3*, 1987: 534; v pojetí Pražského závislostního korpusu je zájmeno *což* chápáno jako slovo výrazně na přechodu k souřadícím spojkám, a proto jsou zde propozice spojené tímto výrazem zpracovány jako souřadné struktury).



Obrázek 10: Ukázka závislostního stromu se sekundárním konektorem z Pražského závislostního korpusu, stromy zachycují souvětí *Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.*

Z obrázku je patrné, že první diskurzí argument (v závislostním stromě reprezentovaný nejvyšším uzlem, tj. slovesem *klesnout*) je v koordinačním vztahu se sekundárním konektorem (reprezentovaným slovesným uzlem *znamenat*), na kterém je syntakticky závislý druhý diskurzí argument (reprezentovaný slovesným uzlem *pracovat*). V závislostním stromě je označen diskurzí vztah explikace zachycený pomocí oranžové šipky mezi uzly *klesnout* a *pracovat* (zkratka *explicat* označuje typ diskurzího vztahu, tj. explikaci). Jako sekundární konektor je označeno spojení *což znamená, že* (s lexikálním jádrem *znamená*). K podrobnému popisu anotace sekundárních konektorů – viz kapitola 5.1.

Podobné syntaktické chování se u primárních konektorů (vzhledem k jejich slovnědruhové příslušnosti a syntaktické povaze) nemůže nikdy vyskytnout a přispívá tak k jedinečnému postavení sekundárních konektorů mezi konektory.

Další zajímavou skutečností je, že některé sekundární konektory mohou mít dokonce formu samostatné věty – srov. např. *Důvod je jednoduchý/jasný/zřejmý*. – srov. příklad 95:

- (95) *S vašimi akciemi se musí obchodovat na burze, ale Wall Street vám nabízí cenu z RMS.*
Důvod je jednoduchý.
V RMS je cena většiny akcií nižší než na burze.

V daném příkladu věta *Důvod je jednoduchý* funguje jako sekundární konektor vyjadřující vztah příčiny a důsledku. V tomto případě může být nahrazen i konektorem primárním – srov.: *S vašimi akciemi se musí obchodovat na burze, ale Wall Street vám nabízí cenu z RMS.* *Jednoduše proto, že v RMS je cena většiny akcií nižší než na burze.*

V tomto ohledu se daný typ sekundárních konektorů výrazně liší od konektorů primárních – sekundární konektory mohou stát jako samostatné věty zcela mimo diskurzí argumenty (úseky textu), které spojují. Tyto sekundární konektory tedy vykazují vysokou míru nezávislosti, protože tvoří syntakticky i sémanticky úplné úseky textu.

d) Modifikovatelnost sekundárních konektorů

Další charakteristickou vlastností sekundárních konektorů je, že jsou poměrně často modifikovatelné (s výjimkou ustálených lexikálních frází typu *rozumějme* či *jak je vidět*). U konektorů, jejichž jádrem je podstatné jméno, je modifikované příslušné jméno – srov. např.

hlavním/jediným/prvním/významným/rozumným... důvodem je, u slovesných frází je rozvíjeno sloveso – srov. např. *jasně to odůvodnil...*

Zajímavé je, že rozvíjeny mohou být i sekundární konektory v podobě předložkových frází – srov. příklad 96 z Pražského závislostního korpusu se sekundárním konektorem *vinou toho* (signalizující diskurzni vztah příčiny a důsledku), který je v tomto případě rozvit výrazem *především*:

- (96) *Za předčasný konec jejich sotva půlhodinového vystoupení si však mohli sami diváci, kteří v průběhu koncertu Luscious Jackson začali vyvolávat jméno Beastie Boys. Především vinou toho nakonec nedošlo ani na největší hit kapely – Keep On Rockin' It.*

V tomto ohledu se sekundární konektory liší od konektorů primárních, které být rozvíjeny v takovém rozsahu většinou nemohou (nelze například říci **významně ale* apod.). Existují ovšem výjimky, jako je např. konektor *proto, že* – srov. *hlavně/jedině/jednoduše proto, že*. Modifikovatelnost je v tomto případě ovšem omezená – srov. **významně proto, že* apod.

Z tohoto hlediska může autor v textu použít sekundární konektory právě proto, že mu umožňují přesněji vyjádřit jeho komunikační záměr (pokud chce např. autor sdělit, že důvod k něčemu byl významný, nemůže použít primární konektor, který modifikovatelnost tímto slovem neumožňuje).

6.1.3 Sémantická charakteristika sekundárních konektorů

Ze sémantického hlediska můžeme popsat sekundární konektory jako **jazykové struktury či fráze obsahující alespoň jedno autosémantické slovo** (na rozdíl od synsémantických primárních konektorů).

Sama jádra jednotlivých konektorů bývají většinou autosémantika – srov. podstatná jména typu *vysvětlení, příčina, podmínka* atd., slovesa typu *zdůvodnit, zapříčinit, znamenat, upřesnit* atd., příslovce typu *přesněji, podobně* atd.

U předložkových frází je slovem signalizujícím příslušný sémantický typ diskurzniho vztahu sekundární předložka (typu *díky, kvůli, na rozdíl od, v důsledku, vinou* atd.), tedy synsémantikem. Nicméně aby se příslušná předložka mohla stát součástí sekundárního konektoru, musí se spojit s plnovýznamovým (anaforickým) slovem – tj. jako sekundární konektor fungují v textu fráze typu *díky tomu, v důsledku toho, vinou toho* apod. (Z tohoto

důvodu nechápeme jako konektor samotné *díky*, *kvůli*, ale celou předložkovou frází *díky tomu*, *kvůli tomu* apod.)

Od příkladu předložkových frází se dostáváme k dalšímu charakteristickému rysu sekundárních konektorů. Všechny sekundární konektory totiž obsahují (explicitně či implicitně) anaforickou referenci odkazující k předchozímu kontextu. Z tohoto hlediska můžeme sekundární konektory rozdělit do tří skupin podle toho, zda:

1) anaforickou referenci v povrchové rovině vyjádřit mohou (např. je možné říci *výsledkem je i výsledkem toho je, důsledkem je i důsledkem toho je* atd.; v PDT tvoří tyto struktury 55 % v rámci 1 161 výskytů sekundárních konektorů);

2) anaforickou referenci v povrchové rovině vyjádřit musí (např. kvůli valenci či rekcii – srov. *kvůli tomu tam půjdu* vs. **kvůli tam půjdu*; *s tím kontrastuje to, že...* vs. **kontrastuje to, že...* atd.; tyto struktury tvoří v PDT 33 %);

3) anaforickou referenci v povrchové rovině zpravidla nevyjadřují (v tomto případě se totiž jedná o ustálená spojení fungující jako sekundární konektory v dané kombinaci a formě, tj. jako konektory se ustálily v určité lexikální podobě a varianta s anaforickou referencí není obvyklá a běžně používaná – srov. např. *jak je vidět* vs. (?)*jak je z toho vidět, jednoduše řečeno* vs. (?)*k tomu jednoduše řečeno* atd.; v PDT tvoří tyto struktury 12 %).

6.2 Porovnání výsledků anotace sekundárních konektorů v češtině a altlexů v angličtině na základě dat PDT a PDTB

Výsledky naší anotace jsme dále porovnali s anotací tzv. alternativních lexikálních vyjádření diskurzních konektorů (altlexů) v angličtině (které se blíží sekundárním konektorům) v korpusu Penn Discourse Treebank (obsahujícím přibližně 49 000 vět, tedy srovnatelný počet jako Pražský závislostní korpus, oba korpusy obsahují publicistické texty).

Penn Discourse Treebank (PDTB) obsahuje celkem 624 výskytů alternativních vyjádření konektorů, tj. 1,5 % z celkového počtu identifikovaných diskurzních vztahů (srov. Prasadová a kol., 2010, 2014). Do celkového počtu diskurzních vztahů byly ovšem započteny také implicitní vztahy, tzv. nulové vztahy (tj. mezi danými sousedními větami není žádný diskurzní vztah) a vztahy mezi jmennými entitami, jejichž anotaci Pražský závislostní korpus zatím neobsahuje. Relevantnější srovnání s výsledky anotace sekundárních konektorů v češtině tedy poskytne pouze podíl diskurzních konektorů a jejich alternativních vyjádření (altlexů) v Penn Discourse Treebanku – srov. tabulka 7 (hodnoty pro PDTB byly převzaty z článku Prasadové a kol., 2014). Pojem altlexy (z PDTB) se ovšem používá pro širší množinu jazykových prostředků než pojem sekundární konektory – v rámci altlexů se například nerozlišují univerzální a neuniverzální konektivní struktury (viz kapitola 3.7.4.5). Pojem altlexy tedy v naší terminologii zahrnuje sekundární konektory a (neuniverzální) volné konektivní fráze zároveň.

Korpus	Konektory (= primární)	%	„Altlexy“ (= sekundární kon. + volné kon. fráze)	%	Celkem
Pražský závislostní korpus	20 255	93,9	1 312	6,1	21 567
Penn Discourse Treebank	18 459	96,7	624	3,3	19 083

Tabulka 7: Srovnání anotace diskurzních vztahů v češtině a angličtině

Penn Discourse Treebank obsahuje 18 459 vztahů vyjádřených (primárními) diskurzními konektory a 624 vztahů vyjádřených altlexy, tj. altlexy tvoří 3,3 % všech explicitních konektivních prostředků, zatímco v Pražském závislostním korpusu 6,1 %. Autoři projektu Penn Discourse Treebank ovšem připouštějí (srov. Prasadová a kol., 2010), že anotace altlexů není v PDTB zcela konzistentní, předpokládají přitom, že skutečný výskyt některých altlexů

může být dvakrát až třikrát vyšší než v jejich současné anotaci. V tomto ohledu by mohly být počty altlexů (tj. sekundárních konektorů a volných konektivních frází) v Pražském závislostním korpusu a Penn Discourse Treebanku velmi podobné.

6.3 Rozdíly mezi primárními a sekundárními konektory na základě analýzy dat PDT

Na základě předchozí analýzy rozsáhlých korpusových dat můžeme dále přesněji vymezit pojem diskurzní konektory a zároveň lépe rozlišit obě podkategorie, tedy primární a sekundární konektory.

Můžeme konstatovat, že v našem pojetí jako diskurzní konektory označujeme takové jazykové prostředky, které mají v textu (téměř) **univerzální platnost ukazatelů diskurzních vztahů** (bez ohledu na syntaktickou povahu argumentů – tedy úseků textu – které jsou konektory spojeny, tj. diskurzní argumenty mohou být i nominalizované).

V rámci kategorie diskurzních konektorů vymezujeme dvě skupiny: primární konektory a sekundární konektory.

Primární konektory jsou synsémantika s univerzální platností ukazatelů diskurzních vztahů v rámci textu (tj. jedná se především o spojky a strukturující částice, příp. o některá příslovce a zájmena typu *potom, což*, která jsou na přechodu ke spojkám či částicím). Jedná se tedy většinou o gramatikalizované, jednoslovné výrazy (některé primární konektory ovšem mohou tvořit korelativní páry typu *bud' _anebo* či sousloví typu *i když*) bez větněčlenské platnosti. Až na výjimky jsou primární konektory nemodifikovatelné (srov. tabulka 8)³⁹.

Sekundární konektory jsou konektivní struktury s univerzální platností ukazatelů diskurzních vztahů obsahující autosémantickou složku. Jedná se o struktury, které zatím neprošly plnou gramatikalizací. Většinou jsou víceslovné, ve větě plní funkci větných členů (některé mohou fungovat i jako větné modifikátory). Sekundární konektory ve formě samostatných vět do syntaxe diskurzních argumentů v tomto smyslu nezasahují. Sekundární konektory jsou často modifikovatelné. Všechny obsahují implicitní anaforickou referenci k prvnímu argumentu, tato reference se podle typu sekundárního konektoru může objevit i v povrchové realizaci konektoru.

Jedinečnost sekundárních konektorů spočívá v tom, že mohou (na rozdíl od konektorů primárních) stát syntakticky výše než druhý diskurzní argument, tj. formálně mohou mít povahu hlavní věty (srov. příklady typu *To je důvod, proč tam nepůjdu.*). Některé sekundární konektory mohou tvořit samostatné věty, tj. syntakticky i sémanticky zcela úplné úseky textu

³⁹ Za primární konektory zároveň považujeme některé grafické značky (dvojtečku a pomlčku), číselné a písmenné výrazy sloužící ke strukturaci textu (např. 1., 2., 3.; a), b), c); i), ii), iii)) a větnou negaci (pokud v daném případě přispívá k vyjádření diskurzního vztahu).

(srov. příklad *Důvod je jednoduchý.*). Struktura některých sekundárních konektorů dovoluje nominalizaci druhého argumentu (srov. např. *Důsledkem je stárnutí obyvatelstva.*).

Díky tomuto idiosynkratickému chování mají sekundární konektory v jazyce jedinečné postavení a diskurzivní analýza, která tyto jazykové prostředky nezohledňuje, je podle našeho názoru neúplná. Z tohoto důvodu je zapotřebí brát označení primární a sekundární konektory především z frekvenčního hlediska (viz kapitola 6.3), nikoli jako označení hodnotící důležitost či význam obou skupin.

Primární konektory	Sekundární konektory
synsémantika	struktury s autosémantickou složkou
univerzální	univerzální
gramatikalizované	negramatikalizované
většinou jednoslovné	většinou víceslovné
bez větněčlenské platnosti	větné členy, větné modifikátory či samostatné věty
nemodifikovatelné (s výjimkami)	modifikovatelné (s výjimkami)
	anaforická reference k prvnímu argumentu (explicitní či implicitní)
	jedinečnost některých struktur: a) syntakticky výše než druhý argument b) forma samostatné věty c) nominalizace druhého argumentu

Tabulka 8: Charakteristika primárních a sekundárních konektorů

6.4 Srovnání primárních a sekundárních konektorů podle četnosti; slovesa říkání s konektivní funkcí

Primární a sekundární konektory s nejvyšším počtem výskytů v PDT jsou uvedeny v tabulce 9⁴⁰. V rámci sekundárních konektorů tabulka zachycuje počet jednotlivých konkrétních realizací sekundárních konektorů, tj. jsou zde např. odděleny výrazy *to znamená*, *že* a *znamená to*, *že*. Jedná se tedy o takové konkrétní realizace sekundárních konektorů, které se v PDT objevily nejčastěji⁴¹ (bez snahy sjednotit je pod zastřešující lemmata). Cílem bylo zjistit, jaké konkrétní podoby sekundárních konektorů jsou v reálných textech nejfrekventovanější (vzhledem k jejich velké formální rozrůzněnosti), tj. na jaké formě se daný typ sekundárního konektoru může postupně ustalovat.

⁴⁰ V tabulce uvádíme všechny formy výskytu sekundárních konektorů, které se v PDT objevily alespoň pětkrát, a odpovídající počet primárních konektorů.

⁴¹ Kompletní abecední a frekvenční seznam primárních a sekundárních konektorů zachycených v současné anotaci v PDT je uveden v příloze 1.

Primární konektory podle četnosti		Sekundární konektory podle četnosti	
<i>a</i>	5 765	<i>dodal</i>	121
<i>však</i>	1 521	<i>podobně</i>	60
<i>ale</i>	1 267	<i>v případě, že</i>	40
<i>když</i>	574	<i>vzhledem k tomu, že</i>	40
<i>protože</i>	525	<i>dodává</i>	36
<i>totiž</i>	460	<i>kromě toho</i>	30
<i>pokud</i>	403	<i>naproti tomu</i>	23
<i>:</i>	395	<i>to znamená, že</i>	22
<i>proto</i>	380	<i>v této souvislosti</i>	17
<i>tedy</i>	307	<i>případně</i>	13
<i>aby</i>	305	<i>příkladem je</i>	12
<i>pak</i>	295	<i>upřesnil</i>	12
<i>ovšem</i>	285	<i>znamená to, že</i>	12
<i>-li</i>	248	<i>z tohoto důvodu</i>	11
<i>také</i>	234	<i>důvodem je</i>	9
<i>neboť</i>	221	<i>díky tomu, že</i>	8
<i>–</i>	218	<i>výjimkou je</i>	8
<i>zatímco</i>	204	<i>výsledkem je</i>	8
<i>nebo</i>	191	<i>za podmínky, že</i>	8
<i>navíc</i>	182	<i>díky tomu</i>	7
<i>přítom</i>	181	<i>dodala</i>	7
<i>i když</i>	178	<i>to znamená</i>	7
<i>což</i>	174	<i>pokračuje</i>	6
<i>sice_ale</i>	165	<i>tím pádem</i>	6
<i>naopak</i>	152	<i>to neznámá, že</i>	6
<i>takže</i>	149	<i>to se týká</i>	6
<i>a tak</i>	141	<i>v takovém případě</i>	6
<i>a to</i>	118	<i>ve skutečnosti</i>	6
<i>dále</i>	117	<i>z toho důvodu, že</i>	6
<i>kdyby</i>	116	<i>jinak řečeno</i>	5
<i>tak</i>	112	<i>právě tak</i>	5
<i>rovněž</i>	106	<i>to souvisí s</i>	5
<i>proto, že</i>	99	<i>to vede k</i>	5

Tabulka 9: Primární a sekundární konektory s nejvyšším počtem výskytů v Pražském závislostním korpusu

Z tabulky je patrné, že výskyty jednotlivých forem primárních konektorů značně převyšují výskyty sekundárních konektorů. Zřetelně se zde také projevuje skutečnost, že primární konektory již prošly gramatikalizací, tj. jsou formálně i funkčně ustálené, zatímco konektory sekundární jsou rozvolněné a jsou teprve v procesu ustalování. V PDT se tak například objevují konektivní struktury se slovem *důvod* ve 180 výskytech (viz tabulka 19 v části o DIALOGU), ovšem nejsou používány v jedné, ustálené formě, ale v mnoha realizacích (např.

z tohoto důvodu; to je důvod, proč atd.). Nejčastěji se z nich objevuje spojení z *tohoto důvodu* (11 výskytů).

Tabulka ukazuje, že nejpočetnějším primárním konektorem v PDT je konektor *a* se 5 765 výskytů (k primárním konektorům jsou v PDT přitom řazeny také dvojtečka, pomlčka a číselné a písmenné výrazy sloužící ke strukturaci textu), nejfrekventovanějším sekundárním konektorem výraz *dodal* se 121 výskytů.

Sloveso *dodat/dodávat* tvoří spolu s dalšími slovesy říkání (tj. slovesy uvozujícími (značenou i neznačenou) přímou či nepřímou řeč) typu *upřesnit/upřesňovat*, *doplnit/doplňovat*, *pokračovat* apod. zvláštní kategorii. Od ostatních primárních i sekundárních konektorů se liší tím, že v sobě spojují dvě složky – prosté sloveso říkání a navazování na předchozí kontext zároveň. Např. tedy sloveso *dodal* můžeme rozložit na *a řekl*⁴² – srov. příklad 97 z PDT:

(97) *S kolegy jsem se seznámil až po prvním dějství, řekl Peter Dvorský.*

*Potom jsem měl plný kalendář, **dodal**.*

Uvedená slovesa říkání tedy předpokládají přítomnost jiné složky v předchozím kontextu (což je obecným rysem diskurzních konektorů). Jinými slovy, zmíněná slovesa se obvykle neobjevují v první větě textu, protože v sobě zahrnují informaci o tom, že již bylo něco řečeno – např. tedy *on dodal* implikuje, že on (příp. někdo jiný) něco řekl již předtím. V tomto smyslu tedy mohou být daná slovesa chápána jako indikátory diskurzních vztahů, protože se zcela zřetelně zapojují do struktury diskurzu.

V příkladě 97 tedy můžeme označit diskurzní vztah konjunkce mezi *řekl Peter Dvorský* a *dodal* vyjádřený slovesem *dodal*. Od ostatních diskurzních konektorů se tedy slovesa říkání liší tím, že v sobě lexikálně zahrnují dvě složky diskurzního vztahu (diskurzní argument – tj. to, že někdo něco říká, a diskurzní konektor – tj. to, že to někdo říká v reakci na něco; *dodal* v sobě tedy zahrnuje složky *a řekl*).

V uvedeném příkladě můžeme zároveň vysledovat dvě roviny diskurzních vztahů – první mezi slovesy říkání (*řekl – dodal*), druhý mezi samotnými obsahy neznačených přímých řečí, tj. mezi *s kolegy jsem se seznámil až po prvním dějství* a *jsem měl plný kalendář* je

⁴² V PDT se objevily i příklady typu *a dodal*, tj. spojení primárního a sekundárního konektoru. Navázání na předchozí kontext je zde podle našeho názoru vyjádřeno dvakrát (primárním konektorem *a* a lexikálním významem slovesa *dodat*).

diskurzni vztah nesoučasnosti (předčasnosti a následnosti) vyjádřený primárním konektorem *potom*.

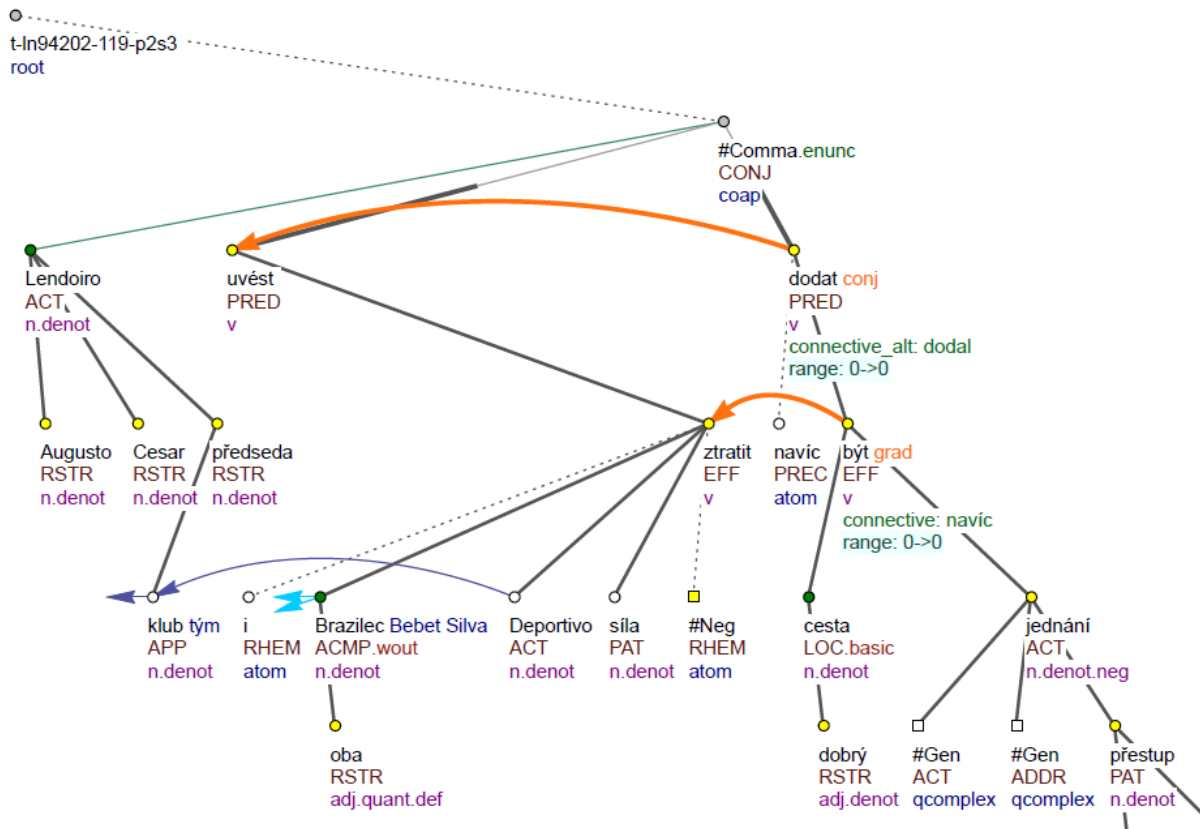
Obečně tedy můžeme říci, že daná slovesa říkání odkazují v textu k jinému slovesu říkání (*řekl – dodal, řekl – upřesnil, řekl – pokračoval* apod.), přičemž mezi vlastními obsahy příslušných přímých či nepřímých řečí může být další, nezávislý diskurzni vztah vyjádřený jinými konektory (či dalšími konektivními prostředky), v příkladě 97 primárním konektorem *potom*. Ve většině případů jsou tyto konektory součástí obsahů daných přímých či nepřímých řečí (tj. význam daných výpovědí a jejich syntaktická realizace jsou v souladu). V některých případech se ovšem může stát, že příslušný konektor se v povrchové rovině objeví syntakticky výše, tj. ve větě řídící, ačkoli sémanticky patří k větě závislé – srov. příklad 98 z PDT:

- (98) *(Loňský vicemistr španělské ligy se tak dostává do nezáviděníhodné situace, protože jeho druhá zahraniční akvizice a Bebetův krajan Mauro Silva má zraněné koleno a týmu bude chybět nejméně čtyři měsíce.)*
Předseda klubu Augusto Cesar Lendoiro uvedl, že i bez obou Brazilců Deportivo na síle neztratí, navíc dodal, že na dobré cestě jsou jednání o přestupu uruguayského útočníka Rubena Sosy z Interu Milán.

V uvedeném příkladě je mezi vyznačenými klauzemi (závislými na řídících větách uvozených slovesy říkání *uvedli* a *dodali*) diskurzni vztah gradace vyjádřený primárním konektorem *navíc* (srov. *I bez obou Brazilců Deportivo na síle neztratí, navíc jsou na dobré cestě jednání o přestupu uruguayského útočníka Rubena Sosy z Interu Milán.*) Významově tedy konektor *navíc* patří mezi tyto dva diskurzni argumenty, ačkoli syntakticky je součástí věty řídící⁴³ – srov. obrázek 11 zachycující příklad 98 v závislostním stromě. Na obrázku jsou diskurzni vztahy zachyceny oranžovou šipkou – první diskurzni vztah je značen mezi řídícími větami (reprezentovanými uzly *řekl* a *dodal*), druhý mezi závislými větami (tj. obsahy nepřímých řečí).

Více o uvedených slovesech říkání typu *dodat/dodávat, upřesnit/upřesňovat, doplnit/doplňovat, pokračovat* viz Rysová (2014).

⁴³ V angličtině pro tento jev používáme termín *connective raising* – srov. Rysová (2014).



Obrázek 11: Ukázka anotace nepřímých řečí v Pražském závislostním korpusu. Obrázek zachycuje souvětí: *Předseda klubu Augusto Cesar Lendoiro uvedl, že i bez obou Brazilců Deportivo na síle neztratí, navíc dodal, že na dobré cestě jsou jednání o přestupu uruguayského útočníka Rubena Sosy z Interu Milán.*

6.5 Sémantické typy diskurzních vztahů vyjádřené primárními a sekundárními konektory

Srovnání toho, jaké sémantické typy diskurzních vztahů vyjadřují primární a sekundární konektory⁴⁴, zobrazují tabulky 10 a 11. Tabulka 11 zachycuje zároveň rozložení jednotlivých typů diskurzních vztahů podle syntaktické povahy diskurzních argumentů, tj. zda jsou příslušné vztahy mezi dvěma slovesnými argumenty (v tabulce jsou označeny jako „VP_VP“) nebo zda je alespoň jeden z argumentů nominální (v tabulce jako „VP_NP“).

Typ vztahu	Primární konektory
Konjunkce	7 386
Opozice	3 171
Příčina a důsledek	2 583
Podmínka	1 351
Přípustka	874
Nesoučasnost	816
Konfrontace	632
Specifikace	625
Rektifikace	443
Gradace	439
Účel	412
Disjunktivní alternativa	261
Restriktivní opozice	266
Explikace	228
Současnost	225
Exemplifikace	147
Generalizace	100
Ekvivalence	99
Konjunktivní alternativa	88
Nepravý kontrast	50
Nepravá příčina a důsledek	41
Nepravá podmínka	16
Jiný vztah	2
Celkem	20 255

Tabulka 10: Typy diskurzních vztahů vyjádřené primárními konektory v Pražském závislostním korpusu

⁴⁴ Frekvenční seznam forem primárních a sekundárních konektorů z Pražského závislostního korpusu s ohledem na vyjadřované významové vztahy je uveden v příloze 1.5.

Typ vztahu	Sekundární konektory	VP_VP	%	VP_NP	%
Konjunkce	344	330	96	14	4
Příčina a důsledek	344	208	60	136	40
Podmínka	100	92	92	8	8
Explicace	67	52	78	15	22
Podmínka	44	43	98	1	2
Nesoučasnost	44	42	95	2	5
Konfrontace	34	32	94	2	6
Exemplifikace	30	9	30	21	70
Restriktivní opozice	28	6	21	22	79
Specifikace	24	14	58	10	42
Gradace	20	18	90	2	10
Generalizace	20	18	90	2	10
Opozice	13	13	100	0	0
Rektifikace	13	13	100	0	0
Ekvivalence	11	9	82	2	18
Disjunktivní alternativa	10	10	100	0	0
Účel	7	7	100	0	0
Nepravá příčina a důsledek	3	3	100	0	0
Konjunktivní alternativa	2	2	100	0	0
Současnost	1	1	100	0	0
Nepravá podmínka	1	1	100	0	0
Jiný vztah	1	1	100	0	0
Nepravý kontrast	0	0	0	0	0
Celkem	1 161	924	80	237	20

Tabulka 11: Typy diskurzních vztahů vyjádřené sekundárními konektory v Pražském závislostním korpusu

Z tabulek je patrné, že rozložení jednotlivých typů diskurzních vztahů⁴⁵ je u primárních a sekundárních konektorů velice podobné. Diskurzní vztahy s vysokým počtem výskytů (u primárních i sekundárních konektorů) jsou vztahy konjunkce, příčiny a důsledku, podmínky a výsledku podmínky, přípustky, nesoučasnosti (předčasnosti a následnosti) a konfrontace. Naopak vztahy s nejnižším (či velice nízkým) počtem jsou u obou typů konektorů tzv. nepravé vztahy (tj. vztahy pragmatické, založené na mimotextových jevech, zkušenosti účastníků komunikace apod.) – jedná se tedy o vztahy nepravé příčiny a důsledku, nepravé podmínky a výsledku podmínky a nepravý kontrast.

⁴⁵ Pro jednotlivé typy diskurzních vztahů jsou v PDT užity následující zkratky: synchr – současnost, preced – nesoučasnost (předčasnost a následnost), reason – příčina a důsledek, f_reason – nepravá příčina a důsledek, cond – podmínka a výsledek podmínky, f_cond – nepravá podmínka a výsledek podmínky, purp – účel, explicat – explicace, confr – konfrontace, opp – opozice, f_opp – nepravý kontrast, restr – restriktivní opozice, conc – přípustka, corr – rektifikace, grad – gradace, conj – konjunkce, exempl – exemplifikace, spec – specifikace, equiv – ekvivalence, gener – generalizace, conjalt – konjunktivní alternativa, disjalt – disjunktivní alternativa, other – jiný typ vztahu.

Primární a sekundární konektory se naopak liší ve vyjadřování vztahů opozice a explikace. Zatímco vztah opozice je u primárních konektorů na druhém místě (tj. s druhým nejvyšším počtem výskytů), u sekundárních konektorů se tento typ diskurzního vztahu objevuje velice zřídka (pouze ve 13 výskytech). Na druhou stranu vztah explikace je u sekundárních konektorů čtvrtým nejpočetnějším typem diskurzního vztahu, zatímco u konektorů primárních je až v druhé polovině tabulky (tj. mezi vztahy s nižším počtem výskytů).

Nižší čísla výskytů pragmatických vztahů ovšem podle našeho názoru neznamenají, že by se tyto vztahy v reálných textech vyskytovaly s tak nízkou frekvencí, ale spíše to, že jim v anotaci PDT nebyl zatím věnovaný dostatečný prostor (klasifikace diskurzních vztahů se prozatím zaměřovala spíše na vztahy sémantické). Dalším krokem při výzkumu koherence textu by tak mohl být pohled z vyšší perspektivy, z pohledu komunikačních strategií. Zatímco jsme do této doby postupovali od studia konkrétních jazykových prostředků směrem k jejich podílu na koherenci textu (tj. zaměřovali jsme se především na vyhledávání konektivních výrazů, na jejich lexikálně-syntaktickou a sémantickou charakteristiku apod.), v další fázi výzkumu by obohacující mohl být opačný pohled či postup z opačného směru – tj. od pragmatické funkce výsledného textu k jevům, které se na komunikačních strategiích bezprostředně podílejí nebo které si autor vybírá právě s ohledem na svůj komunikační záměr.

U některých konektivních prostředků se navíc pragmatický prvek ukazuje jako stěžejní (tj. výrazně se v nich odráží hodnocení a postoje mluvčího směrem k obsahu vyřčeného, komentáře ke způsobu formulování předchozího textu apod.). Na vyjadřování postojů a obecně na výstavbě argumentace v textu se zřetelně podílejí např. sekundární konektory obsahující slovesa smyslového vnímání a myšlení typu *stručně řečeno, jak je vidět, odůvodnit, argumentovat* apod. – srov. příklad 99:

(99) *Jsou již dávno pryč, ostatně jako další statisíce, které již nikdy své poslední slovo k problému neřeknou.*

Cynicky řečeno – vylétli komínem.

V uvedeném příkladě byl diskurzní vztah podle stávající klasifikace vyhodnocen jako ekvivalence (tj. obsah prvního argumentu je v druhém argumentu zopakován jinými slovy). Nicméně vedle prvku sémantického se v tomto případě zcela zřetelně uplatňuje prvek pragmatický (a to ještě výrazněji než u jiných, obdobných struktur typu *jinak řečeno* či *jinými*

slovy, které mají rovněž schopnost signalizovat sémantický vztah ekvivalence), tj. autor subjektivně hodnotí způsob formulace druhého argumentu. V současné době tedy v anotaci PDT není u výrazů typu *jinak řečeno*, *jinými slovy*, *cynicky řečeno* apod. pragmatický prvek blíže značen (ačkoli se vzájemně liší právě mírou subjektivnosti hodnocení mluvčího, srov. méně hodnotící spojení *obecně/krátce/stručně/jinak řečeno* vs. subjektivnější *cynicky/po pravdě/upřímně řečeno* apod.), tj. v PDT v současné době převládá zpracování sémantických diskurzních vztahů nad vztahy pragmatickými. Pragmatické aspekty koherence textu bude tedy zapotřebí v PDT propracovat detailněji (a to v rámci primárních i sekundárních konektorů).

V rámci sekundárních konektorů jsme dále zjišťovali, zda typ diskurzního vztahu souvisí se syntaktickou povahou diskurzních argumentů (tj. tím, zda jsou argumenty vyjádřeny slovesně či jmenně). Z výsledků vyplynulo, že téměř u všech typů vztahů převažují slovesné argumenty – u vztahu konjunkce a vztahu podmínky (tj. dvou ze tří nejpočetnějších vztahů) slovesné argumenty převažují z 96 % a z 92 %. Vztahy s alespoň jedním nominálním argumentem převažují pouze ve dvou případech – ve vztahu restriktivní opozice (79 %) a exemplifikace (70 %). Diskurzní vztah specifikace a vztah příčiny a důsledku se mezi slovesnými argumenty a argumenty, z nichž je jeden nominalizovaný, vyskytují rovnoměrněji (srov. 58 % vs. 42 % u vztahu specifikace a 60 % vs. 40 % u vztahu příčiny a důsledku). U ostatních diskurzních typů převažují vztahy mezi slovesnými argumenty zcela výrazně.

Celkový počet výskytů sekundárních konektorů ve spojení s nominalizovaným diskurzním argumentem je 237. Je zajímavé, že 136 případů z tohoto počtu (tedy více než polovina) vyjadřuje vztah příčiny a důsledku. Počet výskytů tohoto typu vztahu ve spojení s nominalizovaným argumentem tak značně převyšuje ostatní (druhým nejpočetnějším typem vztahu ve spojení s nominalizovaným argumentem je exemplifikace s 21 výskyty). Vysvětlení může spočívat v tom, že vztah příčiny a důsledku je v PDT chápán v širším smyslu, tj. nerozlišují se zde podtypy, jako je např. příčina, důvod, důsledek apod. Nicméně u vztahů mezi dvěma slovesnými argumenty vztah příčiny a důsledku tímto způsobem nedominoje (tvoří přibližně pětinu z celkového počtu diskurzních vztahů mezi slovesnými argumenty, má 208 výskytů mezi 924 případy).

V rámci analýzy sémantických typů diskurzních vztahů v Pražském závislostním korpusu jsme dále zjišťovali, jaké je u jednotlivých typů vztahů procentuální zastoupení primárních a sekundárních konektorů. Výsledky zachycuje tabulka 12.

Z tabulky je patrné, že u všech typů vztahů zřetelně převažují primární konektory (většinou přesahují 90 % výskytů). Poněkud vyšší procentuální zastoupení (v rámci sekundárních konektorů s počtem výskytů přesahujícím 10) mají pouze tři typy vztahů: explikace (22,7 %), exemplifikace (16,9 %) a generalizace (16,7 %). Jejich poměr je mezi primárními a sekundárními konektory vyrovnanější než u všech ostatních typů vztahů, i když i v těchto případech zastoupení primárních konektorů výrazně převažuje.

Typ vztahu	Celkem	Primární konektory %	Sekundární konektory %
Konjunkce	7 730	95,5	4,5
Opozice	3 184	99,6	0,4
Příčina a důsledek	2 927	91,4	8,6
Podmínka	1 451	93,1	6,9
Přípustka	918	95,2	4,8
Nesoučasnost	860	94,9	5,1
Konfrontace	666	94,9	5,1
Specifikace	649	96,3	3,7
Gradace	459	95,6	4,4
Rektifikace	456	97,1	2,9
Účel	419	98,3	1,7
Explikace	295	77,3	22,7
Restriktivní opozice	294	90,5	9,5
Disjunktivní alternativa	271	96,3	3,7
Současnost	226	99,6	0,4
Exemplifikace	177	83,1	16,9
Generalizace	120	83,3	16,7
Ekvivalence	110	90	10
Konjunktivní alternativa	90	97,8	2,2
Nepřavý kontrast	50	100	0
Nepřavá příčina a důsledek	44	93,2	6,8
Nepřavá podmínka	17	94,1	5,9
Jiný vztah ⁴⁶	3	66,7	33,3
Celkem	21 416	94,6	5,4

Tabulka 12: Typy diskurzních vztahů vyjádřené primárními a sekundárními konektory v Pražském závislostním korpusu – procentuální vyjádření

⁴⁶ Jako jiný typ vztahu byly zachyceny případy, ve kterých byl konektor (primární či sekundární) užit na nepatřičném místě, tj. významově se na dané místo nehodí – srov. např. *Je nebezpečné dělat rovnítko mezi bezdomovci a kriminálními živly, avšak opuštěnost a osamělost ve spodních vrstvách společnosti chodí ruku v ruce s kriminalitou. Koho opustila rodina, kdo nemá přátele a zaměstnání, s tím to jde z kopce. Příklady můžeme nalézt v blízkosti tržišť: bezdomovci pomáhají trhovcům, sbírají staré lahve na stadionech a odnášejí do sběru papír.* V daném příkladě spojení *příklady můžeme nalézt* nevyjadřuje vztah exemplifikace, tj. to, že bezdomovci pomáhají trhovcům apod., není příkladem toho, že koho opustila rodina atd., s tím to jde z kopce.

6.5.1 Struktury typu *výjimkou je, příkladem je*

Jak jsme popsali výše, výsledky výzkumu ukázaly, že u vztahů restriktivní opozice a exemplifikace vyjádřených sekundárními konektory převládají výskyty s nominalizací druhého diskurzního argumentu nad výskyty mezi dvěma argumenty slovesnými. Sekundární konektory vyjadřující vztah restriktivní opozice v PDT jsou *výjimku tvoří, výjimkou je, výjimkou zůstává, s výjimkou*. Sekundární konektory vyjadřující v PDT vztah exemplifikace jsou *příkladem je, příklad pochází z, jako příklad uvedl, za příklad může posloužit, podařilo se získat několik příkladů, dokumentuje to na příkladu* – srov. příklad 100 z PDT:

- (100) *Na rozdíl od jiných zachraňujících se mužstev Západočeši výrazně neposilovali.*
Výjimkou je příchod nadějného záložníka Plocka z Pardubic, jehož manažer Jaromír Šeterle vybojoval v konkurenci dalších zájemců – Liberce a Hradce Králové.

V uvedeném příkladě sekundární konektor *výjimkou je* vyjadřuje diskurzní vztah restriktivní opozice mezi prvním slovesným a druhým jmenným diskurzním argumentem. Tento typ výskytu v Pražském závislostním korpusu převažuje nad případy se sekundárními konektory mezi dvěma slovesnými argumenty.

Spojení typu *výjimkou je, výjimku tvoří, příkladem je* apod. (tj. se strukturou: podstatné jméno signalizující příslušný typ vztahu + sloveso s oslabeným lexikálním významem typu *být*) zcela jednoznačně preferují jmenné vyjádření druhého diskurzního argumentu (tj. dominuje případ *výjimkou je příchod nadějného záložníka* nad *výjimkou je to, že přišel nadějný záložník*).

Dále jsme tedy zjišťovali, jak se v reálných textech chovají další struktury tohoto typu (vyjadřující jiné sémantické typy diskurzních vztahů). Vyhledali jsme proto v PDT všechny sekundární konektory (zachycené v současné anotaci), jejichž struktura odpovídá obecnému schématu „podstatné jméno signalizující typ diskurzního vztahu + sloveso *být*“ (např. *podmínkou je, důvodem je, příčinou je, výsledkem je*) a zjišťovali jsme, zda vyjadřují spíše vztahy mezi dvěma slovesnými argumenty či preferují nominalizaci druhého argumentu – srov. tabulka 13 (v tabulce uvádíme počet konkrétních realizací jednotlivých typů sekundárních konektorů, tj. například typ *výjimkou je* se v PDT objevil v různých obměnách – srov. např. *výjimkou je, výjimkou byl, výjimkou jsou* apod.).

Typ sekundárního konektoru	VP_VP	VP_NP	Celkem
<i>důvodem je</i>	8	21	29
<i>příkladem je</i>	0	12	12
<i>podmínkou je</i>	5	6	11
<i>výsledkem je</i>	2	8	10
<i>příčinou je</i>	1	8	9
<i>výjimkou je</i>	0	8	8
<i>důsledkem je</i>	0	3	3
Celkem	16	66	82

Tabulka 13: Sekundární konektory typu *výjimkou je* v Pražském závislostním korpusu

Z tabulky je patrné, že všechny zkoumané typy sekundárních konektorů upřednostňují nominalizaci druhého diskurzivního argumentu (tj. např. *důsledkem je stárnutí obyvatelstva* než *důsledkem je, že obyvatelstvo stárne*), a to celkově v 80 % výskytů⁴⁷. Tato malá sonda ukazuje, že **struktura sekundárních konektorů může mít přímý vliv na syntaktickou realizaci druhého diskurzivního argumentu.**

⁴⁷ Počty nalezených struktur jsou ale poměrně nízké.

6.6 Rozšíření klasifikace diskurzních vztahů vyjádřených sekundárními konektory

Jak jsme popsali v kapitole 3.5.2.1, klasifikace sémantických typů diskurzních vztahů pro sekundární konektory byla převzata z anotace primárních konektorů (vzhledem k tomu, že náš výzkum proběhl na datech Pražského závislostního korpusu, který anotaci primárních konektorů již obsahoval).

Anotátoři sekundárních konektorů dostali zároveň pokyn značit v poznámkách „nové“ typy vztahů (tj. takové typy, které nebyly v nabídce původní klasifikace diskurzních vztahů pro anotaci primárních konektorů). Anotátorské poznámky tímto způsobem upozornily na tři nové typy vztahů – vyplývání (či vyvození výsledků), vyvození závěru a zřetel.

6.6.1 Vztah vyplývání (či vyvození výsledků)

Vztah vyplývání (vyvození výsledků) byl v datech PDT vyjádřen následujícími sekundárními konektory: *z toho vyplývá, že; v každém případě; výsledkem bylo; výsledkem je* – srov. příklad 101 z PDT:

(101) *U nás na škole vysoké se zase mocně krade.*

Pro muzeálnost vybavení univerzity spíš jen nehlídané peněženky nebo odložené fotoaparáty; rozpadající se staré typy počítačů, pokud vůbec nějaké jsou, za to nestojí. Ohlásí se například mladistvý umývač oken s kbelíkem v ruce, a než mu posloužíte smetáčkem na prach, sám je v prachu a s ním něco cenného.

Zavolaná policie odmítne příchod na místo do té doby, než někdo doběhne na náš centrální úřad a opatří pro ni povolení ke vstupu na akademickou půdu.

Nevím o tom, že by se už někdy něco s úspěchem vyšetřilo.

Výsledkem je doporučení – dejte si pokud možno všude pancéřové dveře a klíče od nich do nezczitelného trezoru.

V uvedeném příkladě je sekundárním konektorem spojení *výsledkem je* vztahující se k celému předchozímu odstavci. I v ostatních nalezených případech se většinou jedná o vztah mezi delšími úseky textu (např. celými odstavci).

6.6.2 Vyvození závěru

Dalším typem nového vztahu bylo vyvození závěru – jednalo se o příklady se sekundárními konektory *závěr zněl, že; došel k závěru, že; jednoznačný závěr je, že; dospělo k tomuto závěru* – srov. příklad 102:

- (102) *Umělecký restaurátor a badatel de Wilt vzbudil v uměleckém světě senzaci, když využil rentgenové záření k posouzení stáří obrazu.*
Jeho závěr zněl, že soubor je starý minimálně tři desetiletí a tudíž pravý.
Neboť jaký blázen by měl zájem kolem roku 1895 padělat holandského umělce, jehož plátna byla k dostání po dvou stech francích?
Obžaloba si též najala znalce, který pracoval s rentgenem.
Ten prozkoumal kolekci z hlediska malířského rukopisu a došel k závěru, že ji nenamaloval van Gogh.

Vztah vyvození závěru se v PDT často objevuje opět mezi většími úseky textu.

6.6.3 Zřetel

Nově byl mezi typy diskurzních vztahů, které se objevují v PDT, zařazen také vztah zřetele. Může být vyjádřen například sekundárními konektory *v tomto směru, se zřetelem k tomu* – srov. příklad 103:

- (103) *Přestože Smlouva o hraničním dokumentárním díle nemění průběh státních hranic, upravuje vymezení státního území České republiky.*
Se zřetelem k tomu je třeba ji pokládat za smluvní úpravu, která bezprostředně souvisí se suverenitou státu.

Pro další anotace diskurzních vztahů v PDT bude stávající klasifikace o dané tři typy významových vztahů, tj. vyplývání (či vyvození výsledku), vyvození závěru a zřetel, rozšířena.

6.7 Diskurzní vztahy mezivětné a vnitrovětné

Jak již bylo řečeno výše, Pražský závislostní korpus obsahuje anotaci diskurzních vztahů mezivětných i vnitrovětných (tj. např. *Mám rýmu. Z tohoto důvodu nejdu do školy.* i *Mám rýmu, z tohoto důvodu nejdu do školy.*).

Zjišťovali jsme tedy, jaký je poměr mezivětných a vnitrovětných vztahů vyjádřených primárními a sekundárními konektory. Cílem bylo ověřit, zda některé typy diskurzních vztahů tíhnou spíše k mezivětnému či vnitrovětnému vyjádření nebo zda sekundární konektory převažují u diskurzních vztahů realizovaných mezivětně či uvnitř věty.

Tabulka 14 zobrazuje podíl mezivětných a vnitrovětných vztahů vyjádřených primárními konektory v Pražském závislostním korpusu. Pro lepší přehlednost je v tabulce pro mezivětné vztahy použita zkratka „mezi“, pro vnitrovětné vztahy zkratka „vnitro“. Jednotlivé typy diskurzních vztahů jsou v tabulce seřazeny podle četnosti.

Z tabulky je patrné, že u diskurzních vztahů vyjádřených primárními konektory převažují vztahy vnitrovětné nad mezivětnými (srov. 70 % vs. 30 %). U některých z nich (srov. vztahy podmínky a výsledku podmínky, účelu, nepravé podmínky a výsledku podmínky, disjunktivní alternativy) se vnitrovětné vyjádření blíží téměř 100 %. Mezivětné vztahy výrazně převažují pouze u vztahů generalizace (91 %) a exemplifikace (82 %).

Primární konektory					
Typ vztahu	Mezi	%	Vnitro	%	Celkem
Konjunkce	1 341	18	6 045	82	7 386
Opozice	1 784	56	1 387	44	3 171
Příčina a důsledek	1 013	39	1 570	61	2 583
Podmínka	18	1	1 333	99	1 351
Přípustka	258	30	616	70	874
Nesoučasnost	230	28	586	72	816
Konfrontace	286	45	346	55	632
Specifikace	107	17	518	83	625
Rektifikace	124	28	319	72	443
Gradace	198	45	241	55	439
Účel	1	0	411	100	412
Disjunktivní alternativa	5	2	256	98	261
Restriktivní opozice	172	65	94	35	266
Explicace	128	56	100	44	228
Současnost	51	23	174	77	225
Exemplifikace	120	82	27	18	147
Generalizace	91	91	9	9	100
Ekvivalence	58	59	41	41	99
Konjunktivní alternativa	18	20	70	80	88
Nepřavý kontrast	27	54	23	46	50
Nepřavá příčina a důsledek	28	68	13	32	41
Nepřavá podmínka	1	6	15	94	16
Jiný vztah	1	50	1	50	2
Celkem	6 060	30	14 195	70	20 255

Tabulka 14: Podíl mezivětných a vnitrovětných vztahů vyjádřených primárními konektory v Pražském závislostním korpusu

Tabulka 15 zobrazuje mezivětné a vnitrovětné vztahy vyjádřené sekundárními konektory. Z tabulky je zřejmé, že u sekundárních konektorů je situace opačná než u konektorů primárních, tj. celkově u nich převažují mezivětně realizované diskurzí vztahy nad vnitrovětnými (srov. 63 % vs. 37 %).

Sekundární konektory					
Typ vztahu	Mezi	%	Vnitro	%	Celkem
Konjunkce	229	67	115	33	344
Příčina a důsledek	219	64	125	36	344
Podmínka	29	29	71	71	100
Explicace	46	69	21	31	67
Přípustka	26	59	18	41	44
Nesoučasnost	20	45	24	55	44
Konfrontace	27	79	7	21	34
Exemplifikace	24	80	6	20	30
Restriktivní opozice	19	68	9	32	28
Specifikace	18	75	6	25	24
Gradace	17	85	3	15	20
Generalizace	18	90	2	10	20
Opozice	12	92	1	8	13
Rektifikace	6	46	7	54	13
Ekvivalence	9	82	2	18	11
Disjunktivní alternativa	0	0	10	100	10
Účel	6	86	1	14	7
Nepravá příčina a důsledek	1	33	2	67	3
Konjunktivní alternativa	1	50	1	50	2
Současnost	1	100	0	0	1
Nepravá podmínka	0	0	1	100	1
Jiný vztah	1	100	0	0	1
Nepravý kontrast	0	0	0	0	0
Celkem	729	63	432	37	1 161

Tabulka 15: Podíl mezivětných a vnitrovětných vztahů vyjádřených sekundárními konektory v Pražském závislostním korpusu

V rámci sekundárních konektorů jsme dále zjišťovali, zda na způsob vyjadřování diskurzních vztahů (tj. mezivětně či v rámci věty) má vliv také syntaktická povaha diskurzních argumentů, tedy zda se podíl mezivětných a vnitrovětných vztahů liší u vztahů mezi dvěma slovesnými argumenty (v tabulce 16 pod zkratkou „VP_VP“) a vztahů, u nichž je alespoň jeden z diskurzních argumentů nominální (v tabulce 17 „VP_NP“).

Sekundární konektory: Vztahy VP_VP					
Typ vztahu	Mezi	%	Vnitro	%	Celkem
Konjunkce	220	67	110	33	330
Příčina a důsledek	123	59	85	41	208
Podmínka	25	27	67	73	92
Explikace	36	69	16	31	52
Přípustka	25	58	18	42	43
Nesoučasnost	19	45	23	55	42
Konfrontace	26	81	6	19	32
Gradace	15	83	3	17	18
Generalizace	17	94	1	6	18
Specifikace	10	71	4	29	14
Opozice	12	92	1	8	13
Rektifikace	6	46	7	54	13
Disjunktivní alternativa	0	0	10	100	10
Exemplifikace	9	100	0	0	9
Ekvivalence	8	89	1	11	9
Účel	6	86	1	14	7
Restriktivní opozice	1	17	5	83	6
Nepravá příčina a důsledek	1	33	2	67	3
Konjunktivní alternativa	1	50	1	50	2
Současnost	1	100	0	0	1
Nepravá podmínka	0	0	1	100	1
Jiný typ	1	100	0	0	1
Nepravý kontrast	0	0	0	100	0
Celkem	562	61	362	39	924

Tabulka 16: Podíl mezivětných a vnitrovětných vztahů vyjádřených sekundárními konektory mezi dvěma slovesnými argumenty v Pražském závislostním korpusu

Sekundární konektory: Vztahy VP_NP					
Typ vztahu	Mezi	%	Vnitro	%	Celkem
Příčina a důsledek	96	71	40	29	136
Restriktivní opozice	18	82	4	18	22
Exemplifikace	15	71	6	29	21
Explikace	10	67	5	33	15
Konjunkce	9	64	5	36	14
Specifikace	8	80	2	20	10
Podmínka	4	50	4	50	8
Konfrontace	1	50	1	50	2
Ekvivalence	1	50	1	50	2
Generalizace	1	50	1	50	2
Gradace	2	100	0	0	2
Nesoučasnost	1	50	1	50	2
Přípustka	1	100	0	0	1
Celkem	167	70	70	30	237

Tabulka 17: Podíl mezivětných a vnitrovětných vztahů vyjádřených sekundárními konektory s alespoň jedním nominálním argumentem v Pražském závislostním korpusu

Z tabulek 16 a 17 je patrné, že v obou případech převažují vztahy mezivětné nad vnitrovětnými. U slovesných argumentů celkově v 61 %, u diskurzivních vztahů s alespoň jedním nominálním argumentem v 70 %. Výsledky obou analýz jsou tedy velmi podobné.

Zajímavé je, že u vztahu restriktivní opozice celkově (tj. bez rozlišení povahy diskurzivních argumentů – viz tabulka 15) převažuje mezivětné vyjádření právě díky vztahům s nominálním argumentem. U vztahů mezi slovesnými argumenty totiž u restriktivní opozice převládá vyjádření vnitrovětné (83 %). Celkový počet výskytů tohoto typu vztahu je ovšem v PDT poměrně nízký.

Srovnání podílu mezivětných a vnitrovětných diskurzivních vztahů vyjádřených primárními a sekundárními konektory zobrazuje tabulka 18. Z tabulky je zřejmé, že celkově v Pražském závislostním korpusu převažují vnitrovětné vztahy nad vztahy mezivětnými (vyjádřenými primárními a sekundárními konektory). Ačkoli tedy sekundární konektory tíhnou spíše k vyjádření mezivětnému (celkově ze 63 % – viz tabulka 15), jejich celkově nízký počet se na podílu mezivětných a vnitrovětných vztahů projevuje pouze nepatrně.

Z tabulky je dále patrné, že u mezivětných i vnitrovětných vztahů celkově převažuje vyjádření primárními konektory, procento mezivětných vztahů je ovšem u sekundárních konektorů vyšší než procento vztahů vnitrovětných (srov. 11 % vs. 3 %). V rámci

jednotlivých typů diskurzních vztahů převládají u vnitrovětného vyjádření primární konektory zcela zřetelně v podstatě ve všech případech.

Také u mezivětných vztahů většinou převažuje užití primárních konektorů, v jednom případě (s relevantním počtem výskytů – 100) je tomu ovšem naopak. U mezivětného vyjádření diskurzního vztahu podmínky⁴⁸ převládá užití konektorů sekundárních (v 62 %). Zajímavé je, že obecně je tento diskurzní vztah vyjadřován většinou vnitrovětně (tj. v 97 %). Pokud je ovšem vyjádřen mezivětně, je vyšší pravděpodobnost, že jako konektivní prostředek bude použit sekundární konektor (tj. např. *podmínkou je*).

⁴⁸ Užití sekundárního konektoru převažuje také u mezivětného vztahu účelu. Vzhledem k nízkému počtu výskytů (sekundární konektor pro vztah účelu se objevil pouze v 7 výskytech) ovšem v tomto případě nemůžeme dělat obecné závěry.

Typ vztahu	Mezi %	Primární %	Sekundární %	Vnitro %	Primární %	Sekundární %	Celkem
Konjunkce	20	85	15	80	98	2	7 730
Opozice	56	99	1	44	99,9	0,1	3 184
Příčina a důsledek	42	82	18	58	93	7	2 927
Podmínka	3	38	62	97	95	5	1 451
Přípustka	31	91	9	69	97	3	918
Nesoučasnost	30	92	8	70	96	4	860
Konfrontace	47	91	9	53	98	2	666
Specifikace	19	86	14	81	99	1	649
Rektifikace	28	95	5	72	98	2	459
Gradace	47	92	8	53	99	1	456
Účel	2	14	86	98	99,8	0,2	419
Disjunktivní alternativa	2	100	0	98	96	4	295
Restriktivní opozice	65	90	10	35	91	9	294
Explikace	64	74	26	36	83	17	271
Současnost	23	98	2	77	100	0	226
Exemplifikace	81	83	17	19	82	18	177
Ekvivalence	56	87	13	44	95	5	120
Generalizace	99	83	17	1	82	18	110
Konjunktivní alternativa	21	95	5	79	99	1	90
Nepravý kontrast	54	100	0	46	100	0	50
Nepravá příčina a důsledek	66	97	3	34	87	13	44
Nepravá podmínka	6	100	0	94	94	6	17
Jiný vztah	67	50	50	33	100	0	3
Celkem	32	89	11	68	97	3	21 416

Tabulka 18: Podíl mezivětných a vnitrovětných vztahů vyjádřených primárními a sekundárními konektory v Pražském závislostním korpusu

Z porovnání mezivětných a vnitrovětných vztahů vyjádřených primárními a sekundárními konektory tedy vyplývají následující závěry.

1) Primární konektory vyjadřují častěji vnitrovětné diskurzivní vztahy než vztahy mezivětné (srov. 70 % vs. 30 %), naopak sekundární konektory vyjadřují častěji vztahy mezivětné (srov. 63 % vs. 37 %).

Pokud je tedy diskurzí vztah vyjádřen například spojkou *a*, je její častější užití *Mám rýmu a nepůjdu do školy.* než *Mám rýmu. A nepůjdu do školy.* Naopak pokud je diskurzí vztah vyjádřen například spojením *z tohoto důvodu*, je častěji užito ve sledu vět typu *Mám rýmu. Z tohoto důvodu nepůjdu do školy.* než v souvětí typu *Mám rýmu, z tohoto důvodu nepůjdu do školy.*

Obecně (v absolutních číslech) jsou ale diskurzí vztahy vnitrovětné i mezivětné častěji vyjadřovány konektory primárními. Existují však výjimky, viz následující bod 2).

2) Celkově jsou mezivětné i vnitrovětné diskurzí vztahy vyjadřovány především primárními konektory (celkový počet primárních konektorů v PDT je více než 17krát vyšší než celkový počet sekundárních konektorů). I přesto pro mezivětné vyjádření vztahu podmínky dominuje podle dat PDT užití sekundárního konektoru (v absolutních počtech výskytů).

Podle PDT se tak například častěji užívá *Půjdu ven. V případě, že nebude pršet.* než *Půjdu ven. Jestliže nebude pršet.*

3) U některých typů diskurzí vztahů vyjádřených sekundárními konektory může mít vliv na to, zda budou realizovány mezivětně či vnitrovětně, také syntaktická povaha daných diskurzí argumentů. U vztahu restriktivní opozice (signalizované např. spojeními *výjimkou je* či *výjimku tvoří*) mezi dvěma slovesnými argumenty převažuje vyjádření vnitrovětné nad mezivětným. U vztahů restriktivní opozice s jedním nominálním diskurzím argumentem naopak převládá vyjádření mezivětné. (Ačkoli počet nalezených výskytů v PDT je v tomto případě poměrně nízký, popisovaný jev je tedy třeba chápat pouze jako možnou tendenci.)

Podle PDT se tak například častěji užívá *Fotbalové mužstvo nedostalo žádné nové posily. Výjimkou je příchod nadějného záložníka.* než *Fotbalové mužstvo nedostalo žádné nové posily, výjimkou je příchod nadějného záložníka.*

Ale zároveň se podle PDT častěji používá například *Fotbalové mužstvo nedostalo žádné nové posily, výjimkou je, že přišel nadějný záložník.* než *Fotbalové mužstvo nedostalo žádné nové posily. Výjimkou je, že přišel nadějný záložník.*

6.8 Vybrané sekundární konektory v psané a mluvené češtině

V rámci studia sekundárních konektorů jsme provedli dílčí průzkum zaměřený na výskyt a chování těchto jazykových prostředků v psané a mluvené češtině. Využili jsme přitom data Pražského závislostního korpusu (pro psanou češtinu) a data korpusu mluvené češtiny DIALOG. Oba korpusy obsahují srovnatelný počet textových slov (DIALOG 932 373, PDT 833 195) a texty pocházejí z oblasti publicistiky.

Jako zástupce sekundárních konektorů jsme vybrali takové struktury, které obsahují jádrová slova *důvod* (tj. např. spojení typu *to je důvod, proč*), *příčina* (např. *příčinou bylo*), *z/odůvodnit(ňovat)* (např. *zdůvodnil to tím, že*), *souviset* (např. *s tím souvisí*).

Zkoumali jsme přitom, v kolika procentech z celkových výskytů se v obou korpusech daná slova objevují ve strukturách s konektivní platností (tj. ve funkci sekundárních konektorů a volných konektivních frází, srov. příklad 104 z korpusu DIALOG) a v kolika procentech bez konektivní platnosti (srov. příklad 105 z korpusu DIALOG). Posléze jsme se zaměřili na to, jak se tyto konektivní struktury v reálných textech chovají (zda si např. vynucují nominalizaci druhého diskurzního argumentu apod.).

- (104) *To, co jste teď řekl, je podle mého názoru nesprávné z toho prostého důvodu, že reálná mzda, jak oba víme, je poměr nominální mzdy a míry inflace.*

V příkladě 104 (z mluveného korpusu DIALOG) je diskurzní vztah příčiny a důsledku vyjádřený spojením *z toho prostého důvodu, že* (nahraditelným v tomto případě za rozvitý primární konektor *prostě proto, že*). V následujícím příkladě 105 se naopak slovo *důvod* nevyskytuje v konektivní struktuře, tj. spojení *z důvodu povětrnostních podmínek* (které neodkazuje do předchozího kontextu) nemá v textu spojovací funkci.

- (105) *[Na Šumavě, M. R.] opravdu hrozilo, že by žáci nemohli dojíždět do jiné školy z důvodu povětrnostních podmínek.*

Rozložení vybraných jádrových slov (tj. slov *důvod*, *příčina*, *z/odůvodnit(ňovat)*, *souviset*) ve strukturách s konektivní či nekonektivní funkcí v Pražském závislostním korpusu a v korpusu DIALOG zobrazuje tabulka 19.

Struktury obsahující slovo	S konektivní funkcí		%		Celkový počet výskytů	
	PDT	DIALOG	PDT	DIALOG	PDT	DIALOG
<i>důvod</i>	180	119	55 %	28 %	330	425
<i>příčina</i>	32	14	36 %	21 %	89	66
<i>z/odůvodnit(ňovat)</i>	20	5	51 %	15 %	39	34
<i>souviset</i>	20	42	34 %	28 %	58	148
Celkem	252	180	49 %	27 %	516	673

Tabulka 19: Struktury obsahující vybraná jádrová slova v Pražském závislostním korpusu a korpusu DIALOG

Z tabulky je patrné, že vyšší počet zkoumaných struktur s konektivní funkcí se vyskytuje v Pražském závislostním korpusu (tj. v psaných textech).

U všech vybraných slov vyskytujících se v konektivních strukturách jsme dále provedli podrobnější analýzu. Zkoumali jsme, zda mají dané struktury funkci univerzálního sekundárního konektoru či (neuniverzální) volné konektivní fráze, a zároveň to, zda vyjadřují diskurzí vztah mezi dvěma slovesnými argumenty (v tabulce jsou označeny jako „VP_VP“) či zda je druhý argument nominální (v tabulce „VP_NP“) – srov. tabulka 20 pro jádrové slovo *důvod* v korpusech PDT a DIALOG:

<i>Důvod</i>	VP_VP		VP_NP		Celkem	
	PDT	DIALOG	PDT	DIALOG	PDT	DIALOG
Univerzální sekundární konektory	76	106	43	11	119 (66 %)	117 (98 %)
Neuniverzální volné konektivní fráze	16	0	45	2	61 (34 %)	2 (2 %)
Celkem	92 (51 %)	106 (89 %)	88 (49 %)	13 (11 %)	180 (100 %)	119 (100 %)

Tabulka 20: Konektivní struktury se slovem *důvod* v korpusech PDT a DIALOG

Z porovnání syntaktické realizace diskurzí argumentů vyplývá, že korpus DIALOG obsahuje 89 % konektivních struktur se slovem *důvod* vyjadřujících diskurzí vztah mezi dvěma slovesnými argumenty a pouze 11 % vztahů s nominálním argumentem. V korpusu PDT je situace naopak velice vyrovnaná – srov. 51 % vztahů je mezi slovesnými argumenty, 49 % vztahů se vyskytuje s nominálním argumentem. Z výsledků je tedy patrné, že **v psané češtině je u daných struktur vyšší tendence k nominalizaci (či kondenzaci) diskurzí argumentů než v mluvené češtině** (ve které se nominalizace argumentu objevuje zřídka).

Jinými slovy, v mluvené češtině se častěji (v 89 %) vyskytuje typ příkladu *důvodem je, že obyvateľstvo stárne* než *důvodem je stárnutí obyvateľstva*.

Z porovnání univerzálnosti konektivních struktur se slovem *důvod* vyplývá, že korpus DIALOG obsahuje téměř výhradně univerzální konektivní struktury, tedy sekundární konektory (srov. 98 % univerzálních vs. 2 % neuniverzálních struktur). V korpusu PDT je poměr zkoumaných struktur opět více vyvážený, i když i zde zřetelně převažují struktury univerzální (tj. sekundární konektory nad volnými konektivními frázemi – srov. 66 % vs. 34 %).

Výsledky zároveň ukázaly, že v psané češtině se vyskytují struktury typu *důvodem pro propuštění bylo, důvodem poklesu je, toto počínání je důvodem, proč* apod. častěji než v textech mluvené češtiny, ve kterých se na místě výrazu anaforicky odkazujícího k předchozímu kontextu typicky (tj. z 98 % u struktur se slovem *důvod*) užívá zájmeno *to* – srov. např. *důvodem toho je* či *to je důvod, proč*. **Pro psanou češtinu je tedy příznačnější větší lexikální rozmanitost konektivních struktur, zatímco v mluvené češtině jsou zcela zřetelné tendence k ustalování daných struktur na jedné určité formě** (více k rozmanitosti konektivních struktur obsahujících slovo *důvod* viz Rysová, 2013c).

Lexikálně rozmanitější jsou v PDT také konkrétní varianty konektivních struktur se slovem *příčina*. Zatímco v obou korpusech se objevují příklady typu *příčinou je to, že; příčina je jednoznačná; příčina byla, že* apod., tj. struktury se slovesem *být*, PDT obsahuje i struktury s jinými typy sloves – např. *vidět příčinu v tom, že; hledat příčinu v tom, že; označit za příčinu; příčiny tkví v tom, že* apod.

Podrobnou analýzu týkající se univerzálnosti konektivních struktur a syntaktické povahy argumentů jsme provedli také u ostatních vybraných slov (tj. *příčina, z/odůvodnit(ňovat), souviset*), nicméně počet jejich výskytů je v obou korpusech poměrně nízký (např. v korpusu DIALOG se konektivní *z/odůvodnit(ňovat)* objevilo pouze v pěti výskytech). Z tohoto důvodu uvádíme výsledky našeho výzkumu v jedné shrnující tabulce pro všechna vybraná slova zároveň (tj. pro slova *důvod, příčina, z/odůvodnit(ňovat), souviset*) – srov. tabulka 21.

Typ struktury	VP_VP		VP_NP		Celkem	
	PDT	DIALOG	PDT	DIALOG	PDT	DIALOG
Univerzální sekundární konektory	84	122	72	51	156 (62 %)	173 (96 %)
Neuniverzální volné konektivní fráze	33	2	63	5	96 (38 %)	7 (4 %)
Celkem	117 (46 %)	124 (69 %)	135 (54 %)	56 (31 %)	252 (100 %)	180 (100 %)

Tabulka 21: Shrnující výsledky pro vybrané konektivní struktury v korpusech PDT a DIALOG

Výsledky analýzy dále poukázaly na skutečnost, že vybrané konektivní struktury se v korpusech PDT a DIALOG liší také tím, jaké významové vztahy v textu signalizují, resp. liší se v rozmanitosti významových vztahů. Např. konektivní struktury se slovesem *souviset* slouží v korpusu DIALOG zpravidla k prostému navazování dvou různých témat, zatímco v PDT vyjadřují pestřejší škálu významových vztahů.

Jinými slovy, v korpusu DIALOG se struktury *s tím souvisí, to souvisí s apod.* používají většinou k navození zdánlivě větší plynulosti či návaznosti mluveného textu, význam diskurzního vztahu je v těchto případech spíše zastřený, vágní – srov. příklady *s tím souvisí i druhá věc, to souvisí ještě s druhým a posledním problémem, to trochu souvisí s těmi diskuzemi minulých týdnů, to souvisí s celou řadou dalších otázek, souvisí to znovu s tím druhým tématem, to souvisí eště s jednou věcí, s tím souvisí třetí problém* apod. Z celkového počtu (tj. 42 výskytů – srov. tabulka 19) v korpusu DIALOG se v této funkci objevilo 38 příkladů (tj. 90 %), v PDT pouze 5 (tj. 25 % z celkového počtu 20 výskytů).

Zbylé výskyty (v PDT převažující) slouží ke zřetelnějšímu navazování jednotek textu, tj. mezi danými diskurzními argumenty je „pevnější“ významový vztah než prosté (v některých případech až umělé) napojování dalších částí textu – srov. příklady 106 a 107 z PDT:

- (106) *Těší nás, že se lidé začínají zajímat o své zdraví, výsledky vyšetření však nejsou dobré. Souvisí to s naší životosprávou, zejména s nadměrným požíváním tuků, omezeným pohybem atd., vysvětluje vedoucí oddělení hygieny výživy KHS Č. Budějovice Peter Horan.*

(107) *Za samozřejmost lze už považovat to, že většina vystavovatelů nabízela jak tuzemské, tak i zahraniční značky v nejrůznějších provedení, vybavení i v různých cenových relacích.*

Souvisí to se snahou o získání všech potencionálních zákazníků, jejichž investiční zájmy i možnosti jsou však značně rozdílné.

V uvedených příkladech se objevuje sekundární konektor *souvisí to s* vyjadřující diskurzni vztah mezi prvním slovesným a druhým jmenným argumentem. V obou případech se význam diskurzniho vztahu blíží vztahu příčiny a důsledku (srov. parafráze *...výsledky vyšetření však nejsou dobré, protože máme špatnou životosprávu, zejména nadměrně používáme tuky... a většina vystavovatelů nabízela jak tuzemské, tak i zahraniční značky v nejrůznějších provedení..., protože se snaží získat všechny potencionální zákazníky...*).

Z příkladů je patrné, že struktury *souvisí to s* neslouží v těchto případech k prostému navázání dvou různých témat, význam diskurzniho vztahu je zde zřetelný (v dalších příkladech z PDT se kromě vztahu příčiny a důsledku blíží např. také explikaci). V PDT je tento typ výrazně častější (tvoří 75 % z celkového počtu vztahů vyjádřených konektivními strukturami se slovesným základem *souviset*). V korpusu DIALOG se naopak vyskytuje zřídka (pouze v 10 %).

Na základě analýzy vybraných konektivních struktur v Pražském závislostním korpusu (obsahujícím psané texty) a v korpusu mluvené češtiny DIALOG můžeme vyvodit následující obecnější závěry.

1) Počet vybraných struktur s konektivní funkcí je vyšší v Pražském závislostním korpusu, tj. v psaných textech (PDT obsahuje 252 výskytů daných struktur s konektivní funkcí, tj. 49 % z celkového počtu; DIALOG 180 výskytů, tj. 27 % z celkového počtu – viz tabulka 19). Vysvětlení může spočívat v tom, že některé konektivní struktury typu *toto počínání je důvodem, proč; své chování odůvodnil tím, že; důvodem poklesu je* apod. mohou působit formálněji (než například neutrální konektory, jako jsou *proto, protože* aj.), a proto se objevují častěji v psaných textech.

2) V psané češtině je obecně vyšší tendence k nominalizaci (či kondenzaci) druhého diskurzniho argumentu než v češtině mluvené. Tj. příklady typu *důvodem je jeho napadení*

ředitele (oproti slovesnému vyjádření *důvodem je, že napadl ředitele*) jsou častější v psaných textech.

3) V psané i mluvené češtině se častěji vyskytují univerzální konektivní struktury (tj. sekundární konektory) než neuniverzální (tj. volné konektivní fráze). V mluvené češtině je přitom rozdíl v počtu univerzálních a neuniverzálních struktur zcela jednoznačný – v průměru se v korpusu DIALOG vyskytuje 96 % univerzálních a 4 % neuniverzálních (viz tabulka 21) struktur (obsahujících zkoumaná slova), v korpusu PDT je poměr poněkud vyrovnanější – 62 % vs. 38 %. Obecně tedy můžeme říci, že konektivní struktury typu *důvodem pro propuštění bylo, příčinou poklesu je, toto počínání je důvodem, proč* jsou typičtější pro psané texty. Psaná čeština se tedy v tomto ohledu vyznačuje větší lexikální rozmanitostí než čeština mluvená.

4) Obdobně se konektivní struktury se slovem *příčina* vyskytují v korpusu DIALOG pouze ve spojení se slovesem *být* (tj. např. *hlavní příčinou je; příčina je jednoduchá; příčiny jsou nejrůznější; příčiny jsou ty, že* apod.), zatímco v PDT i s jinými slovesy (srov. *vidět příčinu v tom, že; hledat příčinu v tom, že; označit za příčinu*). Z tohoto hlediska je tedy spektrum konektivních struktur v psaných textech rozmanitější než v textech mluvených.

5) Konektivní struktury obsahující sloveso *souviset* (typu *s tím souvisí, to souvisí s* apod.) se v mluvené češtině používají především (v korpusu DIALOG v 90 % vztahů vyjádřených konektivní strukturou se slovesným základem *souviset*) k prostému, volnějšímú navazování dvou různých témat, přičemž význam daného diskurzního vztahu je spíše vágní, zamlžený – srov. např. *s tím souvisí i druhá věc, to souvisí s celou řadou dalších otázek, to souvisí eště s jednou věcí, s tím souvisí třetí problém* apod.

V psané češtině jsou diskurzní vztahy vyjádřené slovesnými frázemi typu *s tím souvisí* naopak ve většině případů „pevnější“, významově jasnější či zřetelnější (v korpusu PDT jsou tyto vztahy zastoupeny 75 %). Nevyjadřují tedy prosté navazování, ale mohou se významově blížit např. vztahům příčiny a důsledku či explikace.

6.9 Testování univerzálnosti sekundárních konektorů

V naší práci vymezujeme kategorii diskurzních konektorů především podle tzv. principu univerzálnosti (srov. kapitola 3.7.2), tj. jako diskurzní konektory chápeme takové jazykové prostředky, které jsou vhodnými ukazateli diskurzních vztahů ve velkém množství různých kontextů (konektory tedy nejsou vázány pouze na určitý kontext). V následující části práce představíme způsob, jakým jsme univerzálnost jednotlivých konektorů testovali a vyhodnocovali v rámci našeho výzkumu.

Pro testování univerzálnosti jsme vybrali dvě skupiny strukturně odlišných sekundárních konektorů – ty, jejichž jádro tvoří sekundární předložky *díky, kvůli, na základě, vzhledem k, vinou*, a ty, jejichž jádrem je podstatné jméno *důvod*. Jedná se přitom o struktury, které signalizují diskurzní vztah příčiny a důsledku.

Pro testování jsme přitom zvolili sto různých kontextů z Pražského závislostního korpusu, ve kterých se vyskytuje diskurzní vztah příčiny a důsledku. Do těchto kontextů jsme posléze dosazovali jednotlivé konektivní struktury a zjišťovali, pro jaké procento kontextů jsou vhodnými ukazateli diskurzního vztahu.

Zvolili jsme přitom padesát kontextů z PDT, ve kterých se v originálním, autentickém textu objevil sekundární konektor (příp. volná konektivní fráze), a padesát kontextů s primárními konektory *proto* a *a tak*.

V první fázi jsme testovali konektivní spojení (obsahující zmíněná jádrová slova), která se skutečně objevila v PDT (u struktur obsahujících slovo *důvod* jsme vybrali pouze část z nich, cílem bylo získat vyvážený počet kontextů s primárními konektory a s adepty na konektory sekundární) – srov. tabulka 22.

Tabulka zachycuje příslušná jádrová slova daných konektivních struktur, jejich vybrané konkrétní realizace v Pražském závislostním korpusu a počet kontextů, ve kterých se dané struktury objevily. V tabulce jsou dále uvedeny vybrané primární konektory (se kterými budeme srovnávat výsledky u konektorů sekundárních) – v PDT jsme přitom vyhledali padesát kontextů s těmito konektory, tj. počet odpovídající kontextům pro adepty na konektory sekundární.

Adepti na sekundární konektory (seřazeno podle jádrových slov)		Počet kontextů	Primární konektory	Počet kontextů
díky	<i>díky tomu</i>	7	<i>proto</i>	25
	<i>díky neinformovanosti, která je také nadhledem</i>	1	<i>a tak</i>	25
	<i>díky této skulince</i>	1		
	<i>díky všem těmto opatřením</i>	1		
	<i>díky neomezené tréninkové kapacitě</i>	1		
	<i>díky této preciznosti</i>	1		
	<i>díky pomoci příbuzných a přátel z USA</i>	1		
	<i>díky své drzosti a někdy téměř nedostatku dobrého vkusu</i>	1		
	<i>díky jeho chorobě</i>	1		
kvůli	<i>kvůli tomu</i>	3		
	<i>kvůli častějšímu broušení hlav</i>	1		
na základě	<i>na základě těchto faktorů</i>	1		
	<i>na základě čehož</i>	1		
	<i>na základě zmíněných aktivit</i>	1		
	<i>na základě těchto poznatků</i>	1		
	<i>právě na základě svých celonárodních poválečných diskusí</i>	1		
	<i>na základě tohoto koncertu</i>	1		
vzhledem k	<i>vzhledem k tomu</i>	1		
	<i>vzhledem k této zkušenosti</i>	1		
	<i>vzhledem k uvedeným potížím s uznáváním certifikátů</i>	1		
	<i>vzhledem k těmto skutečnostem</i>	1		
vinou	<i>především vinou toho</i>	1		
důvod	<i>z tohoto důvodu</i>	8		
	<i>z toho důvodu</i>	3		
	<i>to je důvod, proč</i>	6		
	<i>z uvedených důvodů</i>	1		
	<i>(právě) z politických důvodů uvnitř Unie</i>	1		
	<i>právě tato skutečnost se stala jedním z důvodů, kvůli kterým</i>	1		
Počet kontextů celkem		50		50

Tabulka 22: Vybrané konektivní prostředky pro testování univerzálnosti

Konektivní struktury obsahující sekundární předložky jsou charakteristické tím, že nabývají konektivní funkce pouze ve spojení s určitým anaforickým výrazem odkazujícím k předchozímu diskurzivnímu argumentu (srov. např. *díky tomu, vzhledem k těmto skutečnostem, díky této preciznosti* apod.). V PDT jsme tedy vybrali konektivní struktury s anaforickými výrazy (vyskytující se v reálných textech), které odkazují k příslušnému diskurzivnímu argumentu – v PDT se tyto případy dají automaticky vyhledávat podle toho, že od příslušného anaforického výrazu vede koreferenční vztah (zachycený v rámci anotace Nedoluzhko a kol., 2011).

V těchto případech bude tedy testování univerzálnosti v podstatě vyhodnocovat míru lexikální naplněnosti či vyprázdňenosti daných anaforických výrazů, resp. jejich schopnost významově pokrýt co největší počet různých kontextů.

Dá se tedy předpokládat, že např. deiktické ukazovací zájmeno *to* bude v tomto ohledu univerzálnější „referencí“ k různým diskurzivním argumentům (tj. dokáže se významově přizpůsobit kontextu či v závislosti na kontextu naplnit významem) než např. výraz *tato preciznost*, který může odkazovat pouze k diskurzivním argumentům, v nichž se o preciznosti hovoří. Z tohoto důvodu bude spojení *díky tomu* pravděpodobně univerzálnějším ukazatelem diskurzivního vztahu příčiny a důsledku než spojení *díky této preciznosti*.

Vyhodnocování univerzálnosti konektivních struktur je tedy podle našeho názoru částečně intuitivní – spojení *kvůli tomu* by pravděpodobně většina čtenářů intuitivně označila jako univerzálnější ukazatel diskurzivního vztahu, zatímco např. *díky této skutečnosti* jako kontextově závislý ukazatel. V těchto případech se ovšem jedná o relativně jasné případy, resp. o dva opačné póly na škále univerzálnosti – nicméně zajímavé bude především vyhodnocení struktur mezi těmito dvěma póly, tj. spojení typu *díky této skutečnosti, na základě těchto faktorů, díky (všem) těmto opatřením* apod. V uvedených případech jsou na místě anaforických výrazů obecná jména (*skutečnost, faktor, opatření*), otázkou ovšem zůstává míra jejich obecnosti, resp. univerzálnosti. Pravděpodobně právě mezi těmito strukturami povede podle našeho názoru hranice oddělující (spíše) univerzální a neuniverzální konektivní struktury, tedy hranice vymezující sekundární konektory.

Pro ilustraci nyní uvedeme několik příkladů a blíže představíme způsob, jakým jsme při testování univerzálnosti postupovali – viz následující příklady originálních kontextů s diskurzivním vztahem příčiny a důsledku z PDT (s původními konektory či konektivními frázemi). Testovaných kontextů bylo celkem 100.

Kontext č. 1: *Každý den bohužel platí to, že míra inflace v ČR je výrazně větší než ve dvou zemích, vůči jejichž měnám máme fixován kurz koruny, tzn. v USA a SRN. Díky tomu dochází k permanentnímu zpevnování koruny a to zpevnění v dnešní chvíli oproti momentu devalvace činí 80–90 procent.*

Kontext č. 2: *Kantor zůstal při představení na jevišti, řídil je způsobem takřka „dirigentským“ a opravoval detaily jako pečlivý malíř, jenž dokončuje svůj obraz. Díky této preciznosti můžeme Zemřelou třídu vnímat jako dílo na pomezí divadla a výtvarného umění.*

Kontext č. 3: *V důvodové zprávě návrhu se uvádí, že k jeho vypracování bylo využito poznatků z evropského zákonodárství. Zdůrazňuje se rovněž soulad s Evropskou konvencí o lidských právech a příslušnými doporučeními Rady Evropy. Vzhledem k těmto skutečnostem by bylo přirozené předpokládat, že návrh „tiskového“ zákona se opírá o praxi západního světa a nemá sporná místa.*

Kontext č. 4: *S prodejem samotných počítačů by však firma neobstála. Proto vedle obvyklého programového vybavení nabízí také některé programy, které jsou nezbytné či užitečné pro podnikatele.*

Z originálních kontextů jsme následně vyjmuli příslušné konektory či konektivní fráze (v daných příkladech vyznačené podtržením) a dosazovali jsme do nich jednotlivé testované výrazy či spojení. Do výsledků testování univerzálnosti jsme přitom pro konkrétní výraz nezapočítávali jeho originální kontext – netestovali jsme, zda je vhodný do kontextu, ve kterém se reálně objevil, ale zda je vhodným konektivním prostředkem pro ostatní kontexty – srov. následující příklady pro jednotlivé konektory či konektivní fráze:

Díky tomu

Kontext č. 2: *Kantor zůstal při představení na jevišti, řídil je způsobem takřka „dirigentským“ a opravoval detaily jako pečlivý malíř, jenž dokončuje svůj obraz.*

Díky tomu můžeme Zemřelou třídu vnímat jako dílo na pomezí divadla a výtvarného umění.

Kontext č. 3: *V důvodové zprávě návrhu se uvádí, že k jeho vypracování bylo využito poznatků z evropského zákonodárství. Zdůrazňuje se rovněž soulad s Evropskou konvencí o lidských právech a příslušnými doporučeními Rady Evropy.*

Díky tomu by bylo přirozené předpokládat, že návrh „tiskového“ zákona se opírá o praxi západního světa a nemá sporná místa.

Kontext č. 4: *S prodejem samotných počítačů by však firma neobstála.*

Díky tomu vedle obvyklého programového vybavení nabízí také některé programy, které jsou nezbytné či užitečné pro podnikatele.

Díky této preciznosti

Kontext č. 1: *Každý den bohužel platí to, že míra inflace v ČR je výrazně větší než ve dvou zemích, vůči jejichž měnám máme fixován kurz koruny, tzn. v USA a SRN.*

**Díky této preciznosti dochází k permanentnímu zpevnování koruny a to zpevnění v dnešní chvíli oproti momentu devalvace činí 80–90 procent.*

Kontext č. 3: *V důvodové zprávě návrhu se uvádí, že k jeho vypracování bylo využito poznatků z evropského zákonodárství. Zdůrazňuje se rovněž soulad s Evropskou konvencí o lidských právech a příslušnými doporučeními Rady Evropy.*

**Díky této preciznosti by bylo přirozené předpokládat, že návrh „tiskového“ zákona se opírá o praxi západního světa a nemá sporná místa.*

Kontext č. 4: *S prodejem samotných počítačů by však firma neobstála.*

**Díky této preciznosti vedle obvyklého programového vybavení nabízí také některé programy, které jsou nezbytné či užitečné pro podnikatele.*

Vzhledem k těmto skutečnostem

Kontext č. 1: *Každý den bohužel platí to, že míra inflace v ČR je výrazně větší než ve dvou zemích, vůči jejichž měnám máme fixován kurz koruny, tzn. v USA a SRN.*

Vzhledem k těmto skutečnostem dochází k permanentnímu zpevnění koruny a to zpevnění v dnešní chvíli oproti momentu devalvace činí 80–90 procent.

Kontext č. 2: *Kantor zůstával při představení na jevišti, řídil je způsobem takřka „dirigentským“ a opravoval detaily jako pečlivý malíř, jenž dokončuje svůj obraz.*

Vzhledem k těmto skutečnostem můžeme Zemřelou třídu vnímat jako dílo na pomezí divadla a výtvarného umění.

Kontext č. 4: *S prodejem samotných počítačů by však firma neobstála.*

Vzhledem k této skutečnosti⁴⁹ vedle obvyklého programového vybavení nabízí také některé programy, které jsou nezbytné či užitečné pro podnikatele.

Proto

Kontext č. 1: *Každý den bohužel platí to, že míra inflace v ČR je výrazně větší než ve dvou zemích, vůči jejichž měnám máme fixován kurz koruny, tzn. v USA a SRN.*

Proto dochází k permanentnímu zpevnění koruny a to zpevnění v dnešní chvíli oproti momentu devalvace činí 80–90 procent.

Kontext č. 2: *Kantor zůstával při představení na jevišti, řídil je způsobem takřka „dirigentským“ a opravoval detaily jako pečlivý malíř, jenž dokončuje svůj obraz.*

Proto můžeme Zemřelou třídu vnímat jako dílo na pomezí divadla a výtvarného umění.

Kontext č. 3: *V důvodové zprávě návrhu se uvádí, že k jeho vypracování bylo využito poznatků z evropského zákonodárství. Zdůrazňuje se rovněž soulad s Evropskou konvencí o lidských právech a příslušnými doporučeními Rady Evropy.*

Proto by bylo přirozené předpokládat, že návrh „tiskového“ zákona se opírá o praxi západního světa a nemá sporná místa.

⁴⁹ Při testování univerzálnosti konektorů a konektivních struktur jsme nerozlišovali jednotné a množné číslo (tj. nerozlišovali jsme mezi strukturami vzhledem k této skutečnosti a vzhledem k těmto skutečnostem).

Z uvedených příkladů je patrné, že výrazy či spojení *díky tomu, vzhledem k těmto skutečnostem* a *proto* se hodí i do všech zbylých tří kontextů, zatímco spojení *díky této preciznosti* se významově nehodí do žádného jiného uvedeného kontextu (než svého původního). Na tomto malém vzorku dat bychom pak vypočetli univerzálnost daných výrazů následujícím způsobem: *díky tomu, vzhledem k těmto skutečnostem* a *proto* jsou vhodné pro 3 (neoriginální) kontexty z celkových 3 (originální kontext nezapočítáváme), tj. jejich univerzálnost (v rámci tohoto malého vzorku) je 1 (3/3), tedy 100 %. Spojení *díky této preciznosti* je vhodné pro 0 (neoriginálních) kontextů z celkových 3, tj. jeho univerzálnost je 0 (0/3), tedy 0 %.

6.9.1 Stanovení „hranice univerzálnosti“ pro diskurzní konektory

Výše popsaným způsobem jsme postupovali u všech vybraných spojení či výrazů – výzkum jsme prováděli na 100 původních kontextech z Pražského závislostního korpusu. Při testování jednotlivých struktur jsme nerozlišovali jednotné či množné číslo (viz výše *vzhledem k této skutečnosti* vs. *vzhledem k těmto skutečnostem*), ale užili jsme takovou formu, která byla pro daný kontext vhodnější (tj. netestovali jsme např. struktury *z tohoto důvodu* a *z těchto důvodů* zvlášť, ale dohromady). V příslušných tabulkách a grafech je nicméně pro ilustraci užita pouze jedna forma, a to ta, ve které se daná konektivní struktura objevila v autentických textech z PDT častěji (tj. např. *z tohoto důvodu, z uvedených důvodů* apod.).

Souhrnné výsledky výzkumu zachycuje tabulka 23 a graf 2 (výzkumu se zúčastnili tři lingvisté s praktickými zkušenostmi se zpracováním a analýzou diskurzních vztahů; následující tabulky a grafy zobrazují konečné výsledky, které jsou průměrem získaných hodnot od tří anotátorů).

Z výsledků je patrné, že výrazný předěl mezi zkoumanými výrazy je za hranicí 75 % přijatelnosti či vhodnosti daných výrazů v nepůvodních kontextech. Nejvýraznější rozdíl se přitom ukázal mezi sousedními spojeními *vzhledem k tomu* (77 %) a *především vinou toho* (34 %). **Z tohoto důvodu jsme se rozhodli stanovit hranici vymezující množinu diskurzních konektorů právě zde, tj. za univerzální diskurzní konektor** (ať už primární nebo sekundární) **považujeme takovou konektivní strukturu, která se hodí alespoň do 75 % kontextů⁵⁰ vyjadřujících příslušný diskurzní vztah** (např. vztah příčiny a důsledku), viz graf 2.

⁵⁰ V některých případech je ale nutné přihlídnout ještě k dalším aspektům, jako jsou sémantické a pragmatické rysy dané konektivní struktury či její stylová platnost.

Zajímavé je, že výraznou roli v tomto případě patrně nehraje délka daného výrazu – viz spojení (*právě*) *tato skutečnost se stala jedním z důvodů, kvůli kterým*, jehož univerzálnost dosáhla v našem experimentu 85 %. Zároveň je ovšem nutné zohlednit funkční styl příslušných textů, do kterých dané výrazy dosazujeme. Spojení typu (*právě*) *tato skutečnost se stala jedním z důvodů, kvůli kterým; z uvedených důvodů* či *vzhledem k těmto skutečnostem* jsou jako diskurzní konektory vhodné spíše pro publicistické či odborné texty⁵¹. Rozdíl v užívání různých typů diskurzních konektorů je patrný také v textech mluvených a psaných (viz kapitola 6.8).

U jednotlivých výrazů či spojení jsme dále zkoumali jejich frekvenci v textech, tj. zaměřili jsme se na to, zda při určování příslušného jazykového prostředku jako diskurzního konektoru můžeme zohledňovat také jeho frekvenci v reálných textech, příp. jaký je vztah frekvence a univerzálnosti. Pro zjišťování frekvence zkoumaných výrazů jsme použili korpus SYN2013PUB⁵² (obsahující 935 milionů slov z výhradně publicistických textů).

Z výsledků je patrné, že nejvyšší frekvenci v reálných textech mají primární konektory – *proto* (733 198 výskytů), *a tak* (228 532) – a dále předložkové fráze obsahující zájmeno *ten* – např. *kvůli tomu* (53 755), *díky tomu* (40 376), *vzhledem k tomu* (39 497). Všechny tyto výrazy mají zároveň vysokou míru univerzálnosti, tj. vhodnosti užití jako konektivního prostředku ve velkém množství různých kontextů.

Většina spojení, která v našem výzkumu dosáhla velmi nízkého stupně univerzálnosti, se zároveň v korpusu SYN2013PUB vyskytuje s malou frekvencí (např. spojení *díky této skulince* má univerzálnost 1 % a zároveň nemá žádný výskyt v SYN2013PUB). V těchto případech je tedy univerzálnost daných jazykových prostředků v souladu s jejich frekvencí v reálných textech.

Na druhou stranu se některá spojení s vysokou mírou univerzálnosti vyskytují v korpusu SYN2013PUB s poměrně nízkou frekvencí (ve srovnání např. s primárními konektory; spojení *vzhledem k těmto skutečnostem* dosáhlo univerzálnosti 88 %, v korpusu SYN2013PUB se vyskytuje ve 145 výskytech), příp. se v uvedeném korpusu nevyskytují vůbec (srov. spojení (*právě*) *tato skutečnost se stala jedním z důvodů, kvůli kterým* s univerzálností 85 %).

⁵¹ Univerzálnost konektorů je tedy třeba testovat na textech odpovídajícího stylu. Skutečnost, že se např. struktura *z tohoto důvodu* nehodí do většiny kontextů z neformálních rozhovorů ještě není důvodem ke konstatování, že tuto strukturu nelze považovat za konektor.

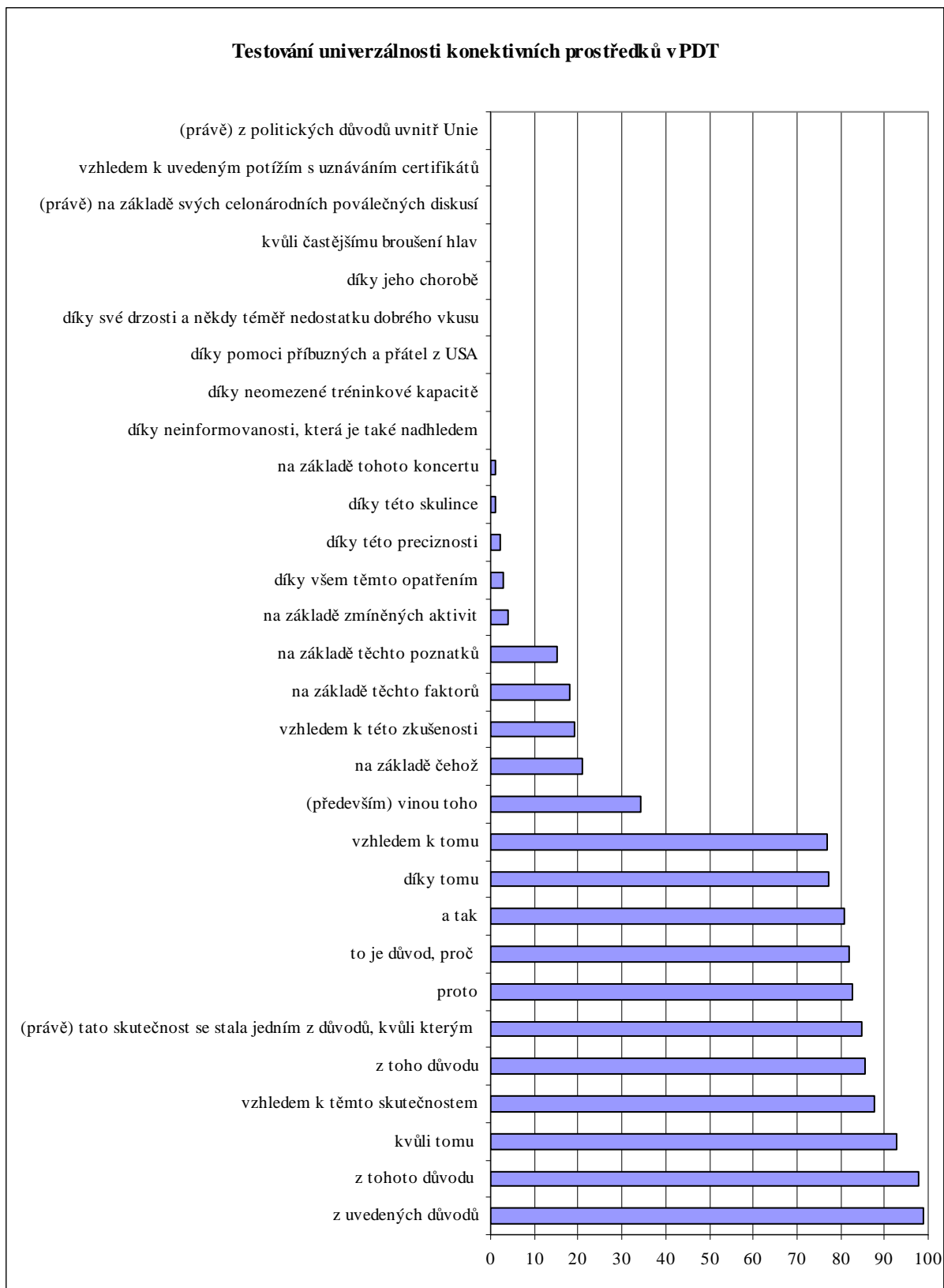
⁵² U jednotlivých výskytů jsme blíže nezkoumali jejich funkci v textu, tj. tabulka zobrazuje absolutní čísla výskytů u jednotlivých výrazů nalezených v korpusu SYN2013PUB.

V korpusu SYN2013PUB jsme zároveň pro ilustraci vyhledali spojení *kvůli své nemoci*, které má pravděpodobně (analogicky k testovanému *díky jeho chorobě*) nízkou univerzálnost – v reálných textech korpusu se objevilo ve 106 výskytech (tj. v textech se vyskytuje s podobnou frekvencí jako např. *vzhledem k těmto skutečnostem*).

Z těchto důvodů nebereme frekvenci jazykových prostředků v reálných textech jako hlavní kritérium pro určování hranice diskurzních konektorů (jako základní kritérium chápeme jejich univerzálnost pro použití v různých kontextech). **Nicméně frekvence příslušných výrazů vypovídá o míře jejich ustálenosti, tj. v rámci kategorie diskurzních konektorů frekvence supluje hranici mezi konektory primárními a sekundárními** (jak jsme popsali v kapitole 6.3, v PDT je mezi všemi výskyty diskurzních konektorů 94,6 % primárních konektorů a 5,4 % sekundárních).

Testované výrazy	Výskyt v originále v PDT	Přijatelnost ve zbylých kontextech v PDT	Univerzálnost %	Výskyty v ČNK v SYN2013PUB
<i>díky tomu</i>	7	72	77	40 376
<i>díky neinformovanosti, která je také nadhledem</i>	1	0	0	0
<i>díky této skulince</i>	1	1	1	0
<i>díky všem těmto opatřením</i>	1	3	3	3
<i>díky neomezené tréninkové kapacitě</i>	1	0	0	0
<i>díky této preciznosti</i>	1	2	2	0
<i>díky pomoci příbuzných a přátel z USA</i>	1	0	0	0
<i>díky své drzosti a někdy téměř nedostatku dobrého vkusu</i>	1	0	0	0
<i>díky jeho chorobě</i>	1	0	0	0
<i>kvůli tomu</i>	3	90	93	53 755
<i>kvůli častějšímu broušení hlav</i>	1	0	0	0
<i>na základě těchto faktorů</i>	1	18	18	0
<i>na základě čehož</i>	1	21	21	39
<i>na základě zmíněných aktivit</i>	1	4	4	0
<i>na základě těchto poznatků</i>	1	15	15	45
<i>(právě) na základě svých celonárodních poválečných diskusí</i>	1	0	0	0
<i>na základě tohoto koncertu</i>	1	1	1	0
<i>vzhledem k tomu</i>	1	76	77	39 497
<i>vzhledem k této zkušenosti</i>	1	19	19	1
<i>vzhledem k uvedeným potížím s uznáváním certifikátů</i>	1	0	0	0
<i>vzhledem k těmto skutečnostem</i>	1	87	88	145
<i>(především) vinou toho</i>	1	34	34	188
<i>z tohoto důvodu</i>	8	90	98	8 801
<i>z toho důvodu</i>	3	83	86	8 623
<i>to je důvod, proč</i>	6	77	82	1 884
<i>z uvedených důvodů</i>	1	98	99	189
<i>(právě) z politických důvodů uvnitř Unie</i>	1	0	0	0
<i>(právě) tato skutečnost se stala jedním z důvodů, kvůli kterým</i>	1	84	85	0
<i>proto</i>	25	62	83	733 198
<i>a tak</i>	25	61	81	228 532

Tabulka 23: Testování univerzálnosti konektivních výrazů v Pražském závislostním korpusu



Graf 2: Graf míry univerzálnosti konektivních výrazů v Pražském závislostním korpusu

Graf 2 dále ukazuje, že prvními dvěma výrazy, které skončily pod hranicí 75 % univerzálnosti, jsou spojení (*především*) *vinou toho* (34 %) a *na základě čehož* (21 %). Ačkoli daná spojení neprokázala vysokou míru univerzálnosti v rámci vybraných textů z PDT, poukázala na další aspekty, které je v rámci vymezování diskurzních konektorů nutné zohledňovat – a to na sémantická omezení daného konektoru s přesahem do pragmatiky a na jeho syntaktické chování.

Spojení *vinou toho* je z hlediska použitelnosti ve funkci diskurzního konektoru značně limitované sémantikou sekundární předložky *vinou*. Její použití je vázáno na negativně vyznívající kontexty, tj. „vyjadřuje příčinu něčeho nežádoucího“ (viz *Slovník spisovné češtiny pro školu a veřejnost*, 2001). Zároveň je posouzení příslušného kontextu (tj. to, zda bude vyhodnocen jako kladný či záporný) na rozhodnutí mluvčího, tedy na tom, co chce mluvčí v dané výpovědi skutečně vyjádřit, jaký je jeho postoj k vyjadřovanému. Podle toho také volí příslušné konektivní prostředky. Zatímco některé výrazy původní vázanost na kladné či záporné kontexty postupně ztrácejí (srov. např. předložku *díky*, která se v reálných textech PDT objevuje i v negativních kontextech typu *díky jeho chorobě*), jiné si tuto vázanost ponechávají (srov. *vinou*). Zatímco tedy spojení *díky tomu* může být použito i v kontextech, ve kterých se původně vyskytovalo *vinou toho*⁵³, obdobná nahraditelnost v opačném směru nefunguje – srov. příklad z PDT, ve kterém jsme spojení *díky tomu* nahradili spojením *vinou toho*:

- (108) *Máme určité kontaktní možnosti ve všech státech, odkud pocházeli zahraniční studenti v bývalém Československu.*

**Vinou toho bychom tam mohli hledat uplatnění pro naše lidi, a naše licence je proto pojata dosti široce.*

Úzká vázanost předložky *vinou* na negativní kontexty pak brání častějšímu používání spojení typu *vinou toho* jako diskurzních konektorů⁵⁴ v autentických textech. Nicméně vzhledem k tomu, že spojení *vinou toho* má analogickou strukturu jako ostatní konektivní prostředky skládající se ze sekundární předložky a zájmena *ten* (např. *kvůli tomu, díky tomu, vzhledem k tomu*), které v našem pojetí chápeme jako konektory (vzhledem k vysoké míře jejich

⁵³ Podle textů PDT.

⁵⁴ Sémantickými a pragmatickými aspekty diskurzních konektorů se podrobně věnuje J. Hoffmannová (srov. např. 1983, 1984). Ve svých pracích přináší podrobnou klasifikaci diskurzních konektorů právě na základě podílu sémantických a pragmatických rysů.

využitelnosti v různých kontextech), budeme analogicky i spojení *vinou toho* řadit mezi diskurzí konektory.

Druhým výrazem pod hranicí univerzálnosti 75 % je spojení *na základě čehož* (21 %). Dané spojení je v tomto případě limitováno svým syntaktickým chováním – v textech se objevuje v rámci vnitrovětných vztahů. Jeho mezivětné použití by mohlo být vnímáno jako syntaktický nedostatek. Na míře univerzálnosti spojení *na základě čehož* se tedy v našem zkoumání projevuje, že většina vybraných kontextů z PDT obsahuje mezivětné vztahy.

Prozkoumali jsme tedy znovu vhodnost spojení *na základě čehož* pouze v rámci příslušných vnitrovětných vztahů (v našem vzorku dat se jednalo o 26 možných kontextů) – v tomto případě dosáhla jeho univerzálnost 81 %. Z tohoto důvodu budeme v našem pojetí chápat i *na základě čehož* jako diskurzí konektor. Mezi primárními konektory jsou na vnitrovětné užití obdobným způsobem vázány výrazy *což* nebo *příčemž*.

Jak jsme již naznačili výše, autor vybírá konektivní prostředky především s ohledem na adresáta, tj. s ohledem na to, aby komunikace proběhla pokud možno co nejefektivněji a adresát dobře porozuměl autorovu komunikačnímu záměru. Na výběr příslušného konektivního prostředku přitom mají vliv různé faktory či omezení. Výsledky naší analýzy např. ukázaly, že jedním z těchto faktorů může být vzdálenost diskurzí argumentů, tj. zda jsou dané argumenty v textu vedle sebe nebo zda je mezi ně vložen další text (v PDT se objevují případy, ve kterých sahá diskurzí vztah např. přes celý odstavec). Z výsledků analýzy je patrné, že v takových případech (pokud hrozí nesprávná interpretace daného diskurzího vztahu) volí mluvčí raději volné konektivní fráze, které opakují část lexikálního významu příslušného argumentu – srov. příklad 109:

(109) ***Po každé směně jsme je [hlavy, M. R.] museli ostřit.***

Po 14 dnech nepřetržitého dolad'ování byl stroj již k použití.

Úspora, vzniklá koupí našich hlav, se však změnila v pasivum, neboť technik seřizoval stroj místo tří dnů čtrnáct.

Kvůli častějšímu broušení hlav vzrostly i náklady na provoz zařízení.

V uvedeném příkladu je diskurzí vztah příčiny a důsledku mezi tučně zvýrazněnými argumenty *Po každé směně jsme je museli ostřit.* a *Kvůli častějšímu broušení hlav vzrostly i náklady na provoz zařízení* signalizovaný spojením *kvůli častějšímu broušení hlav*, díky kterému čtenář příslušný diskurzí vztah snáze interpretuje. Pokud bychom místo daného

spojení použili konektor typu *kvůli tomu* či *proto*, diskurzní vztahy by byly v tomto úseku textu poněkud zamlžené – nebylo by zcela zřejmé, ke kterému argumentu dané konektory odkazují. V úvahu by totiž přicházela i následující interpretace – *Technik seřizoval stroj místo tří dnů čtrnáct, kvůli tomu / proto vzrostly i náklady na provoz zařízení*. Použití volné konektivní fráze (v tomto případě spojení *kvůli častějšímu broušení hlav*) tak může být použito pro snazší interpretaci diskurzních vztahů a lepší orientaci čtenáře v předkládaném textu.

Na druhou stranu četnost užití konektorů a volných konektivních frází v PDT (viz kapitola 6.3) ukazuje, že autor při tvorbě textu (pokud je to možné) raději užívá ustrnulé konektivní prostředky (tedy prostředky, které nemusí sám nově formulovat) než prostředky neustrulé.

Je tedy zřejmé, že na výběr vhodného konektivního prostředku působí několik faktorů zároveň – na jedné straně volí mluvčí takové prostředky, které jsou pro něj při vytváření textu v dané situaci nejjednodušší (tj. v případě konektivních výrazů se jedná častěji o ustálenější, frekventovanější, mnohdy jednoslovné konektory). Zároveň musí brát mluvčí ohled na adresáta, tj. na srozumitelnost svého sdělení (správná interpretace diskurzních vztahů je jedním z jejích klíčových aspektů).

V tomto ohledu působí oba zmíněné faktory, tj. co nejmenší namáhavost procesu vytváření textu na straně mluvčího a zároveň co nejvyšší srozumitelnost textu pro adresáta, do určité míry proti sobě – zároveň existují různé okolnosti či omezení, které tento rozpor mohou převážit na jednu či druhou stranu.

Z analýzy našich dat vyplývá, že nejhodnějšími konektivními prostředky podílejícími se na utváření koherentního, srozumitelného textu jsou (z hlediska frekvence) primární konektory (v PDT se vyskytují v 96 % diskurzních vztahů), tedy struktury plně gramatikalizované, ustrulé, po nich sekundární konektory, tedy struktury, u nichž proces gramatikalizace a lexikalizace teprve probíhá, a naposled volné konektivní fráze, tedy struktury plně negramatikalizované.

Nicméně existují případy, ve kterých se jako nejhodnější prostředky jeví naopak volné konektivní fráze, a to při větší vzdálenosti diskurzních argumentů v textu, kdy použití lexikálně „naplněnější“ struktury zabraňuje víceznačnému výkladu textu.

Tento jev (tedy lexikální naplněnost konektivního prostředku a míra opakování v textu) souvisí zároveň s výzkumem tzv. salience (tj. stupněm aktivovanosti objektů v textu⁵⁵) nebo

⁵⁵ Více k salinci viz např. Hajičová a kol., 2003, 2004.

dále s teorií relevance a mírou explicitnosti a implicitnosti vyjadřování diskurzivních vztahů (tj. s tím, kdy a jakým způsobem musíme v textu vyjadřovat diskurzivní vztahy, aby byl text stále koherentní a srozumitelný pro čtenáře a zároveň aby užívání konektivních prostředků nebylo nadbytečné⁵⁶).

6.9.2 Testování předložkových frází typu *díky tomu* a *na základě těchto skutečností*

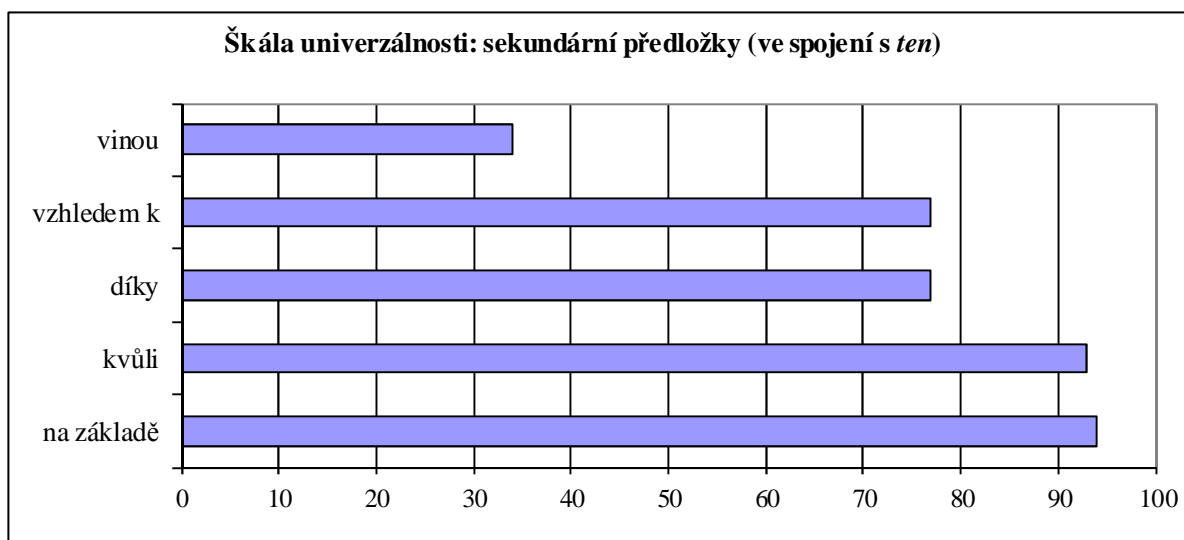
V další fázi testování univerzálnosti konektivních struktur jsme se zaměřili na vybrané sekundární předložky ve spojení se zájmenem *ten*, tj. výrazem s pokud možno co „nejobecnějším významem“. Cílem bylo ověřit vhodnost vybraných předložek pro různé kontexty. Jednalo se o následující spojení – *díky tomu, kvůli tomu, na základě toho, vzhledem k tomu, vinou toho*. Většina těchto struktur se objevila v autentických textech z PDT, s výjimkou *na základě toho* – viz tabulka 24.

Testované výrazy	Výskyt v originále v PDT	Přijatelnost ve zbylých kontextech v PDT	Univerzálnost %	Výskyty v ČNK v SYN2013PUB
<i>na základě toho</i>	0	94	94	2 669
<i>kvůli tomu</i>	3	90	93	53 755
<i>vzhledem k tomu</i>	1	76	77	39 497
<i>díky tomu</i>	7	72	77	40 376
<i>vinou toho</i>	1	34	34	188

Tabulka 24: Univerzálnost vybraných sekundárních předložek ve spojení se zájmenem *ten*

Na základě tabulky 24 můžeme vykreslit škálu univerzálnosti jednotlivých předložek, které se v reálných textech vyskytují v různých konektivních spojeních – srov. graf 3.

⁵⁶ Více k teorii relevance viz D. Sperber, D. Wilsonová (1986); D. Blakemoreová (1987), v českém prostředí např. I. Nebeská (1989, 1991a, 1991b) či J. Hoffmannová (1987, 1990, 1990/1, 1997).



Graf 3: Graf univerzálnosti vybraných sekundárních předložek ve spojení se zájmenem *ten*

Z grafu je patrné, že spojeními, která se hodí do nejvíce vybraných kontextů, jsou podle naší analýzy *na základě toho* (s univerzálností 94 %) a *kvůli tomu* (93 %). Sekundární předložku *na základě* jsme proto použili v další části našeho výzkumu, ve které jsme se zaměřili na testování univerzálnosti obecnějších podstatných jmen (typu *opatření, skutečnost* apod.)⁵⁷, z nichž některé v rámci příslušných předložkových frází prokázaly v první části testování vysokou míru univerzálnosti (srov. např. *vzhledem k těmto skutečnostem* – 88 %). (Jak je patrné z tabulky 23, všechna spojení obsahující konkrétní jména typu *tato preciznost, tato skulinka, jeho choroba* apod. jsou zcela jednoznačně kontextově závislá a nemají obecnou platnost diskurzivních konektorů; jejich vhodnost pro jiné kontexty se pohybuje mezi 0–2 %, a proto by testování dalších podobných spojení nemělo význam.)

V další fázi testování jsme skupinu obecnějších podstatných jmen (tj. *skutečnost, opatření, aktivita, faktor, poznatek, zkušenost*) obohatili o tři nová (*činnost, fakt, iniciativa*), která jsme získali vyhledáním dalších konektivních struktur v PDT obsahujících jiné sekundární předložky (které zároveň signalizují jiný typ sémantického vztahu než vztah příčiny a důsledku, tj. např. předložky *kromě, na rozdíl od, nehledě na* apod.).

Daná obecná jména jsme testovali znovu ve stu kontextech (stejných jako v první fázi testování), tentokrát ovšem ve spojení s jednou sekundární předložkou – *na základě*. Jak jsme

⁵⁷ Pro další testování jsme vybrali sekundární předložku *na základě*, i když spojení *na základě toho* má v korpusu SYN2013PUB nižší frekvenci výskytů než *kvůli tomu* (srov. tabulka 24). Nicméně ve spojení s jinými testovanými jmény (např. *skutečnost* či *zkušenost*) je situace opačná – tj. v korpusu SYN2013PUB je frekventovanější spojení *na základě této skutečnosti* (154 výskytů) než *kvůli této skutečnosti* (38) či *na základě této zkušenosti* (88) než *kvůli této zkušenosti* (23).

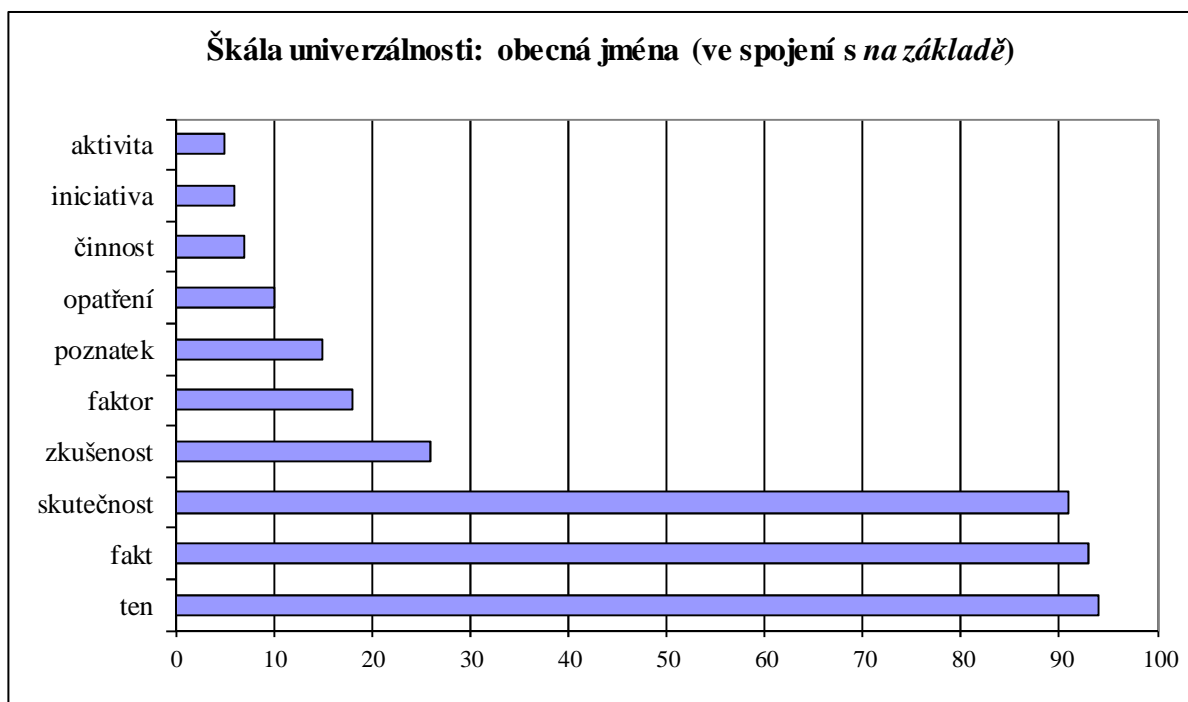
ukázali v předchozím výzkumu, na univerzálnost celého konektivního spojení může mít vliv i daná sekundární předložka.

Testovaná konektivní spojení v další fázi výzkumu byla následující: *na základě toho, na základě této skutečnosti, na základě tohoto opatření, na základě této aktivity, na základě těchto faktorů, na základě těchto poznatků, na základě této zkušenosti, na základě této činnosti, na základě tohoto faktu, na základě této iniciativy*. Při testování jsme opět u užitých obecných jmen nerozlišovali mezi variantou s jednotným a variantou s množným číslem, ale používali jsme vždy tu formu, která byla pro daný kontext přirozená. Výsledky této části výzkumu zobrazuje tabulka 25.

Testované výrazy	Výskyt v originále v PDT	Přijatelnost ve zbylých kontextech v PDT	Univerzálnost %	Výskyty v ČNK v SYN2013PUB
<i>na základě toho</i>	0	94	94	2 669
<i>na základě tohoto faktu</i>	0	93	93	27
<i>na základě této skutečnosti</i>	0	91	91	154
<i>na základě této zkušenosti</i>	0	26	26	88
<i>na základě tohoto faktoru</i>	1	18	18	0
<i>na základě tohoto poznatku</i>	1	15	15	45
<i>na základě tohoto opatření</i>	0	10	10	14
<i>na základě této činnosti</i>	0	7	7	2
<i>na základě této iniciativy</i>	0	6	6	6
<i>na základě této aktivity</i>	0	5	5	5

Tabulka 25: Univerzálnost vybraných obecných jmen ve spojení s předložkou *na základě*

Na základě výsledků v tabulce 25 můžeme vykreslit škálu univerzálnosti pro jednotlivá obecná jména ve spojení se sekundární předložkou *na základě* – srov. graf 4.



Graf 4: Graf univerzálnosti vybraných obecných jmen ve spojení s předložkou *na základě*

Z výsledků je patrné, že pouze tři zkoumaná spojení (*na základě toho*, *na základě tohoto faktu*, *na základě této skutečnosti*) přesahují hranici 75 % univerzálnosti, kterou jsme stanovili pro vymezení diskurzivních konektorů – dané výrazy tuto hranici dokonce výrazně převyšují (srov. *na základě toho* – 94 %, *na základě tohoto faktu* – 93 %, *na základě této skutečnosti* – 91 %). Daná spojení tedy budeme v našem pojetí považovat za diskurzivní konektory. Vhodnost ostatních zkoumaných struktur v různých kontextech se pohybuje mezi 26 % (*na základě této zkušenosti*) a 5 % (*na základě této aktivity*). Mezi spojeními *na základě této skutečnosti* a *na základě této zkušenosti* je tedy výrazný zlom.

Slova *skutečnost* a *fakt* se (především v odborných textech) vyskytují v podobných funkcích jako zájmeno *to* – srov. vyjádření typu *předmětem našeho zájmu je to / skutečnost, zda..., to / tento fakt můžeme doložit na následujícím příkladu* apod., tj. především při anaforickém a kataforickém odkazování v textu. Možnost vzájemné zastupitelnosti zmíněných slov ve velkém množství kontextů se ukázala i v našem výzkumu sekundárních konektorů.

7 Závěry

7.1 Definice diskurzních argumentů a konektorů

Na základě analýzy rozsáhlého jazykového materiálu jsme se pokusili vymezit pojem diskurzní argumenty a konektory. Diskurzní argumenty vymezujeme především sémanticky – diskurzní argument je abstraktní objekt, který může vyjadřovat stavy, události, situace atd. Argumenty bývají v textu realizovány především jednotlivými větami (či posloupností několika vět) nebo nominalizovanými frázemi. Z hlediska syntaxe tak mají nejčastěji podobu vět (jejichž základem je sloveso v určitém tvaru, a proto je nazýváme slovesnými argumenty) a jmenných frází (jejichž základem bývá nejčastěji slovesné podstatné jméno; tyto argumenty nazýváme nominálními).

V textech mohou být diskurzní argumenty realizovány například větou, souvětím, samostatnou vedlejší větou, posloupností několika vět i celým odstavcem.

V některých případech mohou být diskurzními argumenty i objekty vyjádřené jinak než slovně, tj. situace, k nimž mohou odkazovat konektory (např. v počátečních výpovědích typu *Tak jsem doma.*).

Diskurzní argumenty, tj. objekty, mezi kterými existují textové vztahy, tedy mohou být:

i. vyjádřené slovně;

- **slovesně** (realizované jednoduchou větou, souvětím, samostatnou vedlejší větou, posloupností několika vět i celým odstavcem či textem);
- **jmenně** (realizované jmennými frázemi);

ii. situační⁵⁸.

Diskurzní argumenty mohou stát v textu bezprostředně vedle sebe nebo od sebe mohou být „odděleny“ jiným textem, který není součástí ani jednoho z argumentů.

Diskurzní argumenty jsou propojeny vztahy, které mohou být vyjádřené nebo nevyjádřené. Textové vztahy mohou být vyjádřeny diskurzními konektory (primárními a sekundárními) či dalšími spojovacími prostředky. Hranici mezi diskurzními konektory a ostatními konektivními prostředky je možné stanovit na základě toho, zda příslušné výrazy

⁵⁸ Situačním argumentem může být například hluk ve třídě. Slovní reakce *Tak ticho!* obsahuje druhý (slovně vyjádřený) diskurzní argument (*ticho*) a konektor (*tak*), který oba argumenty spojuje.

fungují v textu jako indikátory diskurzních vztahů (ve velké míře) univerzálně či příležitostně (tj. v závislosti na kontextu). Univerzální ukazatele diskurzních vztahů nazýváme souhrnným názvem diskurzní konektory (nezávisle na jejich slovnědruhové příslušnosti, tj. např. *ale; nebo; naopak; to je důvod, proč; kvůli tomu; příčinou je; to znamená* aj.), pro neuniverzální ukazatele diskurzních vztahů používáme označení volné konektivní fráze (srov. např. *díky jeho chorobě; kvůli tomuto zvýšení; díky této skulince; díky této preciznosti; důvodem poklesu je; důvodem tohoto návrhu je*).

7.1.1 Charakteristika primárních a sekundárních konektorů

Vzhledem k tomu, že daná kritéria pro diskurzní konektory splňují značně různorodé jazykové prostředky (srov. např. jednoslovné výrazy *a, také, totiž* a víceslovná spojení typu *to je důvod, proč; příčinou bylo* atd.), vyčleňujeme v rámci kategorie diskurzních konektorů dvě skupiny, pro které zavádíme označení primární a sekundární konektory.

Jako **primární konektory** označujeme především symsémantika s primární spojovací funkcí, tj. výrazy patřící mezi spojky (např. *a, ale, však, i*) a částice strukturující text nebo jeho části (dělení slovních druhů přebíráme podle *Mluvnice češtiny 2* (1986) – srov. např. *totiž, přitom, navíc, pak, potom, nakonec, například*), některá vztažná zájmena (*což*) a vztažná zájmenná příslovce (*proto, přičemž*). Za primární konektory současně považujeme také některé grafické značky (dvojtečku a pomlčku), číselné a písmenné výrazy sloužící ke strukturaci textu, nejčastěji k vyjádření výčtu (např. *1., 2., 3.; a), b), c); i), ii), iii)*) a větnou negaci, pokud se podílí na vyjádření diskurzního vztahu (*Nebudu se jenom dívat na televizi, ale půjdu i na chvíli ven*).

U zmíněných příslovcí a zájmen se spojovací funkcí dochází často k přechodu ve spojky či strukturující částice (srov. např. *potom, pak, což*), dané výrazy tedy ztrácejí větněčlenskou platnost. Proces přechodu v jiný slovní druh ovšem nemusí být plně dokončen, při určování slovních druhů u zmíněných konektivních prostředků tak může docházet k neshodám.

Vzhledem ke zmíněné diskrepanci mezi slovními druhy (především mezi spojkami, částicemi a příslovcí) a ke skutečnosti, že zejména u mnohých příslovcí se spojovací funkcí dochází k přechodu k jinému slovnímu druhu, chápeme primární konektory především jako symsémantika, tj. výrazy bez větněčlenské platnosti.

Primární konektory jsou tedy většinou jednoslovné (některé ovšem mohou tvořit korelativní páry), gramatikalizované a lexikalizované výrazy, ve většině případů nedovolující

modifikaci (nelze například říci **obecně a, *jednoduše ale*, výjimku tvoří výrazy typu *jednoduše/obecně/hlavně proto, že*). Primární konektory bývají v různých studiích užívány jako typičtí zástupci diskurzních konektorů. Frekvence primárních konektorů v textech je mnohem vyšší než frekvence konektorů sekundárních (v Pražském závislostním korpusu vyjadřují primární konektory 94,6 % diskurzních vztahů vyjádřených explicitně). Z tohoto důvodu si můžeme primární konektory představit jako centrum množiny diskurzních konektorů a sekundární konektory jako její periferii.

Sekundární konektory jsou univerzální konektivní struktury obsahující autosémantickou složku (jejíž lexikální význam ale bývá oslaben). Jedná se o struktury, které zatím neprošly gramatikalizací (tj. v plném rozsahu – proces gramatikalizace u nich ovšem probíhá; u různých typů sekundárních konektorů je zároveň proces gramatikalizace v různém stadiu). Sekundární konektory jsou na rozdíl od primárních většinou víceslovné a v některých případech bývá obtížné určit jejich rozsah (při určování rozsahu sekundárních konektorů, tj. např. zda je sekundárním konektorem spojení *podmínkou je* nebo *podmínkou je, že*, bereme jako obecné kritérium možnost náhrady primárním konektorem⁵⁹).

Jednotlivá slova, která dané spojení tvoří, nemívají spojovací funkci ze své podstaty (jako spojky či strukturující částice), ale platnost konektoru nabývá až celá struktura. Ve větě mají sekundární konektory větněčlenskou platnost (některé mohou fungovat i jako větné modifikátory). Sekundární konektory ve formě samostatných vět do syntaxe diskurzních argumentů v tomto smyslu nezasahují. Sekundární konektory jsou zpravidla modifikovatelné (srov. *hlavním/jediným/prvním/významným/rozumným... důvodem je, jasně to odůvodnil* apod.). Všechny obsahují implicitní anaforickou referenci k celému prvnímu argumentu, tato reference se podle typu sekundárního konektoru může objevit i v povrchové realizaci konektoru. Sekundární konektory tak můžeme rozdělit do skupin podle toho, zda anaforickou referenci na povrchu vyjádřit mohou (*příkladem je* vs. *příkladem toho je*), nemohou (*stručně řečeno*) nebo musí (*kvůli tomu*).

Jedinečnost sekundárních konektorů spočívá v tom, že mohou (na rozdíl od konektorů primárních) stát syntakticky výše než druhý diskurzni argument, tj. formálně mohou mít povahu hlavní věty (srov. příklady typu *To je důvod, proč tam nepůjdu.*). Některé sekundární

⁵⁹ Sekundární konektor: *V rámci projektu se může stát promítačem kdokoli a promítat může kdekoli. Podmínkou je, že se na projekci bezplatně zapůjčeného filmu nebude vybírat vstupné.* Po náhradě primárním konektorem: *V rámci projektu se může stát promítačem kdokoli a promítat může kdekoli, pokud se na projekci bezplatně zapůjčeného filmu nebude vybírat vstupné.*

konektory mohou tvořit samostatné věty, tj. syntakticky i sémanticky zcela úplné úseky textu (srov. příklad *Důvod je jednoduchý.*). Struktura některých sekundárních konektorů dovoluje nominalizaci druhého argumentu (srov. např. *Důsledkem je stárnutí obyvatelstva.*) – viz tabulka 26.

Primární konektory	Sekundární konektory
synsémantika	struktury s autosémantickou složkou
univerzální	univerzální
gramatikalizované	negramatikalizované
většinou jednoslovné	většinou víceslovné
bez větňčlenské platnosti	větné členy, větné modifikátory či samostatné věty
nemodifikovatelné (s výjimkami)	modifikovatelné (s výjimkami)
	anaforická reference k prvnímu argumentu (explicitní či implicitní)
	jedinečnost některých struktur: a) syntakticky výše než druhý argument b) forma samostatné věty c) nominalizace druhého argumentu

Tabulka 26: Primární vs. sekundární konektory

Díky tomuto idiosynkratickému chování mají sekundární konektory v jazyce jedinečné postavení a diskurzívní analýza, která tyto jazykové prostředky nezohledňuje, je podle našeho názoru neúplná. Z tohoto důvodu je zapotřebí brát označení primárních konektorů za centrum množiny diskurzívních konektorů a sekundárních konektorů za její periferii především z frekvenčního hlediska, nikoli jako označení hodnotící důležitost či význam obou skupin.

Primární a sekundární konektory se zároveň mohou v textu objevovat vedle sebe (srov. např. *a z tohoto důvodu; a to je důvod, proč* atd.). V sousedství sekundárních konektorů se nejčastěji vyskytují spojky *a, ale, nebo*. Tyto spojky se ale také často objevují i ve spojení s jiným primárním konektorem (např. *a proto, ale také, nebo dokonce, a tudíž*). Ve spojování sekundárních konektorů s konektory primárními můžeme vidět také analogii k historickému vývoji některých primárních konektorů – srov. např. konektory *ale, avšak* apod. (které vznikly ze dvou komponentů *a + -le; a + však*).

V naší práci tedy kromě primárních (gramatikalizovaných) a sekundárních (dosud ne zcela gramatikalizovaných) konektorů rozlišujeme ještě konektory **složené** a **nesložené**, **modifikované** a **nemodifikované**.

Nesložený konektor v sobě zahrnuje (ze synchronního hlediska) konektor pouze jeden (ať už primární nebo sekundární), např. *a*; *ale*; *avšak*; *z tohoto důvodu*; *za těchto podmínek*. Z příkladů je patrné, že nesložený konektor může být jednoslovný i víceslovný. **Složený konektor** v sobě obsahuje více než jeden konektor (ať už jde o kombinaci dvou primárních konektorů nebo kombinaci primárního a sekundárního konektoru), např. *a tak*; *i když*; *a proto*; *a z tohoto důvodu*. Složené konektory jsou vždy víceslovné.

V některých případech může vyvstat otázka, zda je dané slovní spojení složeným konektorem nebo není. U složených konektorů (typu *a proto*, *i když*) dochází podle našeho názoru k oslabení jedné významové složky (často se jedná o konektory *a*, *i* s konjunktivním významem), tj. složený konektor vnímáme jako jazykový prostředek sloužící k vyjádření primárně jednoho typu významového vztahu (např. tedy složený konektor *a proto* vyjadřuje „silněji“ či zřetelněji vztah příčiny a důsledku než vztah konjunkce, *i když* vztah přípustky apod.). U daných konektorů tedy můžeme pozorovat postupný proces gramatikalizace – *a*, *i* v těchto případech postupně ztrácejí svůj původní význam a stávají se součástí konektoru vyjadřujícího jiný typ diskurzního vztahu – podobně je tomu např. u výše zmíněných konektorů *ale*, *avšak*, *anebo*, které v současné době vyjadřují v textu zcela zřetelně vždy pouze jeden vztah (*ale*, *avšak* vztah opozice, *anebo* vztah disjunktivní či konjunktivní alternativy). Významový vztah konjunkce byl v těchto případech zcela potlačen. Jako složený konektor tedy chápeme výrazy skládající se z více konektorů, u kterých dominuje jeden významový vztah. Pokud si každý ze společně se vyskytujících konektorů zcela zřetelně ponechává svůj původní význam (tj. na daném místě textu silně vnímáme dva různé významové vztahy), považujeme je za navzájem různé diskurzní konektory a u každého z nich označujeme daný významový typ a dosah diskurzního vztahu zvlášť (srov. např. *ale potom*).

Některé konektory mohou být zároveň **modifikované** – srov. např. *hlavně proto*; *právě kvůli tomu*; *z jednoduchých důvodů*; *vlastně proto*, *že*; *především pak*; *jediným důvodem bylo*; *s jasným cílem*, *možná díky tomu* apod. Modifikace se přitom (vzhledem k jejich lexikálně-syntaktické povaze) vyskytuje častěji u sekundárních konektorů.

V některých případech pak může být obtížnější určit, zda se u víceslovného konektoru jedná o konektor složený či modifikovaný. Rozlišovacím kritériem je tedy to, že složené konektory se skládají z více konektorů (tj. všechny části složeného konektoru jsou samy o sobě konektorem – srov. *a tak*, *a proto*, *ale naopak*, *a dokonce*), modifikované konektory naopak obsahují výraz, jehož hlavní funkce není spojovací, ale modifikační, tj. svým

významem upřesňuje příslušný diskurzí vztah (říká například, že je daná podmínka hlavní, příčina jednoduchá, cíl jasný apod.). Význam těchto výrazů je tedy především hodnotící (srov. *hlavně, právě, jednoduchý, jediný, jasný...*) či modální (např. *možná, vlastně...*).

V naší práci tedy vymezujeme **tři základní opozice**, podle kterých je možné diskurzí konektory dále členit. Konektory mohou být:

a) primární (*však, a, nebo, ale, proto...*) **či sekundární** (*důvodem je, kvůli tomu, s jasným cílem*) – hlavním rozlišovacím kritériem je v tomto případě míra gramatikalizace daného výrazu;

b) složené (*a tudíž, a kvůli tomu, ale naopak*) **či nesložené** (*nebo, ale, z tohoto důvodu*) – složené konektory se skládají z více konektorů, nesložené se skládají (ze synchronního hlediska) z jednoho konektoru (nesložené konektory mohou být jednoslovné (*ale*) i víceslovné (*z tohoto důvodu*)); složené konektory se v textu mohou vyskytovat vedle sebe (*a proto*), nebo odděleně (*bud'nebo, když_tak, v případě, že_pak*), v tomto případě tvoří tzv. korelativní páry;

c) modifikované (*hlavně proto, s jasným cílem, jedině kvůli tomu, především vinou toho*) **či nemodifikované** (*proto, s cílem, kvůli tomu, vinou toho*) – modifikované konektory obsahují hodnotící či modální výraz, který upřesňuje význam příslušného diskurzí vztahu.

Uvedené aspekty se mohou vzájemně kombinovat – tj. např. složený konektor může být zároveň modifikovaný apod. (srov. *a hlavním důvodem bylo, a právě proto*). Ve složený konektor se mohou spojovat konektory primární (*a proto*) i primární a sekundární (*a z tohoto důvodu*).

7.1.2 Hranice mezi primárními a sekundárními konektory

Sekundární konektory můžeme obecně chápat jako jazykové prostředky, které jsou z jedné strany ohraničené univerzální platností ukazatelů diskurzí vztahů v textu (čímž se odlišují od volných konektivních frází) a z druhé strany (ne)gramatikalizací (čímž se liší od gramaticky i lexikálně ustálených konektorů primárních).

Jednotlivé konektivní prostředky si tak můžeme představit na následující škále či ose zobrazující implicitní a explicitní diskurzní vztahy:

(univerzálnost)

implicitnost → **volné konektivní fráze** → **sekundární konektory** → **primární konektory**
(gramatikalizace)

Jeden pól osy tvoří primární (tj. ustálené, gramatikalizované) konektory, druhý výrazy bez konektivní funkce (diskurzní vztah je pouze implicitní). Mezi těmito dvěma póly stojí sekundární konektory (negramatikalizované, univerzální struktury) a volné konektivní fráze (negramatikalizované, neuniverzální struktury).

Zároveň je podle našeho názoru hranice mezi primárními a sekundárními konektory prostupná. Mezi sekundárními konektory můžeme vysledovat různé podskupiny, které jsou svou podstatou a chováním blízké konektorům primárním. Některé z nich mohou tuto hranici překročit a stát se konektory primárními.

Velkou skupinu sekundárních konektorů tvoří předložkové fráze, tj. struktury obsahující sekundární předložku (např. *díky*, *kvůli*, *vinou* atd.) a anaforický výraz odkazující k prvnímu diskurznímu argumentu. Jako sekundární konektory tedy v textu působí spojení typu *díky tomu*, *kvůli tomu*, *vinou toho* atd. Tyto sekundární konektory jsou podle našeho názoru blízké konektorům primárním.

Podobnou strukturu má z historického hlediska např. primární konektor *proto* vzniklý z předložky *pro* a zájmena *to*. Původně byl tedy tento konektor dvouslovný, posléze prošel gramatikalizací a ustálil se na jednoslovné formě *proto* s primární konektivní funkcí. Produktivnost tohoto tvoření dokládají protějšky českého *proto*, které vznikly podobným způsobem – srov. např. anglický konektor *therefore* (vzniklý spojením slova *there* a předložky *fore*, staroanglické a středoanglické kolaterální formy předložky *for*; používaný ve významu ‚v důsledku toho‘) či podobný vývoj u konektorů *daarvoor* z holandštiny, *dafür* z němčiny či *derfor* z dánštiny. Obdobný vývoj tohoto konektoru tedy proběhl v několika jazycích, a není tudíž vázaný pouze na jeden konkrétní jazyk. Podobný vývojový proces by tedy mohl nastat i u dalších konektivních struktur, především u těch, které mají podobnou strukturu, jako měl původně konektor *proto*.

Z tohoto důvodu nechápeme primární a sekundární konektory jako dvě striktně oddělené kategorie, ale jako kategorie s prostupnou hranicí.

7.1.3 Vznik primárních konektorů

Vzhledem k předchozím úvahám o prostupnosti hranic mezi primárními a sekundárními konektory a procesu ustalování konektorů jsme se dále zabývali otázkou vzniku primárních konektorů obecně. Na základě prací J. Zubatého a J. Bauera (zabývajících se vznikem vybraných spojek a částic v češtině, např. *a*, *ale*, *avšak*, *přece* apod.) vymezujeme primární konektory jako výrazy vzniklé původně z jiných slovních druhů, tj. dané výrazy či spojení nabývaly spojovací funkci postupně – v procesu vznikání primárních konektorů docházelo ke ztrátě či oslabování původního lexikálního významu příslušného slova (či jednotlivých složek daného spojení).

Podstatou vzniku primárních konektorů je tedy podle našeho názoru postupné ustalování formy i funkce dané přechodem těchto výrazů či spojení (patřících zpravidla mezi jiné slovní druhy) ke spojkám či strukturujícím částicím a jejich následná gramatikalizace, přičemž tento proces není u některých z nich zatím dokončen.

Podobný proces formování primárních konektorů můžeme sledovat i v jiných jazycích – srov. např. anglický konektor *because* (*protože*) vzniklý ze spojení „by case“ (doslovně ‚podle / vzhledem k případu‘), konektor *nevertheless* (*nicméně*) z víceslovné fráze „never the less“ (ve významu ‚o nic méně‘, analogické k frázi „never the worse“ (ve významu ‚o nic horší/hůře‘)) či primární konektor *for* (*nebot*) vzniklý pravděpodobně ze staroanglických frází typu „for þon þy“ (lit. ‚for the (reason) that‘ – ‚z toho důvodu, že‘) apod. (viz Harper, 2001).

Z tohoto hlediska bychom se mohli dívat na sekundární konektory (tj. výrazy typu *to je důvod*, *proč*) jako na jazykové prostředky, u kterých můžeme sledovat proces postupného ustalování, který ovšem zatím není v takové fázi, že by mohlo dojít k úplné gramatikalizaci. Diskurzívní konektory tedy nechápeme jako uzavřenou množinu výrazů, ale jako škálu jazykových prostředků mapující proces jejich gramatikalizace.

7.1.4 Oslabování lexikálního významu u sekundárních konektorů

V rámci procesu ustalování jistého výrazu či spojení jako primárního konektoru dochází k oslabování či úplné ztrátě jeho původního lexikálního významu (*přece* ztratilo význam ‚dále kupředu‘, *a* význam ‚hle‘ apod.). Postupným procesem gramatikalizace tedy dané výrazy či spojení ztrácejí svoji plnovýznamovost (jak jsme popsali výše, spojky a částice, které považujeme za primární konektory, vznikaly z jiných slovních druhů – především příslovcí a citoslovcí, z předložkových frází apod.). Původně se tedy jednalo

o plnovýznamová slova či o spojení s plnovýznamovou složkou. Lexikální význam byl ovšem v těchto případech oslaben a dané výrazy se v jazyce ustálily jako synsémantika (příp. jsou stále v procesu ustalování, viz některá příslovce či relativní zájmena s konektivní funkcí).

K postupnému oslabování lexikálního významu dochází také u sekundárních konektorů, tj. postupně se stírají rozdíly mezi plnovýznamovostí a neplnovýznamovostí. Jednotlivé složky sekundárních konektorů postupně ztrácejí svůj původní věcný význam a nový (tj. konektivní) nabývá celá struktura. Jinými slovy, u příslušných autosémantických složek je oslabován jejich původní význam ve prospěch celého spojení. Některé z nich se v jazyce ustalují natolik, že jsou významově zcela nahraditelné primárními konektory – srov. např. sekundární konektory *z toho důvodu*, *z té příčiny*, *pro tu příčinu* a primární konektor *proto* – daná slovní spojení uvádí např. *Slovník spisovného jazyka českého* (1989) jako významové ekvivalenty slova *proto*. Příslušné autosémantické složky samy o sobě konektivní funkci nemají a v textu mohou vystupovat samostatně v roli plnovýznamových slov.

V kategorii sekundárních konektorů tedy dochází k postupnému ustalování formy, k ustalování konektivní funkce a k oslabování vlastního věcného významu jednotlivých složek příslušné struktury. Tyto struktury pak mají jako celek v textu platnost sekundárních diskurzních konektorů.

7.2 Výsledky a evaluace anotací PDT – primární konektory, sekundární konektory, volné konektivní fráze

Výsledky anotace ukázaly, že Pražský závislostní korpus obsahuje v současné době 21 416 diskurzních vztahů vyjádřených explicitně (tj. pomocí konkrétních jazykových prostředků). Těmito prostředky jsou z 94,6 % primární konektory, z 5,4 % sekundární konektory. Z výsledků je patrné, že označení primární a sekundární je pro dané výrazy v tomto ohledu příznačné – k vyjadřování diskurzních vztahů v textu slouží primárně (tj. z 94,6 %) primární konektory. Počet sekundárních konektorů ale není zanedbatelný. Sekundární konektory obohatily diskurzní anotaci o 1 161 nových vztahů.

Anotace sekundárních konektorů může dále usnadnit výzkum implicitních diskurzních vztahů v textu (tj. vztahů, které nejsou v povrchové rovině vyjádřeny žádným jazykovým prostředkem). Díky dokončené anotaci primárních i sekundárních konektorů bude možné snadněji vyhledávat implicitní vztahy podle míst, ve kterých se nevyskytuje žádný primární

ani sekundární konektor, ale u kterých přítomnost diskurzního vztahu předpokládáme (vzhledem k tomu, že text je chápán jako síť vzájemně propojených vztahů, ve kterých žádná jednotka nezůstává osamocena).

Podle výsledků současné anotace obsahuje PDT dále 151 výskytů volných konektivních frází. Obecně se tedy ukazuje, že pro mluvčího je patrně jednodušší použít pokud možno co nejustálenější (tj. gramatikalizovaný) a většinou také kratší, mnohdy jednoslovný výraz. Mluvčí má tedy (ve většině případů) na výběr různé typy konektivních prostředků (s různou mírou ustálenosti a lexikálního „naplnění“). Nejčastěji přitom volí ten, který je z jeho pohledu při výběru nejjednodušší, tj. nejméně namáhavý při tvorbě textu. Zároveň však mluvčí volí konektivní prostředky s ohledem na čtenáře (adresáta), tj. aby čtenář co nejlépe a nejsnadněji interpretoval obsah celého sdělení, tedy i příslušné diskurzní vztahy. Existují případy, kdy je mluvčí právě naopak nucen užít konektivní prostředek s konkrétnějším lexikálním naplněním, který do jisté míry opakuje část předchozího kontextu (např. pokud jsou od sebe diskurzní argumenty v textu příliš vzdáleny či odděleny jiným textem⁶⁰), aby čtenář lépe pochopil, k jakému úseku textu daný konektivní prostředek odkazuje, a aby se tak zabránilo víceznačností. Volba konektivního prostředku se tedy v první řadě řídí snahou o to, aby bylo sdělení srozumitelné pro adresáta.

Z hlediska syntaktické povahy diskurzních argumentů převažují u sekundárních konektorů a volných konektivních frází v PDT vztahy mezi dvěma slovesnými argumenty (v 76 %) – na výsledku se ovšem odráží skutečnost, že ne všechny konektivní struktury dovolují nominalizaci druhého argumentu (srov. např. předložkové fráze, některá ustálená spojení typu *stručně/jednoduše/obecně... řečeno* apod.). K nominalizaci druhého diskurzního argumentu tak tíhnou především určité struktury (obsahující predikaci), např. *příčinou je, důsledkem je, výsledkem je, podmínkou je* apod. (u uvedených struktur v PDT nominalizace druhého argumentu dokonce převažuje, a to celkově v 80 % výskytů). Je tedy zřejmé, že struktura sekundárních konektorů má v tomto případě přímý vliv na syntaktickou realizaci druhého diskurzního argumentu.

Spolehlivost anotace sekundárních konektorů v PDT (která proběhla především pro účely této práce) ukazuje míra anotátorské shody. Shoda na rozpoznání diskurzního vztahu

⁶⁰ Srov. *Po každé směně jsme je [hlavy, M. R.] museli ostříit. Po 14 dnech nepřetržitého doladování byl stroj již k použití. Úspora, vzniklá koupí našich hlav, se však změnila v pasivum, neboť technik seřizoval stroj místo tří dnů čtrnáct. Kvůli častějšímu broušení hlav vzrostly i náklady na provoz zařízení.*

vyjádřeného sekundárním konektorem dosáhla 0,7 míry F1 a shoda na rozpoznání sémantického typu diskurzního vztahu 82 % (resp. 0,78 Cohenovy kappy). Výsledky mezinotátorské shody pro sekundární konektory jsou srovnatelné s obdobným měřením pro primární konektory (viz Poláková a kol., 2013) – srov. 0,83 F1 pro rozpoznání diskurzního vztahu a 77 % (resp. 0,71 C. k.) pro rozpoznání sémantického typu vztahu.

Výsledky anotace sekundárních konektorů a volných konektivních frází v PDT pro češtinu jsme srovnali s anotací podobných výrazů (nazývaných souhrnně „altlexy“) pro angličtinu v korpusu Penn Discourse Treebank (PDTB). Zatímco v PDT tvoří sekundární konektory a volné konektivní fráze 6,1 % (v rámci všech explicitně vyjádřených diskurzních vztahů), v PDTB pouze 3,3 %. Rozdílný počet těchto jazykových prostředků zřejmě není způsoben odlišností či specifičností češtiny a angličtiny, ale patrně do jisté míry odráží skutečnost, že dané výrazy nejsou v PDTB anotovány zcela konzistentně (srov. Prasadová a kol., 2010). Sami autoři očekávají, že se počet některých „altlexů“ po detailnější anotaci může dvakrát až třikrát zvýšit.

7.3 Nejčastěji používané primární a sekundární konektory v PDT

Z vyhodnocení výsledků anotace vyplývá, že výskyty jednotlivých primárních konektorů značně převyšují výskyty sekundárních konektorů. Sekundární konektory se v reálných textech navíc objevují v mnoha variantách, tj. ne v jedné ustálené formě jako konektory primární (srov. např. *to je důvod, proč; to je důvodem; je to důvodem toho, že* apod.). Zřetelně se zde tedy projevuje skutečnost, že primární konektory již prošly gramatikalizací, tj. jsou formálně i funkčně ustálené, zatímco konektory sekundární jsou zatím stále v procesu ustalování.

Výsledky ukazují, že nejpočetnějším primárním konektorem v PDT je konektor *a* se 5 765 výskyty, nejpočetnějším sekundárním konektorem výraz *dodal* se 121 výskyty. Sloveso *dodat/dodávat* tvoří spolu s dalšími slovesy říkání (tj. slovesy uvozujícími (značenou i neznačenou) přímou či nepřímou řeč) typu *upřesnit/upřesňovat, doplnit/doplňovat, pokračovat* apod. zcela unikátní kategorii. Od ostatních primárních i sekundárních konektorů se liší tím, že v sobě spojují dvě složky – diskurzní argument, tj. to, že někdo něco říká, a diskurzní konektor – tj. to, že to někdo říká v reakci na něco (srov. *dodal = a řekl*). Obecně

tedy můžeme říci, že daná slovesa říkání odkazují v textu k jinému slovesu říkání (*řekl – dodal, řekl – upřesnil, řekl – pokračoval* apod.) či jinak vyjádřenému uvození přímé či nepřímé řeči (*podle jeho slov – doplnil* atd.), přičemž mezi vlastními obsahy příslušných přímých či nepřímých řečí je další, samostatný diskurzivní vztah.

7.4 Srovnání sémantických typů diskurzivních vztahů pro primární a sekundární konektory

Výsledky anotace dále ukázaly, že rozložení jednotlivých sémantických typů diskurzivních vztahů je u primárních a sekundárních konektorů velice podobné. Diskurzivní vztahy s vysokým počtem výskytů (u primárních i sekundárních konektorů) jsou vztahy konjunkce, příčiny a důsledku, podmínky, přípustky, nesoučasnosti (předčasnosti a následnosti) a konfrontace. Naopak vztahy s nejnižším (či velice nízkým) počtem jsou u obou typů konektorů tzv. nepravé (pragmatické) vztahy, tj. vztahy nepravé příčiny a důsledku, nepravé podmínky a výsledku podmínky a nepravý kontrast.

Nižší čísla výskytů pragmatických vztahů ovšem neznamenaají, že by se v reálných textech vyskytovaly s tak nízkou frekvencí, ale spíše to, že jim v PDT nebyl zatím věnovaný dostatečný prostor (klasifikace diskurzivních vztahů se prozatím zaměřovala spíše na vztahy sémantické). Dalším krokem při výzkumu koherence textu by tak mohl být pohled z vyšší perspektivy, z pohledu komunikačních strategií autora.

Primární a sekundární konektory se naopak liší u vztahů opozice a explikace. Zatímco vztah opozice je u primárních konektorů na druhém místě (tj. s druhým nejvyšším počtem výskytů), u sekundárních konektorů se tento typ diskurzivního vztahu objevuje velice zřídka. Na druhou stranu vztah explikace je u sekundárních konektorů čtvrtým nejpočetnějším typem diskurzivního vztahu, zatímco u konektorů primárních je až v druhé polovině tabulky četnosti vztahů (tj. mezi vztahy s nižším počtem výskytů).

Z výsledků analýzy dále vyplynulo, že téměř u všech sémantických typů diskurzivních vztahů vyjádřených sekundárními konektory převažují slovesné argumenty, u vztahu konjunkce a vztahu podmínky a výsledku podmínky (tj. dvou ze tří nejpočetnějších vztahů) dokonce v 96 %, resp. 92 % všech výskytů. U vztahů s jedním nominálním argumentem převažují pouze dva sémantické typy – vztah restriktivní opozice (79 %) a exemplifikace (70 %). Diskurzivní vztahy specifikace a příčiny a důsledku jsou v rámci slovesných a jmenných argumentů vyrovnanější (srov. 58 % vs. 42 % u vztahu specifikace a 60 %

vs. 40 % u vztahu příčiny a důsledku). U ostatních diskurzních typů převažují vztahy mezi slovesnými argumenty zcela výrazně.

Celkový počet výskytů sekundárních konektorů ve spojení s nominalizovaným diskurzním argumentem je 237. Je zajímavé, že 136 případů z tohoto počtu (tedy více než polovina) vyjadřuje vztah příčiny a důsledku. Počet výskytů tohoto typu vztahu ve spojení s nominalizovaným argumentem tak značně převyšuje ostatní (druhým nejpočetnějším typem vztahu ve spojení s nominalizovaným argumentem je exemplifikace s 21 výskyty). Vysvětlení může spočívat v tom, že vztah příčiny a důsledku je v PDT chápán v širším smyslu, tj. nerozlišují se zde podtypy, jako je např. příčina, důvod, důsledek apod. Nicméně u vztahů mezi dvěma slovesnými argumenty vztah příčiny a důsledku tímto způsobem nedominuje (tvoří přibližně pětinu z celkového počtu diskurzních vztahů mezi slovesnými argumenty, má 208 výskytů mezi 924 případy).

Obecně můžeme shrnout, že všechny sémantické typy diskurzních vztahů jsou v PDT vyjadřovány častěji primárními konektory (většinou ve více než 90 % všech výskytů) – ani u jednoho typu vztahu tedy nepřevážil výskyt sekundárních konektorů. Nepatrně vyšší procentuální zastoupení sekundárních konektorů (ve srovnání s ostatními typy vztahů) mají pouze vztahy explikace (sekundární konektory tvoří 22,7 %), exemplifikace (16,9 %) a generalizace (16,7 %), i když i v těchto případech zastoupení primárních konektorů výrazně převažuje.

7.4.1 Možnosti rozšíření klasifikace diskurzních vztahů vyjadřovaných sekundárními konektory

Klasifikace sémantických typů diskurzních vztahů pro sekundární konektory byla převzata z anotace primárních konektorů. V průběhu anotace se ovšem ukázalo, že sekundární konektory vyjadřují i jiné typy vztahů, než se objevují ve zmíněné klasifikaci. Anotátoři upozornili celkem na tři: a) vyplývání (či vyvození výsledků – vyjadřované sekundárními konektory *z toho vyplývá, že; v každém případě; výsledkem bylo; výsledkem je*), b) vyvození závěru (*závěr zněl, že; došel k závěru, že; jednoznačný závěr je, že; dospělo k tomuto závěru*), c) zřetel (*v tomto směru, se zřetelem k tomu*).

Klasifikace diskurzních vztahů by tak (na základě zkušenosti s reálnými texty) měla být pro sekundární konektory rozšířena.

7.5 Srovnání diskurzních konektorů v rámci mezivětných a vnitrovětných vztahů

V rámci analýzy primárních a sekundárních konektorů jsme dále zjišťovali, jaký je u nich poměr mezivětných a vnitrovětných vztahů, tj. zda primární a sekundární konektory, případně některé sémantické typy diskurzních vztahů tíhnou spíše k mezivětnému či vnitrovětnému vyjádření. Z analýzy jsme vyvodili následující závěry.

1) Primární konektory vyjadřují častěji vnitrovětné diskurzní vztahy než vztahy mezivětné (srov. 70 % vs. 30 %), naopak sekundární konektory vyjadřují častěji vztahy mezivětné (srov. 63 % vs. 37 %).

Pokud je tedy diskurzní vztah vyjádřen například spojkou *a*, je její častější užití *Mám rýmu a nepůjdu do školy.* než *Mám rýmu. A nepůjdu do školy.* Naopak pokud je diskurzní vztah vyjádřen například spojením *z tohoto důvodu*, je častěji užito ve sledu vět typu *Mám rýmu. Z tohoto důvodu nepůjdu do školy.* než v souvětí typu *Mám rýmu, z tohoto důvodu nepůjdu do školy.*

Obecně (v absolutních číslech) jsou ale diskurzní vztahy vnitrovětné i mezivětné častěji vyjadřovány konektory primárními. Existují však výjimky, viz následující bod 2).

2) Celkově jsou mezivětné i vnitrovětné diskurzní vztahy vyjadřovány především primárními konektory (celkový počet primárních konektorů v PDT je více než 17krát vyšší než celkový počet sekundárních konektorů). I přesto pro vyjádření vztahu podmínky dominuje podle dat PDT užití sekundárního konektoru (v absolutních počtech výskytů).

Podle PDT se tak například častěji užívá *Půjdu ven. V případě, že nebude pršet.* než *Půjdu ven. Jestliže nebude pršet.*

3) U některých typů diskurzních vztahů vyjádřených sekundárními konektory může mít vliv na to, zda budou realizovány mezivětně či vnitrovětně, také syntaktická povaha daných diskurzních argumentů. U vztahu restriktivní opozice (signalizované např. spojenými *výjimkou je* či *výjimku tvoří*) mezi dvěma slovesnými argumenty převažuje vyjádření vnitrovětné nad mezivětným. U vztahů restriktivní opozice s jedním nominálním diskurzním argumentem naopak převládá vyjádření mezivětné. (Ačkoli počet nalezených výskytů v PDT je v tomto případě poměrně nízký, popisovaný jev je tedy třeba chápat pouze jako možnou tendenci.)

Podle PDT se tak například častěji se užívá *Fotbalové mužstvo nedostalo žádné nové posily*. Výjimkou je *příchod nadějného záložníka*, než *Fotbalové mužstvo nedostalo žádné nové posily*, výjimkou je *příchod nadějného záložníka*.

Ale zároveň se podle PDT častěji používá například *Fotbalové mužstvo nedostalo žádné nové posily*, výjimkou je, že *přišel nadějný záložník*, než *Fotbalové mužstvo nedostalo žádné nové posily*. Výjimkou je, že *přišel nadějný záložník*.

7.6 Srovnání sekundárních konektorů v psané a mluvené češtině

Na základě analýzy vybraných konektivních struktur v Pražském závislostním korpusu (obsahujícím psané texty) a v korpusu mluvené češtiny DIALOG můžeme vyvodit následující závěry.

1) Počet vybraných struktur s konektivní funkcí je vyšší v Pražském závislostním korpusu, tj. v psaných textech (PDT obsahuje 252 výskytů daných struktur s konektivní funkcí, tj. 49 % z celkového počtu; DIALOG 180 výskytů, tj. 27 %). Vysvětlení může spočívat v tom, že některé konektivní struktury typu *toto počínání je důvodem, proč; své chování odůvodnil tím, že; důvodem poklesu je* apod. mohou působit formálněji (než například neutrální konektory, jako jsou *proto, protože* aj.), a proto se objevují častěji v psaných textech.

2) V psané češtině je obecně vyšší tendence k nominalizaci (či kondenzaci) druhého diskurzivního argumentu než v češtině mluvené. Tj. příklady typu *důvodem je jeho napadení ředitele* (oproti slovesnému vyjádření *důvodem je, že napadl ředitele*) jsou častější v psaných textech.

3) V psané i mluvené češtině se častěji vyskytují univerzální konektivní struktury (tj. sekundární konektory) než neuniverzální (tj. volné konektivní fráze). V mluvené češtině je přitom rozdíl v počtu univerzálních a neuniverzálních struktur zcela jednoznačný – v průměru se v korpusu DIALOG vyskytuje 96 % univerzálních a 4 % neuniverzálních struktur (obsahujících zkoumaná jádrová slova), v korpusu PDT je poměr poněkud vyrovnanější – 62 % vs. 38 %.

Obecně tedy můžeme říci, že konektivní struktury typu *důvodem pro propuštění bylo, příčinou poklesu je, toto počínání je důvodem, proč* jsou typičtější pro psané texty. Psaná čeština se tedy v tomto ohledu vyznačuje větší lexikální rozmanitostí než čeština mluvená.

4) Obdobně se konektivní struktury se slovem *příčina* vyskytují v korpusu DIALOG pouze ve spojení se slovesem *být* (tj. např. *hlavní příčinou je; příčina je jednoduchá; příčiny jsou nejrůznější; příčiny jsou ty, že* apod.), zatímco v PDT i s jinými slovesy (srov. *vidět příčinu v tom, že; hledat příčinu v tom, že; označit za příčinu*). Z tohoto hlediska je tedy spektrum konektivních struktur v psaných textech rozmanitější než v textech mluvených.

5) Konektivní struktury obsahující sloveso *souviset* (typu *s tím souvisí, to souvisí s* apod.) se v mluvené češtině používají především (v korpusu DIALOG v 90 % vztahů vyjádřených konektivní strukturou se slovesným základem *souviset*) k prostému, volnějšímú navazování dvou různých témat, přičemž význam daného diskurzního vztahu je spíše vágní, zamlžený – srov. např. *s tím souvisí i druhá věc, to souvisí s celou řadou dalších otázek, to souvisí eště s jednou věcí, s tím souvisí třetí problém* apod.

V psané češtině jsou diskurzní vztahy vyjádřené slovesnými frázemi typu *s tím souvisí* naopak ve většině případů „pevnější“, významově jasnější či zřetelnější (v korpusu PDT jsou tyto vztahy zastoupeny 75 %). Nevyjadřují tedy prosté navazování, ale mohou se významově blížit např. vztahům příčiny a důsledku či explikace.

7.7 Testování vhodnosti sekundárních konektorů pro různé kontexty – princip univerzálnosti

Diskurzní konektory vymezujeme v této práci na základě jejich vhodnosti (univerzálnosti) jako konektivních prostředků v různých kontextech, tj. na základě toho, zda daný výraz může být použit jako konektivní prostředek v textu z velké míry univerzálně nebo je závislý na určitém specifickém kontextu. Pro testování univerzálnosti jsme vybrali dvě skupiny strukturně odlišných sekundárních konektorů – ty, jejichž jádro tvoří sekundární předložky *díky, kvůli, na základě, vzhledem k, vinou*, a ty, jejichž jádrem je podstatné jméno *důvod*. Jedná se přitom o konektivní spojení, která signalizují diskurzní vztah příčiny a důsledku.

Pro testování jsme zvolili sto různých kontextů z PDT, ve kterých se vyskytuje diskurzivní vztah příčiny a důsledku. Do těchto kontextů jsme posléze dosazovali jednotlivé konektivní struktury a zjišťovali, pro jaké procento kontextů jsou vhodnými ukazateli diskurzivního vztahu.

Zvolili jsme přitom padesát kontextů z PDT, ve kterých se v originálním textu objevil sekundární konektor (příp. volná konektivní fráze), a padesát kontextů s primárními konektory *proto* a *a tak*.

Na základě výsledků výzkumu jsme pro vymezení diskurzivních konektorů stanovili hranici 75 % univerzálnosti, která se ukázala být významným zlomem oddělujícím výrazy typu *z tohoto důvodu; vzhledem k těmto skutečnostem* apod. od výrazů jako *díky této skulince, díky jeho chorobě* apod. Nejvyššího stupně univerzálnosti dosáhly v našem výzkumu sekundární konektory *z uvedeného důvodu / z uvedených důvodů* (99 %), *z tohoto důvodu / z těchto důvodů* (98 %) a *kvůli tomu* (93 %). Zároveň je ovšem zapotřebí zohlednit skutečnost, že jsme příslušné výrazy testovali na publicistických textech a že na užívání a chování konektivních prostředků může mít vliv právě funkční styl zkoumaných textů.

U jednotlivých výrazů či spojení jsme zároveň zkoumali jejich frekvenci v textu v korpusu SYN2013PUB, tj. zaměřili jsme se na to, zda při určování příslušného jazykového prostředku jako diskurzivního konektoru můžeme zohledňovat také jeho frekvenci v reálných textech, příp. jaký je vztah frekvence a univerzálnosti.

Výsledky analýzy ukázaly, že nejvyšší frekvenci v reálných textech mají primární konektory – *proto* (733 198 výskytů), *a tak* (228 532) a dále předložkové fráze obsahující zájmeno *ten* – např. *kvůli tomu* (53 755), *díky tomu* (40 376), *vzhledem k tomu* (39 497). Všechny tyto výrazy mají zároveň vysokou míru univerzálnosti, tj. vhodnosti užití jako konektivního prostředku ve velkém množství různých kontextů.

Většina spojení, která dosáhla velmi nízké univerzálnosti, se zároveň v korpusu SYN2013PUB vyskytuje s malou frekvencí (např. spojení *díky této skulince* má univerzálnost 1 % a zároveň nemá žádný výskyt v SYN2013PUB). V těchto případech je tedy univerzálnost daných jazykových prostředků v souladu s jejich frekvencí v reálných textech.

Na druhou stranu se některá spojení s vysokou mírou univerzálnosti vyskytují v korpusu SYN2013PUB s poměrně nízkou frekvencí (ve srovnání např. s primárními konektory; spojení *vzhledem k těmto skutečnostem* dosáhlo univerzálnosti 88 %, v korpusu SYN2013PUB se vyskytuje ve 145 výskytích), příp. se v uvedeném korpusu nevyskytují

vůbec (srov. spojení *(právě) tato skutečnost se stala jedním z důvodů, kvůli kterým* s univerzálností 85 %).

V korpusu SYN2013PUB jsme zároveň pro ilustraci vyhledali spojení *kvůli své nemoci*, které má pravděpodobně (analogicky k testovanému *díky jeho chorobě*) nízkou univerzálnost – v reálných textech korpusu se objevilo ve 106 výskytech (tj. v textech se vyskytuje s podobnou frekvencí jako např. *vzhledem k těmto skutečnostem*).

Z těchto důvodů nebereme frekvenci jazykových prostředků v reálných textech jako hlavní kritérium pro určování hranice diskurzních konektorů (jako základní kritérium chápeme jejich univerzálnost pro použití v různých kontextech). Nicméně frekvence příslušných výrazů vypovídá o míře jejich ustálenosti, tj. v rámci kategorie diskurzních konektorů frekvence supluje hranici mezi konektory primárními a sekundárními (jak jsme popsali výše, v PDT je mezi všemi výskyty diskurzních konektorů 94,6 % primárních konektorů a 5,4 % sekundárních).

Vzhledem k tomu, že sekundární konektory často obsahují explicitní anaforickou referenci, zkoumali jsme dále to, která jména se na této pozici mohou objevit, aby měla celá struktura v textu funkci diskurzního konektoru. Zaměřili jsme se přitom na obecnější jména typu *skutečnost, činnost, opatření* apod., protože významově konkrétnější jména typu *skulinka* či *choroba* se v rámci konektivních struktur ukázala zcela jednoznačně jako kontextově závislá (tj. s univerzálností 0–2 %). Vybraná obecná jména jsme testovali ve spojení se sekundární předložkou *na základě* (tj. např. *na základě této skutečnosti, na základě této činnosti, na základě tohoto opatření* apod.), která v našem výzkumu dosáhla ve spojení se zájmenem *ten* univerzálnosti 94 %, tj. nejvíce mezi ostatními obdobnými předložkovými frázemi *kvůli tomu* (93 %), *díky tomu* (77 %), *vzhledem k tomu* (77 %) a *vinou toho* (34 %) (použití spojení *vinou toho* je limitováno úzkým sémantickým vymezením předložky *vinou* vázanou na negativní kontext, přesto dané spojení považujeme za diskurzní konektor – vzhledem k jeho analogické struktuře ke konektorům typu *kvůli tomu*).

Výsledky výzkumu ukázaly, že konektivními prostředky s obecnými jmény, které dosahují nejvyšší univerzálnosti, jsou spojení *na základě toho* (94 %), *na základě tohoto faktu / těchto faktů* (93 %) a *na základě této skutečnosti / těchto skutečností* (91 %). Tato spojení jako jediná mezi ostatními zkoumanými strukturami s obecnými jmény přesahují hranici 75 % univerzálnosti, tj. použitelnosti ve funkci konektivního prostředku v různých kontextech. Z tohoto důvodu považujeme (v rámci zkoumaného vzorku dat) za diskurzní konektory pouze takové struktury s explicitně vyjádřenou anaforickou referencí, na jejíž pozici stojí anaforické

výrazy *ten, tento fakt* či *tato skutečnost* (případně vnitrovětné *což*, viz níže), tj. např. *kvůli tomu; vzhledem k těmto skutečnostem; díky tomuto faktu; to je důvod, proč; tyto skutečnosti jsou hlavní příčinou toho, že* apod.

Výsledky výzkumu dále poukázaly na skutečnost, že některé konektivní prostředky jsou vázané syntakticky – výraz *na základě čehož* byl vhodný pouze pro 21 % vybraných kontextů. V testovaných kontextech však dominovaly mezivětné diskurzní vztahy. Když jsme daný výraz testovali znovu pouze na vztazích vnitrovětných, výsledek vzrostl na 81 %. Z tohoto důvodu považujeme spojení *na základě čehož* za sekundární diskurzní konektor vyjadřující diskurzní vztahy primárně uvnitř věty, podobně jako některé primární konektory typu *což* nebo *příčemž*.

Z výsledků výzkumu je tedy zřejmé, že vedle univerzálnosti diskurzních konektorů (tj. jejich vhodnosti v rámci velkého vzorku kontextů) je při výběru konektivních prostředků zapotřebí zohledňovat i další faktory.

Výsledky naší analýzy ukázaly, že jedním z nich může být také vzdálenost diskurzních argumentů, tj. zda jsou dané argumenty v textu vedle sebe nebo zda je mezi ně vložen další text, a je tedy zapotřebí užít (pokud hrozí nesprávná interpretace daného vztahu) konektivní prostředek s vyšší lexikální naplněností (který jednoznačně odkazuje k příslušnému diskurznímu argumentu).

Na základě našeho výzkumu tedy můžeme shrnout hlavní faktory uplatňující se při výběru vhodných konektivních prostředků. Jsou jimi vzájemná vzdálenost diskurzních argumentů, sémanticko-pragmatické aspekty konektorů (tj. např. míra jejich subjektivity), syntaktické chování konektorů (tj. vázanost na vnitrovětné či mezivětné vztahy), funkční styl příslušného textu (tj. zohlednění stylistické platnosti konektoru) a míra lexikální naplněnosti konektoru. Některé z uvedených faktorů jsou přitom vzájemně provázené. Zároveň právě kvůli těmto různým faktorům a omezením většinou nedosahují primární ani sekundární konektory úplné univerzálnosti 100 %.

Závěrem

V předkládané disertační práci jsme se zaměřili na jazykové prostředky, které se významně podílejí na strukturaci textu – tzv. diskurzní konektory. Ukázali jsme, že diskurzní konektory představují široké spektrum jazykových výrazů, od plně gramatikalizovaných po dosud neustálené. Proces jejich ustalování probíhá v jazyce již několik staletí, z tohoto hlediska nejsou diskurzní konektory uzavřenou množinou jazykových prostředků, ale jako celek procházejí neustále vývojem. Diskurzní konektory si tak můžeme představit jako množinu výrazů, která má své více méně neměnné centrum (tzv. primární konektory) a svoji periferii (tzv. sekundární konektory). Provedený výzkum těchto prostředků realizovaný na rozsáhlém jazykovém materiálu by mohl přispět ke zlepšení porozumění vztahům v rámci textu, a tedy k prohloubení obecného povědomí o celkové koherenci textu.

Literatura

- ADAMEC, Přemysl. Konektivní částice a jiné textově propojovací výrazy v současné češtině. In *Přednášky z 37. a 38. běhu LŠSS*. Praha: Univerzita Karlova, 1995, s. 59–64.
- AFANTENOS, Stergos D.; ASHER, Nicholas; BENAMARA, Farah a kol. An empirical resource for discovering cognitive principles of discourse organization: the ANNODIS corpus. In CALZOLARI, N. a kol. (eds.). *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul: European Language Resources Association (ELRA), 2012, s. 2727–2734.
- AIJMER, Karin. English Discourse Particles. Evidence from a corpus. *Studies in Corpus Linguistics 10*. Amsterdam/Philadelphia: John Benjamins, 2002.
- AL-SAIF, Amal; MARKERT, Katja. The Leeds Arabic Discourse Treebank: Annotating discourse connectives for Arabic. In CALZOLARI, N. a kol. (eds.). *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Valletta: European Language Resources Association (ELRA), 2010, s. 2046–2053.
- ARIEL, Mira. Pragmatic operators. *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press, 1994, s. 3250–3253.
- ASHER, Nicholas a kol. Modelling Strategic Conversation: The STAC Project. In BROWN-SCHMIDT, S.; GINZBURG, J.; LARSSON, S. (eds.). *Proceedings of 16th Workshop on the Semantics and Pragmatics of Dialogue*, 2012, s. 145–146.
- ASHER, Nicholas. *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer Academic Publishers, 1993.
- ASHER, Nicholas; LASCARIDES, Alex. *Logics of Conversation*. Cambridge University Press, 2003. ISBN 9780521659512.
- ASR, Fatemeh Torabi; DEMBERG, Vera. On the Information Conveyed by Discourse Markers. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistic*. Association for Computational Linguistics, Sofia, Bulgaria, 2013, s. 84–93. ISBN 978-1-937284-61-9.
- AUSTIN, John Langshaw. *How to Do Things with Words*. Cambridge, MA: Harvard University Press, 1962.
- BAJZÍKOVÁ, Eugénia. O textových konektorech. *Zborník Pedagogickej fakulty v Nitre. 23. Jazyk a literatúra*. Bratislava: Slovenské pedagogické nakladateľstvo, 1975, s. 45–55.
- BAJZÍKOVÁ, Eugénia. *Úvod do textovej syntaxe*. Bratislava, 1979.
- BAUER, Jaroslav. Podíl citoslovcí na vzniku českých spojek. *Sborník prací Filosofické fakulty brněnské university A 11*, 1963, s. 21–28.
- BAUER, Jaroslav. Relativa a spojky. *Acta Universitatis Carolinae – Philologica, Slavica Pragensia IV*, 1962, s. 221–226.
- BAUER, Jaroslav. Spojky a částice. *Sborník prací Filosofické fakulty brněnské university A 12*, 1964, s. 131–139.
- BAUER, Jaroslav. Spojky a příslovce. *Sborník prací Filosofické fakulty brněnské university A 10*, 1962, s. 29–37.
- BAUER, Jaroslav. Spojky. *Český jazyk 8*, 1958, s. 220–226.
- BAZANELLA, Carla. Phatic connectives as intonational cues in contemporary spoken Italian. *Journal of Pragmatics 14 (4)*, 1990, s. 629–647.
- BEJČEK, Eduard a kol. *Prague Dependency Treebank 2.5*. Data/software, Univerzita Karlova v Praze, MFF, ÚFAL, Praha, 2011.
- BLAKEMORE, Diane. Or-parenthicals, that is-parenthicals and the pragmatics of reformulation. *Journal of Linguistics 43 (2)*, 2007, s. 311–339. ISSN 0022-2267.

- BLAKEMORE, Diane. *Relevance and Linguistic Meaning. The Semantics of Discourse Markers*. Cambridge: Cambridge University Press, 2002.
- BLAKEMORE, Diane. *Semantic constraints on relevance*. Oxford: Blackwell, 1987.
- BLAKEMORE, Diane. *Understanding Utterances*. Blackwell: Oxford, 1992.
- BLASS, Regina. *Relevance Relations in Discourse: A Study with Special Reference to Sissala*. Cambridge: Cambridge University Press, 1990. ISBN-10 0521385156. ISBN-13 978-0521385152.
- BROWN, Gillian; YULE, George. *Discourse analysis*. Cambridge: Cambridge University Press, 1983.
- CARDOSO, Paula C. F. a kol. CSTNews – A Discourse-Annotated Corpus for Single and Multi-Document Summarization of News Texts in Brazilian Portuguese. In *Proceedings of Anais do III Workshop "A RST e os Estudos do Texto"*, 2011, s. 88–105.
- CELLE, Agnes; HUART, Ruth. *Connectives as Discourse Landmarks*. Amsterdam: John Benjamins Publishing Company, 2007. ISBN 9789027254047.
- CRIBLE, Ludivine; DUMONT, Amandine; NOTARRIGO, Ingrid. *Situational features*. Technical Report. Université Catholique de Louvain, 2014.
- ČERMÁK, František. *Frazeologie a idiomatika: česká a obecná*. Praha: Karolinum, 2007. ISBN 9788024613710.
- ČERVENKA, Miroslav. *Významová výstavba literárního díla*. 1.vyd. Praha: Karolinum, Acta Universitatis Carolinae, 1992.
- ČMEJRKOVÁ, Světlá. *Reklama v češtině. Čeština v reklamě*. Praha: Leda 2000. ISBN 8085927756.
- ČMEJRKOVÁ, Světlá; DANEŠ, František; SVĚTLÁ, Jindra. *Jak napsat odborný text*. 1. vyd. Praha: Leda, 1999. ISBN 8085927691.
- DANEŠ, František. *Intonace a věta ve spisovné češtině*. Praha: ČSAV, 1957.
- DANEŠ, František. Typy tematických poslopností v textu (na materiále českého textu odborného). *Slovo a slovesnost* 29 (2), 1968, s. 125–141. ISSN 0037-7031.
- DANEŠ, František. *Věta a text*. Praha: Academia, 1985.
- DANEŠ, František. Zum Status der Textebene. In: DANEŠ, F.; VIEHWEGER, D. (eds.). *Probleme der Textgrammatik II*. Berlin: Akademie-Verlag, 1977, s. 153–158.
- DANEŠ, František; VIEHWEGER, Dieter (eds). *Ebenen der Textstruktur*. Berlin: Akademie der Wissenschaften der DDR, Zentralinstitut für Sprachwissenschaft, 1983. ISSN: 0138-4694.
- DANLOS, Laurence et. al. Vers le FDTB: French Discourse Tree Bank. In ANTONIADIS, G. a kol. (eds.). *Actes de la conférence conjointe JEP-TALN-RECITAL*. Svazek 2: TALN. Grenoble: Association Francophone pour la Communication Parlée (AFCP) et Association pour le Traitement Automatique des Langues (ATALA), 2012, s. 471–478.
- DE BEAUGRANDE, Rober-Alain; DRESSLER, Wolfgang. *Introduction to Text Linguistics*. London: Longman, 1981.
- DEGAND, Lisbeth a kol. LOCAS-F: un corpus oral multigenres annoté. In *CMLF 2014 – 4ème congrès mondial de linguistique française*. Berlin: EDP Sciences, 2014.
- DOLEŽEL, Lubomír. *Narativní způsoby v české literatuře*. 1. vyd. Praha: Český spisovatel, 1993. ISBN 80-202-0418-0.
- DOLEŽEL, Lubomír. *O stylu moderní české prózy: výstavba textu*. 1. vyd. Praha: Nakladatelství Československé akademie věd, 1960.
- DOLK, Simone. *Patient Information Leaflets: Evaluation of the European translation policy and the MHRA suggestion to increase user's awareness of key safety issues*. Diplomová práce, 2010 [cit. 2015-05-15]. Online vydání dostupné z WWW: <<http://igitur-archive.library.uu.nl/student-theses/2010-0301-200230/UUindex.html>>.

- DRESSLER, Wolfgang. *Einführung in die Textlinguistik*. Tübingen: Niemeyer, 1972.
- FASOLD, Ralph. *Sociolinguistics of Language*. Oxford: Blackwell, 1990. ISBN 978-0-631-13825-9.
- FIRBAS, Jan. *Functional sentence perspective in written and spoken communication*. Cambridge: Cambridge University Press, 1992.
- FIRBAS, Jan. Poznámka k místu funkční perspektivy větné ve struktuře jazyka. *Otázky slovanské syntaxe 1*, 1962, s. 201–203.
- FIRBAS, Jan. *Souhra činitelů aktuálního členění větného*. Brno: Masarykova univerzita, 1978.
- FISCHER, Kerstin (ed.). *Approaches to discourse particles*. Amsterdam: Elsevier, 2006.
- FRASER, Bruce. An approach to discourse markers. *Journal of Pragmatics 14*, 1990, s. 383–395.
- FRASER, Bruce. Pragmatic markers. *Pragmatics 6 (2)*, 1996, s. 167–90.
- FRASER, Bruce. Types of English discourse markers. *Acta Linguistica Hungarica 38 (1–4)*, 1988, s. 19–33.
- FRASER, Bruce. What are discourse markers? *Journal of Pragmatics 31 (7)*, Elsevier, 1999, s. 931–952.
- GASTEL, Anna a kol. Annotation of explicit and implicit discourse relations in the TüBa-D/Z treebank. In HEDELAND, H. a kol. (eds.). *Multilingual Resources and Multilingual Applications, Proceedings of the German Society of Computational Linguistics and Language Technology (GSCL) 2011*. Hamburg: University of Hamburg, Center for Language Corpora, 2011 s. 99–104.
- GRICE, Herbert Paul. Logic and conversation. In COLE, P.; MORGAN, J. (eds.). *Syntax and Semantics 3: Speech Acts*. New York: Academic Press, 1975, s. 41–58.
- GRICE, Herbert Paul. *Studies in the way of words*. Cambridge, MA: Harvard University Press, 1989.
- GRICE, Herbert Paul. The causal theory of perception. In *Proceedings of the Aristotelian Society, Supplementary Volume 35*, 1961, s. 121–152.
- HAIJČ, Jan a kol. *Prague Dependency Treebank 2.0*. Philadelphia: Linguistic Data Consortium, 2006.
- HAIJČOVÁ, Eva; HAVELKA, Jiří; SGALL, Petr. Topic and focus, anaphoric relations and degrees of salience. In *Prague Linguistic Circle Papers / Travaux du cercle linguistique de Prague N.S.* Amsterdam: John Benjamins, 2004.
- HAIJČOVÁ, Eva; PARTEE, Barbara H.; SGALL, Petr. *Topic-Focus Articulation, Tripartite Structures and Semantic Content*. 1. vyd. Dordrecht/Boston: Kluwer Academic Publishers, 1998. ISBN 0-7923-5289-0.
- HAIJČOVÁ, Eva; SGALL, Petr; BURÁŇOVÁ, Eva. Topic-Focus Articulation and degrees of salience in the Prague Dependency Treebank. In *Formal Approaches to Function in Grammar. In honor of Eloise Jelinek Arizona*. Amsterdam/Philadelphia: John Benjamins, 2003, s. 165–177.
- HAKULINEN, A. The use of Finnish *nyt* as a discourse particle. In JUCKER, A. H.; ZIV, Y. (eds.). *Discourse markers. Description and theory*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1998, s. 83–96.
- HALLIDAY, Michael Alexander Kirkwood; HASAN, Ruqaiya. *Cohesion in English*. London: Longman, 1976.
- HANSEN, Maj-Britt Mosegaard. *The Function of Discourse Particles. A Study with Special Reference to Spoken Standard French*. Amsterdam/Philadelphia: John Benjamins, 1998.
- HARPER, Douglas. *Online Etymology Dictionary*. 2001 [cit. 2015-05-15]. Dostupné z WWW: <<http://www.etymonline.com>>.
- HARRIS, Zellig Sabbetai. Discourse Analysis. *Language 28: 1*, 1952, s. 1–30.

- HARRIS, Zellig Sabbetai. *Methods in Structural Linguistics*. Chicago: University of Chicago Press, 1951.
- HAUSENBLAS, Karel. Gramatická syntax textu. In *Referati od X. zasedanie na slovenskíte literaturni jazyci*. Skopje, 1979, s. 175–192.
- HAUSENBLAS, Karel. O studiu syntaxe běžně mluvených projevů. In *Otázky slovanské syntaxe I*. Praha, 1962, s. 313–325.
- HAUSENBLAS, Karel. On the characterization and classification of discourses. In *Travaux linguistiques de Prague 1*, 1964.
- HAUSENBLAS, Karel. Text, komunikáty a jejich komplexy. Zamyšlení pojmoslovné. *Slovo a slovesnost 45 (1)*, 1984, s. 1–7.
- HAUSENBLAS, Karel. *Výstavba jazykových projevů a styl*. Praha: Univerzita Karlova: Mir, 1971.
- HAUSENBLAS, Karel; HOFFMANNOVÁ, Jana. Od syntaxe k stavbě textu. In *Syntax a jej vyučovanie*. Bratislava, 1982, s. 427–440.
- HOBBS, Jerry R. Coherence and Coreference. *Cognitive Science 3*, 1979, s. 27–90.
- HOEY, Michael. *Patterns of Lexis in Text*. Oxford: Oxford University Press, 1991.
- HOEY, Michael. *Signalling in discourse*. Birmingham UK: English Language Research Unit, University of East Anglia, 1979.
- HOEY, Michael. *Textual Interaction: An Introduction to Written Discourse Analysis*. London: Routledge, 2001.
- HOFFMANNOVÁ, Jana. Lingvistická pragmatika a interpretace textu. *Jazykové aktuality 24*, 1987, s. 95–96.
- HOFFMANNOVÁ, Jana. Pojem relevance v současné lingvistice, v jazykovém vyučování a v pedagogické komunikaci. *Český jazyk a literatura 41*, č. 1/2, 1990/1, s. 7–15.
- HOFFMANNOVÁ, Jana. Pragmatická perspektiva a současná lingvistika. *Slovo a slovesnost 51*, 1990, s. 131–139.
- HOFFMANNOVÁ, Jana. *Sémantické a pragmatické aspekty koherence textu*. Praha: Ústav pro jazyk český, 1983.
- HOFFMANNOVÁ, Jana. *Stylistika a... Současná situace stylistiky*. Praha: Trizonia, 1997.
- HOFFMANNOVÁ, Jana. Typen der Konektoren und deren Anteil an der Organisierung des Textes. In: KOŘENSKÝ, J.; HOFFMANNOVÁ, J. (eds.). *Text and pragmatic aspects of language*, 1984, s. 101–140.
- HOVY, Eduard H. Parsimonious or profligate approaches to the question of discourse structure relations. In *Proceedings of the 5th International Workshop on Natural Language Generation*, 1990, s. 128–34.
- HRBÁČEK, Josef. *Nárys textové syntaxe spisovné češtiny*. Praha: Trizonia, 1994. ISBN 80-85573-51-2.
- HRBÁČEK, Josef. Text, promluva a nadpromluvvové útvary. *Slovo a slovesnost 52 (2)*, 1991, s. 123–129.
- CHAFE, Wallace. Cognitive constraints on information flow. In *Coherence and grounding in discourse*. Amsterdam: John Benjamins, 1987, s. 21–51.
- CHOMSKY, Avram Noam. *Knowledge of language*. New York: Prager, 1986.
- IRWIN, Judith W.; PULVER, Cynthia J. Effects of explicitness, clause order and reversibility on children's comprehension of causal relationships. *Journal of Educational Psychology 76 (3)*, 1984, s. 399–407.
- JAKOBSON, Roman. Closing statement: linguistics and poetics. In: *Style in language*. New York, 1960.
- JAKOBSON, Roman. *Poetická funkce*. 1. vyd.. Jinočany: H & H, 1995. ISBN 80-85787-83-0.
- JAWORSKI, Adam; COUPLAND, NIKOLAS (eds). *The Discourse Reader*. Routledge, 2006. ISBN-13 978-0415346320.

- JEDLIČKA, Alois; FORMÁNKOVÁ, Věra; REJMÁNKOVÁ, Miloslava. *Základy české stylistiky*. 1. vyd. Praha: Státní pedagogické nakladatelství, 1970.
- JELÍNEK, Milan. Stylistika. In *Příruční mluvnice češtiny*. Praha: Nakladatelství lidové noviny, 1995. ISBN 80-7106-134-4.
- JÍNOVÁ, Pavlína; MÍROVSKÝ, Jiří; POLÁKOVÁ, Lucie. Semi-Automatic Annotation of Intra-Sentential Discourse Relations in PDT. In *Proceedings of the Workshop on Advances in Discourse Analysis and its Computational Aspects (ADACA) at Coling 2012*, Mumbai, India, 2012, s. 43–58.
- JOHANNSEN, Anders; SØGAARD, Anders. Disambiguating explicit discourse connectives without oracles. In *The 6th International Joint Conference on Natural Language Processing (IJCNLP)*. Association for Computational Linguistics, Nagoya, Japan, 2013, s. 997–1001.
- JOHNSTONE, Barbara. *Discourse analysis*. London: Blackwell Publisher, 2002. ISBN 978-1-4051-4427-8.
- KADERKA, Petr. Pragmatika situace. *Slovo a slovesnost* 74, 2013, s. 13–40.
- KNOTT, Alistair. *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. Edinburgh: University of Edinburgh, disertační práce, 1996.
- KNOTT, Alistair; DALE, Robert. Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes* 18(1), 1994, s. 35–62.
- KOŘENSKÝ, Jan. *Komunikace a čeština*. Jinočany: H+H, 1992. ISBN 8085467925.
- KRAUS, Jiří. *Rétorika v dějinách jazykové komunikace*. 1. vyd. Praha: Academia, 1981.
- LAPSHINOVA-KOLTUNSKI, Ekaterina; KUNZ, Kerstin; AMOIA, Marilisa. Compiling a Multilingual Corpus. In MELLO, H. a kol. (eds.). *Proceedings of the VIIth GSCP-2012, CP-2012 International Conference: Speech and Corpora*. Firenze: Firenze University Press, 2012, s. 29–34.
- LONGACRE, Robert E. *The Grammar of Discourse*. New York: Plenum Press, 1983.
- MACUROVÁ, Alena; MAREŠ, Petr. *Text a komunikace: jazyk v literárním díle a ve filmu*. Praha: Univerzita Karlova, 1993.
- MANN, William C.; THOMPSON, Sandra A. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text* 8 (3), 1988, s. 243–281.
- MANN, William C.; THOMPSON, Sandra A. *Rhetorical Structure Theory: Description and Construction of Text Structures*. Defense Technical Information Center, 1986.
- MARTIN, James R. *English text: System and structure*. Amsterdam: Benjamins, 1992.
- MASCHLER, YAEL. *Discourse Markers in Bilingual Conversation*. Kingston Press, Middlesex, 2000.
- MATHESIUS, Vilém. Jazykozpytné poznámky k řečnické výstavbě souvislého výkladu. *Slovo a slovesnost* 7, 1943, s. 114–129.
- MATHESIUS, Vilém. O tak zvaném aktuálním členění větném. *Slovo a slovesnost* 5, 1939, s. 171–174.
- MEURANT, Laurence; SINTE, Aurélie. Towards a corpus of French Belgian Sign Language (LSFB) discourses. In BOLLY, C.; DEGAND, L. (eds). *Across the Line of Speech and Writing Variation. Corpora and Language in Use – Proceedings 2*. Louvain-la-Neuve: Presses universitaires de Louvain, 2013, s. 199–212.
- MILTSAKAKI, Eleni a kol. Annotating discourse connectives and their arguments. In *Proceedings of the HLT/NAACL Workshop on Frontiers in Corpus Annotation*, 2004, s. 9–16.
- MISTRÍK, Jozef. Hypersyntax a stylistika. In *Československé přednášky pro VIII. mezinárodní sjezd slavistů v Záhřebu*. *Lingvistika*. Praha: Academia, 1978, s. 213–220.
- MISTRÍK, Jozef. *Kompozícia jazykového prejavu*. 1. vyd. Bratislava: Slovenské pedagogické nakladateľstvo, 1968.

- MISTRÍK, Jozef. Skladba textu. *Slovenský jazyk a literatúra v škole* 22, (1975/6), s. 209–214.
- MISTRÍK, Jozef. *Štruktúra textu*. Československý rozhlas, 1975.
- Mluvnice češtiny* 2. Praha: Academia, 1986.
- Mluvnice češtiny* 3. Praha: Academia, 1987.
- MUKAŘOVSKÝ, Jan. *Kapitoly z české poetiky*. Praha: Melantrich, 1948.
- NEBESKÁ, Iva. Může být relevance postačujícím principem komunikace? *Slovo a slovesnost* 52, 1991a, s. 104–107.
- NEBESKÁ, Iva. Relevance and communicative norms. In *Kommunikationstagung 1989. Linguistische Studien, A. Arbeitsberichte 199*. Berlin: Akademie der Wissenschaften der DDR, Zentralinstitut für Sprachwissenschaft, 1989, s. 86–90.
- NEBESKÁ, Iva. Relevance jako hlavní princip komunikace. *Slovo a slovesnost* 52, 1991b, s. 74–76.
- NEKULA, Marek. *System der Partikeln im Deutschen und Tschechischen*. Tübingen: Max Niemeyer Verlag, 1996.
- OZA, Umangi a kol. The Hindi discourse Relation bank. In *Proceedings of the Third Linguistic Annotation Workshop*. Association for Computational Linguistics, 2009, s. 158–161.
- PALEK, Bohumil. *Cross-Reference. A Study from Hyper-Syntax*. 1. vyd. Praha: Universita Karlova, 1968.
- PEŠEK, Ondřej. *Argumentativní konektory v současné francouzštině a češtině*. České Budějovice: Jihočeská univerzita v Českých Budějovicích, 2011. ISBN 978-80-7394-305-9.
- PETÖFI, János Sándor. On the problems of co-textual analysis of texts. In *Proceeding COLING '69 Proceedings of the 1969 conference on Computational linguistics*, 1969, s. 1–45.
- PETÖFI, János Sándor. *Per una teoria semiotica integrata del testo. Dalla TeSWeST alla Testologia Semiotica*. Carabba, 2011. ISBN-10 8863442037. ISBN-13 978-8863442038.
- PETÖFI, János Sándor. *Scrittura e interpretazione. Introduzione alla Testologia Semiotica dei testi verbali*. Roma: Carocci, 2004. ISBN 8843031333.
- PETÖFI, János Sándor. Towards an Empirically Motivated Grammatical Theory of Verbal Texts. *Studies in Text Grammar* 19, 1973, s. 205–275.
- PETÖFI, János Sándor. *Transformationsgrammatiken und eine ko-textuelle Texttheorie. Grundfragen und Konzeptionen*. Frankfurt am Main: Athenäum Verlag, 1971.
- POKORNÝ, Jan. *Lingvistická antropologie: jazyk, mysl a kultura*. Grada Publishing, 2010.
- POLÁKOVÁ, Lucie a kol. Introducing the Prague Discourse Treebank 1.0. In *Proceedings of the 6th International Joint Conference on Natural Language Processing*. Asian Federation of Natural Language Processing, Nagoya, Japan, 2013, s. 91–99. ISBN 978-4-9907348-0-0.
- POLÁKOVÁ, Lucie a kol. *Manual for Annotation of Discourse Relations in Prague Dependency Treebank*. Technical report no. 2012/47, Prague: ÚFAL MFF UK, 2012a.
- POLÁKOVÁ, Lucie a kol. *Prague Discourse Treebank 1.0*. Data/software, Prague: ÚFAL MFF UK, 2012b.
- PRASAD, Rashmi a kol. The Biomedical Discourse Relation Bank. *BMC Bioinformatics*, 2011, s. 1–18.
- PRASAD, Rashmi a kol. *The Penn Discourse TreeBank 2.0 Annotation Manual*. 2007 [cit. 2015-05-15]. Online vydání dostupné z WWW: <<http://www.seas.upenn.edu/~pdtb/PDTBAPI/pdtb-annotation-manual.pdf>>.

- PRASAD, Rashmi a kol. The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation*. Marrakech, Morocco, 2008.
- PRASAD, Rashmi; JOSHI, Aravind; WEBBER, Bonnie. Realization of Discourse Relations by Other Means: Alternative Lexicalizations. In *Coling 2010: Posters (August 2010)*, 2010, s. 1023–1031.
- PRASAD, Rashmi; WEBBER, Bonnie; JOSHI, Aravind. Reflections on the Penn Discourse TreeBank, Comparable Corpora, and Complementary Annotation. *Computational Linguistics* 40 (4), 2014, s. 921–950.
- Pravidla českého pravopisu*. Praha: Academia, 1993.
- PRINCE, Ellen F. Discourse analysis: a part of the study of linguistic competence. In *Linguistics: the Cambridge survey, Vol. II*. Cambridge: Cambridge University Press, 1988, s. 164–82.
- Příruční mluvnice češtiny*. Praha: Nakladatelství Lidové noviny, 1995.
- QUIRK, Randolph a kol. *A Comprehensive Grammar of the English Language*. London: Longman, 1985.
- REDEKER, Gisela. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14(3), 1990, s. 367–381.
- REDEKER, Gisela. Linguistic markers of discourse structure. *Linguistics* 29 (6), 1991, s. 1139–72.
- REESE, Brian. *Reference Manual for the Analysis and Annotation of Rhetorical Structure (Version 1.0)*. Departments of Linguistics and Philosophy University of Texas at Austin, 2007.
- RIEBER, Steven. Conventional Implicatures as Conventional Implicatures. *Linguistics and Philosophy* 20 (1), 1997, s. 50–72.
- ROUCHOTA, Villy. Discourse connectives: what do they link? *UCL Working Papers in Linguistics* 8, 1996, s. 1–15.
- RYSOVÁ, Magdaléna. Alternative Lexicalizations of Discourse Connectives in Czech. In CALZOLARI, N. a kol. (eds.). *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul, Turecko: European Language Resources Association (ELRA), 2012, s. 2800–2807. ISBN 978-2-9517408-7-7.
- RYSOVÁ, Magdaléna. Jazykové prostředky vyjadřující textové vztahy v češtině a jejich zpracování v Pražském závislostním korpusu. *Bohemistika* 13 (1), 2013b, s. 57–71. ISSN 1642–9893.
- RYSOVÁ, Magdaléna. K explikativním vztahům v češtině. In NESWETHA, J. a kol. (eds.). *Grenzüberschreitungen – Polnische, tschechische und deutsche Sprache, Literatur und Kultur: Beiträge zur VIII. Internationalen Westslawistischen interFaces-Konferenz in Leipzig*. Hildesheim, Zürich, New York: Olms, 2013a, s. 331–341. ISBN-13 978-3487150048. ISBN-10 3487150042.
- RYSOVÁ, Magdaléna. Verbs of Saying with a Textual Connecting Function in the Prague Discourse Treebank. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavík, Island: European Language Resources Association (ELRA), 2014, s. 930–935. ISBN 978-2-9517408-8-4.
- RYSOVÁ, Magdaléna. Způsoby vyjadřování diskurzivních vztahů v češtině a jejich anotace v Pražském závislostním korpusu. In CZOK, M. a kol. (eds.). *Między regionalizmami a kosmopolityzmem. Polska, niemiecka i czeska literatura, język i kultura. Materiały IX Międzynarodowej Konferencji interFaces we Wrocławiu i w Karpaczu*. Leipzig –

- Dresden – Wrocław: ATUT, Neisse Verlag, 2015, s. 233–241. ISSN 1426-7241. ISBN 978-83-7977-110-3. ISBN 978-3-86276-161-6.
- RYSOVÁ, Magdaléna. Způsoby vyjadřování vztahu příčiny a důsledku v textu – alternativy pro „proto“ a „protože“. *Didaktické studie* 5 (2), 2013c, s. 39–53. ISSN 1804-1221.
- RYSOVÁ, Magdaléna; MÍROVSKÝ, Jiří. Use of Coreference in Automatic Searching for Multiword Discourse Markers in the Prague Dependency Treebank. In *Proceedings of The 8th Linguistic Annotation Workshop (LAW-VIII)*. Dublin, Ireland: Dublin City University, 2014. s. 11–19. ISBN 978-1-941643-29-7.
- RYSOVÁ, Magdaléna; RYSOVÁ, Kateřina. The Centre and Periphery of Discourse Connectives. In AROONMANAKUN, W. a kol. (eds.). *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing (PACLIC 28)*. Bangkok, Thailand: Department of Linguistics, Faculty of Arts, Chulalongkorn University, 2014, s. 452–459. ISBN 978-616-551-887-1.
- SALKIE, Raphael. *Text and discourse analysis*. Routledge, 1995. ISBN-10 0415092787. ISBN-13 978-0415092784.
- SANDERS, Ted J. M.; NOORDMAN, Leo G. M. *The role of coherence markers and their linguistic markers in text processing*. *Discourse Processes*, 2000, s. 37–60.
- SANDERS, Ted J. M.; SPOOREN, Wilbert P. M.; NOORDMAN, Leo G. M. Towards a taxonomy of discourse relations. *Discourse Processes* 15, 1992, s. 1–35.
- SGALL, Petr. *Generativní popis jazyka a česká deklinace*. 1. vyd. Praha: Academia, 1967.
- SGALL, Petr. Generativní systémy v lingvistice. *Slovo a slovesnost* 25 (4), 1964, s. 274–282.
- SGALL, Petr. K programu lingvistiky textu. *Slovo a slovesnost* 34 (1), 1973, str. 39–43.
- SGALL, Petr; HAJIČOVÁ, Eva; BURÁŇOVÁ, Eva. *Aktuální členění věty v češtině*. 1. vyd. Praha: Academia, 1980.
- SGALL, Petr; HAJIČOVÁ, Eva; PANEVOVÁ, Jarmila. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspect*. Dordrecht: Reidel Publishing Company, Praha: Academia, 1986.
- SHLOUSH, Shelley. A unified account of Hebrew bekicur in short: relevance theory and discourse structure considerations. In JUCKER, A.; ZIV, Y. (eds.). *Discourse Markers: Descriptions and Theory*. Amsterdam, Philadelphia: John Benjamins Publishing Company, 1998, s. 61–82.
- SCHIFFRIN, Deborah. *Approaches to Discourse*. Malden (MA): Blackwell, 1994.
- SCHIFFRIN, Deborah. *Discourse Markers*. Cambridge: Cambridge University Press, 1987.
- SCHMIDT, Siegfried J. *Texttheorie*. München: Fink, 1973.
- SCHOURUP, Lawrence Clifford. *Common discourse particle in English conversation*. New York: Garland, 1985.
- SKALIČKA, Vladimír. Promluva jako lingvistický pojem. *Slovo a slovesnost* 3 (3), 1937, s. 163–166.
- SKALIČKA, Vladimír. Syntax promluvy (enunciace). *Slovo a slovesnost* 21 (4), 1960, s. 241–249.
- Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia, 2001. ISBN 80-200-0493-9.
- Slovník spisovného jazyka českého*. 2., nezměněné vyd. Praha: Academia, 1989.
- SPERBER, Dan; WILSON, Deirdre. *Relevance*. Cambridge, Massachusetts: Harvard University Press, 1986.
- STEDE, Manfred; NEUMANN, Arne. Potsdam Commentary Corpus 2.0: Annotation for discourse research. In CALZOLARI, N. a kol. (eds.). *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavík: European Language Resources Association (ELRA), 2014, s. 925–929.
- STEMPEL, Wolf-Dieter (ed.). *Beiträge zur Textlinguistik*. München: Wilhelm Fink Verlag, 1971.

- STUBBS, Michael. *Discourse analysis*. Chicago, IL: The University of Chicago Press, 1983.
- TABOADA, Maite. Implicit and explicit coherence relations. In RENKEMA, J. (ed.). *Discourse, of course*. Amsterdam: John Benjamins, 2009, s. 125–138.
- TÁRNYIKOVÁ, Jarmila. Discourse markers in cross-cultural perspective. In *Perspectives* 8, 1998, s. 37–44.
- TÁRNYIKOVÁ, Jarmila. *From Text to Texture: An Introduction to Processing Strategies*. Olomouc: Universita Palackého v Olomouci, 2002. ISBN 80-244-0438-9.
- TÁRNYIKOVÁ, Jarmila. Sentence adverbials or discourse markers? In *Abstracts. International Conference Bridges and Interfaces: Form, Meaning and Function*, Prague, 1998.
- TONELLI, Sara a kol. Annotation of Discourse Relations for Conversational Spoken Dialogs. In CALZOLARI, N. a kol. (eds.). *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Valletta: European Language Resources Association (ELRA), 2010, s. 2084–2090.
- URGELLES-COLL, Miriam. *Continuum Studies in Theoretical Linguistics: Syntax and Semantics of Discourse Markers*. London: Continuum International Publishing, 2010.
- VAN DIJK, Teunen A. *Macrostructures. An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Erlbaum, 1980a.
- VAN DIJK, Teunen A. Pragmatic connectives. In *Journal of Pragmatics* 3. North-Holland Publishing Company, 1979, s. 447–456.
- VAN DIJK, Teunen A. *Some aspects of text grammar*. The Hague: Mouton, 1972.
- VAN DIJK, Teunen A. *Studies in the pragmatics of discourse*. The Hague/Berlin: Mouton, 1981.
- VAN DIJK, Teunen A. *Textwissenschaft*. Tuebingen: Niemeyer, 1980b.
- VONDRÁČEK, Miloslav. Citoslovce a částice – hranice slovního druhu. *Naše řeč* 81 (1), 1998, s. 29–38.
- VONDRÁČEK, Miloslav. Příslovce a částice – hranice slovního druhu. *Naše řeč* 82 (2), 1999, s. 72–79.
- WIDDOWSON, Henry Great. *Discourse analysis*. Oxford: Oxford University Press, 2007. ISBN-10 0194389219. ISBN-13 978-0194389211.
- WILSON, Deirdre; SPERBER, Dan. *Linguistic form and relevance*. *Lingua* 90.1/2, 1993, s. 1–25.
- WOLF, Florian a kol. *Discourse Graphbank*. Linguistic Data Consortium, Philadelphia, 2004. ISBN 1-58563-320-8.
- WOLF, Florian; GIBSON, Edward. Representing discourse coherence: A corpus-based study. *Computational Linguistics* 31 (2), 2005, s. 249–287.
- ZEYREK, Deniz a kol. The annotation scheme of the Turkish Discourse Bank and an evaluation of inconsistent annotations. In N. XUE, N.; POESIO, M. (eds.). *Proceedings of the Fourth Linguistic Annotation Workshop (LAW-IV)*. Uppsala: The Association for Computer Linguistics, 2010, s. 282–289.
- ZHOU, Yuping; XUE, Nianwen. PDTB-style discourse annotation of Chinese text. In HAIZHOU, L. a kol. (eds.). *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012)*. Jeju: The Association for Computer Linguistics, 2012, s. 69–77.
- ZIKÁNOVÁ, Šárka a kol. Zachycení výstavby textu v Pražském závislostním korpusu. *Slovo a slovesnost*, 2015. [v tisku]
- ZUBATÝ, Josef. Náli, nali-ť. *Listy filologické* 37, 1910, s. 217–228.
- ZUBATÝ, Josef. O jistém způsobu užívání zájmena an a on a o jeho původě. *Listy filologické* 36 (1), 1909, s. 25–34.
- ZWICKY, Arnold M. Clitics and particles. *Language* 61 (2), 1985, s. 283–305.

Zdroje dat

BEJČEK, Eduard a kol. *Prague Dependency Treebank 3.0*. Data/software, Univerzita Karlova v Praze, MFF, ÚFAL, Praha, 2013.

Český národní korpus – SYN2013PUB. Ústav Českého národního korpusu FF UK, Praha, 2013. Dostupné z WWW: <<https://ucnk.ff.cuni.cz/syn2013pub.php>>.

Korpus DIALOG 1.1. Praha: Ústav pro jazyk český, AV ČR, 2012. Dostupné z WWW: <<http://ujc.dialogy.cz>>.

Přílohy

Příloha 1: Seznamy konektorů

V následujících přehledech jsou uvedeny frekvenční a abecední seznamy primárních a sekundárních konektorů anotovaných v Pražském závislostním korpusu. V seznamech jsou zachycené jednotlivé formy konektorů tak, jak se objevily v textech PDT, tj. neuvádíme zde lemmata či základní tvary konektorů, ale všechny konkrétní varianty, které se vyskytly v reálných textech (tj. např. *na straně jedné_na straně druhé, na jedné straně_na druhé straně* apod.). Ve frekvenčním seznamu forem je pak možné vysledovat, která z variant je častější, tj. která má vyšší pravděpodobnost, že se právě v této podobě daný konektor ustálí.

Je zajímavé, že u některých konektorů jejich forma stále poměrně kolísá – především to, zda se v korpusu objevují psané dohromady nebo zvlášť. Například konektor *přesto* se objevuje jak ve variantě *přesto*, tak *přes to*, obě formy jsou přitom užity ve stejné funkci konektoru strukturujícího text (např. *Už jsem obědval, přesto / přes to mám ještě hlad.*). Zatímco *Pravidla českého pravopisu* (1993) u některých diskurzivních konektorů připouští pouze jednu z forem (*přesto*), u jiných akceptují obě varianty (*zaprvé* i *za prvé*). V příložených seznamech konektorů ponecháváme vždy všechny varianty, které se objevily v textech PDT bez ohledu na to, zda jsou kodifikované nebo nekodifikované. Proto se v seznamu primárních konektorů objevují vedle sebe formy *nedosti na tom* i *ne dosti na tom*; *proto, že* i *pro to, že*; *zato* i *za to*; *za druhé* i *zadruhé*; *přesto* i *přes to* (vždy však mají funkci konektoru strukturujícího text, užití typu *Vzala jsem si letní šaty a přes to lehký svetrík* v seznamech zachycena nejsou). Je zajímavé, že v rámci jednoho diskurzivního vztahu se v PDT v jednom souvětí vedle sebe vyskytl i korelativní pár s odlišným zápisem: *zaprvé_a za druhé*. Je tedy zřejmé, že u těchto konektorů dochází k poslední fázi gramatikalizace, nicméně způsob jejich zápisu (v reálných textech) ještě kolísá.

Konektory jsou v seznamech rozděleny na primární a sekundární, zároveň jsou zde (už bez explicitního rozlišení) zahrnuty i konektory složené⁶¹ a modifikované. Konektory, které jsou složené ze dvou či více primárních konektorů (srov. konektory složené ze dvou konektivních částí: *a i, ale i, ale však, : proto*⁶²..., konektory složené ze tří složek: *a i proto, a poté i, a tedy i, i tak ale, i tak však, – dokonce i...*), jsou řazeny mezi primární konektory. Konektory složené z primárních a sekundárních konektorů (např. *a pokračuje, a kromě toho,*

⁶¹ V seznamu jsou zahrnuty i korelativní páry, jsou zapisovány s podtržítkem (*bud'_nebo*).

⁶² Vzhledem k tomu, že mezi primární konektory řadíme i dvojtečku a pomlčku, považujeme za složený konektor např. i výrazy *: proto, : totiž, – proto* apod. Mezi primární konektory je řazena zároveň větná negace (v seznamech označovaná jako #neg) – srov. příklady typu *Nebudu se jenom dívat na televizi, ale půjdu i na chvíli ven.*

však na druhé straně, a v tomto případě...) jsou v seznamech řazeny mezi konektory sekundární.

V seznamech jsou zahrnuty také modifikované primární (např. *právě proto, obdobně pak, zvláště když*) a sekundární (např. *z jednoduchého důvodu; hlavní příčinou je to, že; hlavním důvodem byla skutečnost, že*) konektory, i když jejich univerzálnost pro různé kontexty není tak vysoká jako u konektorů nemodifikovaných. Ze seznamů zároveň vyplývá, že modifikované primární konektory (na rozdíl od modifikovaných sekundárních konektorů) se v reálných textech objevují jen velice zřídka, což podporuje naši myšlenku, že primární a sekundární konektory se ve způsobu a míře modifikace vzájemně liší.

Jak již bylo řečeno, hranice mezi primárními a sekundárními konektory není ostrá a v některých případech je obtížné určit, do které kategorie zkoumaný výraz patří. Jinými slovy, mezi oběma kategoriemi existuje jakési přechodné pásmo obsahující výrazy, u kterých je zařazení mezi primární či sekundární konektory předmětem diskuze. Při určování primárních a sekundárních konektorů jsme uplatňovali kritéria popsaná v této práci, tj. především míru gramatikalizace daného konektoru (odrážející např. to, zda se konektor v textu píše dohromady nebo zvlášť, skládá-li se z více komponentů), jeho větněčlenskou platnost a slovnědruhovou příslušnost.

Obtížné je např. zařazení konektorů, které se skládají z primární předložky a zájmena *ten* (příp. *což*) – tj. výrazy *k tomu, na to, s čímž, za což* aj. analogické k jednoslovným konektorům typu *proto, přitom, zato, předtím* aj. (u kterých navíc v některých případech v reálných textech PDT do jisté míry kolísá způsob jejich zápisu – srov. *při tom, za to, před tím*). V těchto případech jsme jako rozlišovací kritérium vzali skutečnost, zda se příslušné konektory píší (ve většině případů a podle úzu) dohromady (a pak jsme je označili jako primární konektory) nebo zvlášť (ty byly označeny jako konektory sekundární, protože proces gramatikalizace u nich zcela zřetelně nebyl zatím dokončen).

Další obtížně zařaditelnou skupinu tvoří příslovce typu *vzápětí, nakonec, současně, následně, posléze* apod. (tj. zejména příslovečná určení času, pokud tuto funkci ve větě ještě zastávají). V těchto případech je opět obtížné určit, zda je u daných výrazů proces gramatikalizace v takové fázi, že je možné označit je jako primární konektory. Při rozřazování do skupin jsme tedy zohledňovali opět to, že se jedná o slovní spřežky, tj. dané výrazy byly gramatikalizovány do té míry, že se píší dohromady, příp. ustrnuly (v rámci konektivní funkce) v jedné své formě – srov. *později, nejdříve*. Z tohoto důvodu řadíme příslovce typu *vzápětí, nakonec, následně, posléze, později* mezi konektory primární, analogicky ke konektorům

potom, pak. Daná příslovce jsou nicméně na přechodu mezi konektory primárními a sekundárními, protože ve větě mají např. stále větněčlenskou platnost, i když oslabenou.

Problematické je dále zařazení příslovcí typu *prostě, stejně, obdobně*. Dané výrazy stojí opět mezi konektory primárními a sekundárními – jsou na přechodu k částicím, nicméně jejich lexikální význam (většinou modální) u nich není podle našeho názoru oslaben do takové míry, abychom je mohli považovat za gramatikalizované primární konektory. Z tohoto důvodu řadíme podobné výrazy mezi konektory sekundární. Zajímavé je, že kromě funkce konektorů (srov. příklad *Nepůjdu na oběd, stejně nemám hlad.* – v podobných výskytech považujeme tato příslovce za sekundární konektory) mohou mít dané výrazy také funkci modifikátorů (srov. *stejně tak, obdobně pak* – tj. podobné příklady považujeme za modifikované primární konektory).

Ze seznamů konektorů je tedy zřejmé, že mezi primárními a sekundárními konektory existuje přechodné pásmo výrazů, které mají zároveň některé rysy primárních a některé rysy sekundárních konektorů a z tohoto hlediska je jejich zařazení do jedné či druhé kategorie předmětem diskuze.

Seznamy zahrnují celkem přibližně 400 různých forem sekundárních konektorů a 570 forem primárních konektorů. Zajímavé je, že se v reálných textech ve velké míře opakují pouze některé z nich – většina forem se v PDT objevuje naopak velice zřídka, často v jednom⁶³ či ve dvou výskytech (přibližně 310 různých forem sekundárních konektorů (tj. téměř 78 % ze 400 forem sekundárních konektorů) a 350 primárních konektorů (tj. přibližně 61 % z 570 forem primárních konektorů) se v textech PDT vyskytuje pouze v jednom, příp. ve dvou výskytech).

Zajímavé je, že naopak výskyty prvních pěti nejfrekventovanějších primárních konektorů (*a, však, ale, když, protože*) tvoří 48 % všech výskytů primárních konektorů v PDT, výskyty prvních pěti forem nejfrekventovanějších sekundárních konektorů (*dodal; podobně; v případě, že; vzhledem k tomu, že; dodává*) tvoří 26 % všech výskytů sekundárních konektorů v PDT.

Prvních deset primárních konektorů (*a, však, ale, když, protože, totiž, pokud, :, proto, tedy*) se v PDT vyskytuje celkem v 11 597 případech, tj. v 57 % z celkového počtu výskytů všech primárních konektorů (tj. z 20 255). Prvních deset forem sekundárních konektorů (*dodal; podobně; v případě, že; vzhledem k tomu, že; dodává; kromě toho; naproti tomu; to*

⁶³ Primární i sekundární konektory s nízkým počtem výskytů jsou v seznamu označeny a v rámci daného výskytu jsou seřazeny podle abecedy.

znamená, že; v této souvislosti; případně) se v PDT objevuje celkem ve 402 výskytech, tj. ve 35 % z celkového počtu výskytů všech sekundárních konektorů (tj. z 1 161 výskytů).

Ze seznamů primárních a sekundárních konektorů tedy můžeme vyvodit (zjednodušeně řečeno), že se v reálných textech objevují k vyjádření diskurzních vztahů téměř stále stejné (většinou primární) konektory. Resp. ačkoli je škála konektorů v češtině (primárních i sekundárních) značně bohatá, s vysokou frekvencí se v reálných psaných textech objevuje pouze několik z nich. Můžeme si také všimnout, že prvních pět nejfrekventovanějších primárních konektorů (*a, však, ale, když, protože*) zároveň významově pokryje všechny čtyři základní diskurzní vztahy anotované v PDT – a to vztahy navazovací (*a*), kontrastivní (*však, ale*), časové (*když*) a kauzální (*protože*).

Příloha 1.1: Abecední seznam forem primárních konektorů z Pražského závislostního korpusu

–	#neg_a spíše jenom	a nakonec
– a	#neg_ale	a naopak
– a nakonec	#neg_ale daleko spíše	a naopak vzápětí
– a přitom	#neg_ale i	a například
– a to	#neg_ale jen	a navíc
– a zároveň	#neg_ale pouze	a ne
– aby	#neg_ale především	a nebo
– což	#neg_ale spíš	a nikoli
– dokonce	#neg_ale spíše	a nikoliv
– dokonce i	#neg_ale také	a pak
– leda	#neg_ale zato	a pokud
– např.	#neg_ba naopak	a posléze
– například	#neg_i	a posléze i
– pokud	#neg_jako	a potažmo
– proto	#neg_jen	a poté
– přesto	#neg_jen i	a poté i
– přitom	#neg_jenom	a potom
– tedy	#neg_mnohem spíše	a později
– totiž	#neg_naopak	a proto
– tzn.	#neg_nýbrž	a přece
– vždyť ⁶⁴	#neg_pouze	a především
#neg ⁶⁴	#neg_přitom	a předtím
#neg a	#neg_spíš	a přesto
#neg ani_jako spíš ⁶⁵	#neg_spíše	a při tom
#neg jen_a	#neg_však	a přímo
#neg jen_ale	:	a přitom
#neg jen_ale i	: aby	a rovněž
#neg jen_ale především	: např.	a současně
#neg jen_ale také	: například	a stejně
#neg jen_ale též	: ovšem	a stejně tak
#neg jen_nýbrž	: proto	a tak
#neg jen_především	: přece	a také
#neg jenom_ale i	: tedy	a také proto
#neg nic_jen	: totiž	a taky
#neg pouze_ale	a	a tedy
#neg pouze_ale i	a ani	a tedy i
#neg pouze_i	a bezprostředně poté	a teprve potom
#neg pouze_nýbrž	a dále	a též
#neg sice_však	a dokonce	a tím
#neg_a naopak	a dokonce i	a tím i
#neg_a pouze	a hlavně	a to
#neg_a spíš	a hned pak	a to i
#neg_a spíše	a i	a to jen
	a i proto	a to ještě
	a ještě	a to proto
	a konečně	a tudíž
	a nadále	a vzápětí

⁶⁴ #neg je značka pro větnou negaci.

⁶⁵ Korelativní páry jsou zapisovány s podtržítkem (*bud'_nebo*).

a zároveň	avšak pouze	i tak ale
a zase	až	i tak však
a zejména	až když	i tehdy, kdy
a zrovna tak	až na	i však
a_b	až nakonec	jak
aby	až poté	jak_tak
aby nakonec	až poté, kdy	jakkoli
aby posléze	až tehdy, až	jakkoliv tak
aby tak	ba	jakmile
ač	ba co víc	jako
ačkoli	ba i	jako vždy, když
ačkoliv	bud'_anebo	jednak_a
ale	bud'_či	jednak_a dále
ale ani	bud'_nebo	jednak_a hlavně
ale dokonce	bud'to	jednak_a jednak
ale i	bud'to_nebo	jednak_jednak
ale i tak	byť	jednak_především
ale jen	což	však
ale nakonec	což ale	jednak_také
ale naopak	což ovšem	jelikož
ale například i	což však	jen
ale ne	čemuž	jen když
ale přece	či	jen tehdy_-li
ale přece jen	či spíše	jen však
ale především	čili	jenom
ale přesto	čímž	jenomže
ale přitom	dál	jenže
ale spíše	dále	jenže i potom
ale stejně	dále pak	jenže také
ale také	dále_a nakonec	jestli
ale třeba i	dokonce	jestli pak
ale však	dokonce ani	jestli_tak
ale zároveň	dokonce aniž	jestliže
ale zato	dokonce i	jestliže potom
alespoň	dokonce i když	jestliže_pak
aneb	dokud	jestliže_tak
anebo	dokud_tak	ještě
ani	dříve, než	ještě předtím
ani ne_ale spíše	hned vzápětí	ještě spíše
ani ne_jako spíše	i	ježto
ani_ani	i ale	jinak
aniž	i kdyby	jinak však
aniž_tak	i kdyby_tak	již dříve
aspoň	i když	kdežto
ať	i když pouze	kdežto naopak
ať již	i když přesto	kdyby
ať už	i nebo spíše právě	kdyby_pak
ať už_-li	i přesto	kdyby_potom
avšak	i přesto, že	kdyby_tak
avšak jen	i tak	kdykoli

kdykoliv	nebo spíš	nikoli
když	nebo spíše	nikoli pouze
když předtím	neboli	nikoli tedy
když už	neboť	nikoli však
když_když	nejdřív_pak	nikoli_ale
když_tak	nejdřív_potom	nikoliv
koneckonců	nejdříve_pak	nikoliv jenom
leč	nejdříve_potom	nikoliv ovšem
leč přece jen	nejen	nikoliv však
ledaže	nejen navíc i	nikoliv_ale
-li	nejen, že_ale zejména	nikoliv_nýbrž
-li_jestliže	nejen:	nýbrž
-li_pak	nejen_ale	obdobně pak
-li_potom	nejen_ale ani	ostatně i
-li_tak	nejen_ale dokonce	ovšem
mezi tím	nejen_ale i	ovšem pouze
mezitím	nejen_ale naopak	ovšem přeci jen
mimo jiné	nejen_ale především	ovšem také
mimo to	nejen_ale přímo	ovšemže
mimoto	nejen_ale rovněž	pak
mj.	nejen_ale rovnou	pak ovšem
nadále	nejen_ale také	pak totiž
nadto	nejen_ale též	pakliže
nakonec	nejen_ale zároveň	po tom, co
nakonec totiž	nejen_nýbrž i	po tom, kdy
naopak	nejenom i	pokud
naopak však	nejenom_ale i	pokud ovšem
naprosto ne	nejenom_ale také	pokud potom
např.	nejenom_ale zejména	pokud_-li
napřed_a teprve potom	nejenomže_ale	pokud_pak
napřed_pak	nejenže	pokud_tak
například	nejenže navíc	poněvadž
například, když	nejenže_ale	posléze
následně	nejenže_ale ani	poté
natož	nejenže_ale i	poté, co
natož pak	nejenže_ale ještě	poté, kdy
navíc	nejenže_ale také	potom
ne	nejenže_ale rovněž	potom, co
ne však	nejprve	potom, kdy
ne_ale	nejprve_a pak	pouze
ne_nýbrž	nejprve_a poté	pouze tehdy, pokud
ne_pouze	nejprve_dále	pouze tehdy_-li
neb	nejprve_pak	pouze_nikoli
nebo	nejprve_poté	právě proto
nebo alespoň	nejprve_potom	pro to, aby
nebo dokonce	nejprve_později	proto
nebo i	nejprve_vzápětí	proto také
nebo ještě	nejprve_záhy	proto, aby
nebo naopak	než	proto, že
nebo pouze	nicméně	proto_-li

protože
protože pokud
protože_tak
přece
přece jen
přece jenom
před tím
před tím, než
především pak
především pokud
především však
předtím
předtím, než
přesto
přesto ale
přesto však
přesto, že
přestože
při tom
příčemž
příčemž i
příčemž ještě
příčemž navíc
příčemž zároveň
přítom
přítom jen
přítom navíc
přítom také
přítom však
přítom zase
resp.
respektive
rovněž
sice
sice_a
sice_ale
sice_ale aspoň
sice_ale i tak
sice_ale pouze
sice_ale přesto
sice_ale taky
sice_ale zároveň
sice_ale zato
sice_avšak
sice_jenže
sice_jenže také
sice_nicméně
sice_nicméně však

sice_ovšem
sice_přesto
sice_přesto však
sice_však
sice_zato
snad jen
sotva
souběžně
současně ale
současně však
spíš
spíše
spíše však
stejně jako
stejně tak
stejně tak ale
stejně tak i
tak
tak ale
tak např.
tak například
tak tedy
tak, aby
také
také proto
taky
takže
takže nakonec
takže například
takže vlastně zase
teda
tedy
tehdy
tehdy, až
tehdy, jestliže
tehdy, když
tehdy, pokud
tehdy_-li
tehdy_pokud
teprve
teprve poté
teprve potom
též
tím
tím ovšem
tím, že
tj.
totiž

totiž navíc
třeba
třebaže
tudíž
už
už jen
už ovšem
už předtím
vlastně totiž
však
však i
však mezitím
však naopak
však navíc
však nejen
však pouze
však přece i
však přeci jen
však přesto
však současně
však také
však zároveň
vzápětí
vždyť
vždyť také
za to
záhy poté
zároveň
zároveň ale
zároveň naopak
zároveň však
zas
zase
zase naopak
zatím
zatím ale
zatímco
zatímco naopak
zato
zato ale
zato však
zčásti_zčásti
zda
zdali
zejména
zejména pokud
zprvu_později
zvláště když

Příloha 1.2: Frekvenční seznam forem primárních konektorů z Pražského závislostního korpusu

a	jenže	a navíc
však	avšak	a potom
ale	a pak	předtím
když	aniž	nejen_ale
protože	potom	pro to, aby
totiž	a také	#neg_pouze
pokud	ačkoli	třeba
:	vždyť	tudíž
proto	zase	zároveň však
tedy	sice_však	třebaže
aby	až	vzápětí
pak	ačkoliv	přesto však
ovšem	než	a poté
-li	a i	stejně tak
také	ani	ne_ale
neboť	dokud	nýbrž
–	jakmile	sice_ovšem
zatímco	pokud_pak	mezitím
nebo	poté	ať už
navíc	#neg_nýbrž	: totiž
přitom	zato	jestli
i když	ač	pouze
což	#neg_jen	a zároveň
sice_ale	byť	ne
naopak	ať	posléze
takže	anebo	čili
a tak	přesto, že	ale zároveň
a to	buď_nebo	a ještě
dále	čímž	kdežto
kdyby	jinak	tehdy, když
tak	nejen_ale i	– totiž
rovněž	a tedy	ale také
proto, že	a tudíž	těž
přesto	jelikož	tj.
#neg_ale	nakonec	a tím
přestože	pokud_tak	#neg_však
například	a ne	kdykoliv
zároveň	a přitom	sice_avšak
příčemž	a dokonce	jestliže_pak
a proto	a ani	jak
či	a naopak	předtím, než
poté, co	jen	kdyby_tak
i	i když	nikoli
jestliže	#neg_spíše	jenomže
dokonce	např.	jakkoli
#neg	-li_pak	při tom
nicméně	tím, že	ale přesto

a přesto
sotva
poněvadž
-li_tak
však také
mimo jiné
nejen_ale také
a rovněž
poté, kdy
taky
ale i
přece
ještě
buď_anebo
jednak_jednak
a stejně tak
tehdy_-li
a posléze
a přece
#neg a⁶⁶
#neg jen_ale i
#neg_jen i
#neg_jenom
a hlavně
a nakonec
a vzápětí
a zase
ale ani
ani_ani
čemuž
i přesto
jak_tak
jenom
jestli_tak
následně
nebo dokonce
nejdříve_pak
nejen_ale dokonce
nejprve_pak
ovšem také
před tím, než
stejně jako
tím
však i
však přesto
však zároveň
#neg jen_ale⁶⁷

⁶⁶ Od této značky k další mají všechny formy konektorů v PDT shodně 4 výskyty a jsou řazeny abecedně.

#neg_ale jen
#neg_naopak
#neg_spíš
a nikoliv
a později
a současně
a stejně
a taky
ale i tak
ale přece
i tak
jestli pak
jestliže_tak
kdyby_pak
když_když
leč
mimoto
mj.
nadto
ne však
nebo alespoň
nejen_ale především
nejenže_ale i
nejprve_potom
potom, co
právě proto
proto také
proto, aby
přesto ale
sice
sice_ale přesto
sice_nicméně
sice_zato
souběžně
současně však
spíše
zas
- a⁶⁸
- dokonce
- tedy
#neg jen_ale především
#neg jen_ale také
#neg_a naopak

⁶⁷ Od této značky k další mají všechny formy konektorů v PDT shodně 3 výskyty a jsou řazeny abecedně.

⁶⁸ Od této značky k další mají všechny formy konektorů v PDT shodně 2 výskyty a jsou řazeny abecedně.

#neg_ale spíše
: např.
: tedy
a dokonce i
a konečně
a například
a nebo
a potažmo
a především
ale ne
ale především
ale přitom
ale zato
aneb
až když
až nakonec
ba
buďto
dále pak
dokonce i
dokud_tak
dříve, než
i přesto, že
i tehdy, kdy
jednak_a
jen když
kdyby_potom
když_tak
koneckonců
ledaže
mimo to
nadále
natož
natož pak
ne_nýbrž
nejen
nejen_ale rovněž
nejenom_ale i
nejenže
nejenže_ale
nejenže_ale také
nejprve_a pak
nejprve_a poté
nikoli však
nikoli_ale
nikoliv_nýbrž
pakliže
po tom, co
pokud potom
pouze tehdy_-li

přičemž ještě
 přitom však
 resp.
 sice_ale zato
 sice_přesto
 spíš
 tak, aby
 tehdy, pokud
 už
 už předtím
 však pouze
 zároveň ale
 zejména pokud
 zprvu_později
 – a nakonec⁶⁹
 – a přitom
 – a to
 – a zároveň
 – aby
 – což
 – dokonce i
 – leda
 – např.
 – například
 – pokud
 – proto
 – přesto
 – přitom
 – tzn.
 – vždyť
 #neg ani_jako spíš
 #neg jen_a
 #neg jen_ale též
 #neg jen_nýbrž
 #neg jen_především
 #neg jenom_ale i
 #neg nic_jen
 #neg pouze_ale
 #neg pouze_ale i
 #neg pouze_i
 #neg pouze_nýbrž
 #neg sice_však
 #neg_a pouze
 #neg_a spíš
 #neg_a spíše
 #neg_a spíše jenom
 #neg_ale daleko spíše

#neg_ale i
 #neg_ale pouze
 #neg_ale především
 #neg_ale spíš
 #neg_ale také
 #neg_ale zato
 #neg_ba naopak
 #neg_i
 #neg_jako
 #neg_mnohem spíše
 #neg_přitom
 : aby
 : například
 : ovšem
 : proto
 : přece
 a bezprostředně poté
 a dále
 a hned pak
 a i proto
 a nadále
 a naopak vzápětí
 a nikoli
 a pokud
 a posléze i
 a poté i
 a předtím
 a při tom
 a přímo
 a také proto
 a tedy i
 a teprve potom
 a též
 a tím i
 a to i
 a to jen
 a to ještě
 a to proto
 a zejména
 a zrovna tak
 a_b
 aby nakonec
 aby posléze
 aby tak
 ale dokonce
 ale jen
 ale nakonec
 ale naopak
 ale například i
 ale přece jen

ale spíše
 ale stejně
 ale třeba i
 ale však
 alespoň
 ani ne_ale spíše
 ani ne_jako spíše
 aniž_tak
 aspoň
 ať již
 ať už_-li
 avšak jen
 avšak pouze
 až na
 až poté
 až poté, kdy
 až tehdy, až
 ba co víc
 ba i
 buď_či
 buďto_nebo
 což ale
 což ovšem
 což však
 či spíše
 dál
 dále_a nakonec
 dokonce ani
 dokonce aniž
 dokonce i když
 hned vzápětí
 i ale
 i kdyby_tak
 i když pouze
 i když přesto
 i nebo spíše právě
 i tak ale
 i tak však
 i však
 jakkoliv tak
 jako
 jako vždy, když
 jednak_a dále
 jednak_a hlavně
 jednak_a jednak
 jednak_především
 však
 jednak_také
 jen tehdy_-li
 jen však

⁶⁹ Od této značky k další mají všechny formy konektorů v PDT shodně 1 výskyt a jsou řazeny abecedně.

jenže i potom
jenže také
jestliže potom
ještě předtím
ještě spíše
ježto
jinak však
již dříve
kdežto naopak
kdykoli
když předtím
když už
leč přece jen
-li_jestliže
-li_potom
mezi tím
nakonec totiž
naopak však
naprosto ne
napřed_a teprve potom
napřed_pak
například, když
ne_pouze
neb
nebo i
nebo ještě
nebo naopak
nebo pouze
nebo spíš
nebo spíše
neboli
nejdřív_pak
nejdřív_potom
nejdříve_potom
nejen navíc i
nejen, že_ale zejména
nejen:
nejen_ale ani
nejen_ale naopak
nejen_ale přímo
nejen_ale rovnou
nejen_ale též
nejen_ale zároveň
nejen_nýbrž i
nejenom i
nejenom_ale také
nejenom_ale zejména
nejenomže_ale
nejenže navíc
nejenže_ale ani

nejenže_ale ještě
nejenže_rovněž
nejprve
nejprve_dále
nejprve_poté
nejprve_později
nejprve_vzápětí
nejprve_záhy
nikoli pouze
nikoli tedy
nikoliv
nikoliv jenom
nikoliv ovšem
nikoliv však
nikoliv_ale
obdobně pak
ostatně i
ovšem pouze
ovšem přeci jen
ovšemže
pak ovšem
pak totiž
po tom, kdy
pokud ovšem
pokud_-li
potom, kdy
pouze tehdy, pokud
pouze_nikoli
proto_-li
protože pokud
protože_tak
přece jen
přece jenom
před tím
především pak
především pokud
především však
přičemž i
přičemž navíc
přičemž zároveň
přítom jen
přítom navíc
přítom také
přítom zase
respektive
sice_a
sice_ale aspoň
sice_ale i tak
sice_ale pouze
sice_ale taky

sice_ale zároveň
sice_jenže
sice_jenže také
sice_nicméně však
sice_přesto však
snad jen
současně ale
spíše však
stejně tak ale
stejně tak i
tak ale
tak např.
tak například
tak tedy
také proto
takže nakonec
takže například
takže vlastně zase
teda
tehdy
tehdy, až
tehdy, jestliže
tehdy_pokud
teprve
teprve poté
teprve potom
tím ovšem
totiž navíc
už jen
už ovšem
vlastně totiž
však mezitím
však naopak
však navíc
však nejen
však přece i
však přeci jen
však současně
vždyť také
za to
záhy poté
zároveň naopak
zase naopak
zatím
zatím ale
zatímco naopak
zato ale
zato však
zčásti_zčásti
zda

zdali

zejména

zvláště když

Příloha 1.3: Abecední seznam forem sekundárních konektorů z Pražského závislostního korpusu

– ostatně	ani v tomto případě
– prostě	argumentoval
– vlastně	argumentoval tím, že
#neg_místo toho	argumentovali tím, že
#neg_raději	argumentovalo tím, že
#neg_rozhodně	argumentuje, že
#neg_to jenom	argumentují tím, že
a bude pokračovat	argumentují, že
a do třetice	avšak na druhé straně naopak
a dodal jsem	bude pokračovat
a dodejme	byla podmínkou v opačném případě
a doplňuje	což neznámá, že
a chtělo by se dodat	což znamená, že
a jako důvod uvedla, že	což znamenalo, že
a ještě k tomu	dále dodal
a k tomu	dám příklad
a k tomu ještě	díky tomu
a kromě toho	díky tomu, že
a kvůli tomu	dlužno dodat, že
a měly k tomu pochopitelné důvody	do třetice
a na závěr	dodal
a naproti tomu	dodala
a pokračoval	dodává
a pokračovali	dodávám
a pokračovat bude	dodejme
a pokračuje	dodejme rovněž
a stejně	doplnil
a tím pádem	doplňuje
a to je důvod, proč	dospěl k závěru, že
a to znamená, že	došli zkrátka k závěru, že
a v důsledku toho	důsledkem je
a v tomto případě	důvod je jednoduchý
a ve skutečnosti	důvod je složitější
a vlastně	důvod je zřejmý
a z toho důvodu	důvod spočívá v tom, že
a z tohoto důvodu	důvodem bylo to, že
a za druhé	důvodem bylo, že
aby pokračovala	důvodem je
ač ve skutečnosti	důvodem je skutečnost, že
ale místo toho	důvodem je to, že
ale na druhé straně	důvodem je, že
ale o to víc	důvodů je mnoho
ale stejně	důvodů proč je více
ale to neznámá, že	důvody jsou dva
ale ve skutečnosti	důvody jsou jiné
anebo ještě přesněji	existuje řada důvodů, aby
anebo přesněji	existuje více důvodů, proč

existují pro to tři důvody
fakt neznamená, že
hlavní důvod, proč
hlavní příčinou je to, že
hlavním důvodem byla skutečnost, že
hned z několika důvodů
chtěli bychom dodat
i přes tato fakta
jak dodal
jak dodali
jak dodává
jak je vidět
jak upřesňuje
jak_tak případně
jako poslední
jako příklad uvádí to, že
je to důsledek toho, že
je to důvod, proč
je to z toho důvodu, že
je to způsobeno tím, že
je třeba dodat
jednak a kromě toho
jednak na straně druhé zase
jednak_a za druhé
jedním z těchto důvodů je skutečnost, že
jednoduše řečeno
jen v případě, že
ještě k tomu
jinak řečeno
jinými podmínkami bylo, že
jinými slovy
jmenovitě
jmenujme dva příklady
k čemuž
k tomu
k tomu přistupuje skutečnost, že
k tomu třeba dodat
k tomu, aby
k tomuto účelu
když ve skutečnosti
konkrétně
krátce nato
krátce řečeno
kromě toho
kromě toho ovšem i
kromě toho také
kromě toho však
kromě toho, že
kvůli tomu
kvůli tomu, že

lépe řečeno
má pokračovat
má to své důvody
mají k tomu dobrý důvod
mají pokračovat
mezi tím
místo toho
místo toho však
místo toho, aby
místo, aby
mohl bych pokračovat
mohli bychom pokračovat
na druhé straně
na druhé straně ovšem
na druhé straně však
na druhé straně zase
na druhou stranu
na jedné straně_a na druhé straně
na jedné straně_ale na druhé straně
na jedné straně_jenže na druhé straně
na jedné straně_na druhé straně
na jedné straně_na druhé straně však
na jednu stranu
na první pohled však
na prvním místě
na rozdíl od toho
na rozdíl od toho, co
na straně druhé
na straně jedné_a na straně druhé
na to, aby
na základě čehož
na základě toho, že
naproti tomu
následoval poté, co
následovat bude
navzdory tomu
navzdory tomu však
navzdory tomu, že
ne dosti na tom
nedosti na tom
nedosti na tom:
nehledě na to, že
nejen kromě toho
nemluvě o tom, kdy
nemluvě o tom, že
není to z důvodu, že
neznamená to, že
nutno ovšem dodat
o což
o čemž

o důvod víc
o důvodech není třeba příliš hovořit
o to víc
o to víc, že
o to více
o to více, že
od čehož
od okamžiku, kdy
od toho, aby
oproti tomu
ostatně
ovšem jen v případě, že
ovšem s tím, že
ovšem za podmínky, že
pak dodal
podmínka je to nevyhnutelná
podmínkou je
podmínkou je, aby
podmínkou je, že
podobně
podobně i
pokračoval
pokračovala
pokračovat by mělo
pokračuje
pokračují
pokud budou pokračovat
popř.
popřípadě
poslužme příkladem
pouze v případě, že
pouze v případě, že nebo když
pouze v těch případech, kdy
pravděpodobnější verzí je, že
právě tak
právě tak jako
pro případ, že
pro takový případ
pro tento účel
prostě
především vinou toho
přeloženo
přes tyto skutečnosti
přesněji
přesněji řečeno
příčin proč je několik
příčinu vidí v tom, že
příkladem je
příklady můžeme nalézt
případně

připustil v této souvislosti, že
přispívá k tomu fakt, že
rozdíl je v tom, že
rozdílem je skutečnost, že
s jednou podmínkou
s odůvodněním, že
s podmínkou, že
s rozdílem, že
s tím
s tím kontrastuje skutečnost, že
s tím, aby
s tím, že
s tím, že pak
s tím, že zároveň
s výjimkou případů, kdy
sice_ale na druhé straně
sice_ale ve skutečnosti
spolu s tím rovněž
srozumitelně řečeno
srozumitelněji řečeno
stejně
stejným dechem
stručně řečeno
tato skutečnost se stala jedním z
důvodů, kvůli kterým
tím argumentoval
tím pádem
tím spíš, že
tím spíše
tím spíše, že
to ale neznamená, že
to ale znamenalo, že
to byl důvod, proč
to byl hlavní důvod, proč
to byl jeden z hlavních důvodů, proč
to je důvod, proč
to je důvodem k tomu, aby
to je výklad možných příčin toho, že
to lze přičíst skutečnosti, že
to má svůj důvod
to má závažné důsledky
to může znamenat pouze jediné
to nebyl důvod, proč
to neznamená, že
to odůvodnil tím, že
to ovšem neznamená, že
to se dá přičíst tomu, že
to se týká
to souvisí s
to vede k

to však neznamená, že
to znamená
to znamená jediné
to znamená, že
to znamenalo, že
upřesnil
upřesňuje
uvedeme příklad
uvedl příklad, kdy
uvedu příklad
v čemž
v každém případě
v poslední řadě
v opačném případě
v podstatě
v podstatě to znamená, že
v první řadě
v prvním případě_ve druhém případě
v případě
v případě toho, že
v případě, kdy
v případě, kdyby
v případě, že
v případě, že a pokud
v případech, kdy
v souvislosti s tím, že
v takovém případě
v této souvislosti
v této souvislosti nezbyvá dodat, že
v tom pokračují
v tom případě
v tom případě, že
v tom, že zatímco
v tomto případě
v závěru
ve skutečnosti
ve spojení s tím, že
ve stejném okamžiku
vede k tomu skutečnost, že
vedle toho
vedle toho, že i
vlastně
však dodat stejně tak
však na druhé straně
však navzdory těmto faktům
však nutno dodat
však ve skutečnosti
výjimkou je
vyjma případu, že

výsledkem je
výsledkem je, že
vysvětlují si to dvěma důvody
vzhledem k těmto skutečnostem
vzhledem k tomu
vzhledem k tomu, co
vzhledem k tomu, že
z čehož
z čehož lze odhadnout, že
z důvodu, že
z faktu nevyplývá, že
z jedné strany_ze strany druhé
z jednoduchého důvodu
z jednoduchých důvodů
z mnoha důvodů
z přiznání plyne, že
z různých důvodů, mezi nimiž
z těchto důvodů
z toho důvodu
z toho důvodu, aby
z toho důvodu, že
z toho plyne, že
z toho vyplývá, že
z toho, když
z tohoto důvodu
z tohoto typu plyne
z uvedených důvodů
za což
za druhé
za podmínky, kdyby
za podmínky, že
za tím účelem
za tímto účelem
za to, že
zprvce_a za druhé
zásluhou toho, že
závěr je
závěr je tedy následující
zdůvodnil to tím, že
zdůvodňoval by to tím, že
ze dvou důvodů
ze stejného důvodu
ze zcela jiných příčin totiž
zjednodušeně řečeno
zkrátka
znamená to toto
znamená to, že
znamenal to, že

Příloha 1.4: Frekvenční seznam forem sekundárních konektorů z Pražského závislostního korpusu

dodal
podobně
v případě, že
vzhledem k tomu, že
dodává
kromě toho
naproti tomu
to znamená, že
v této souvislosti
případně
příkladem je
upřesnil
znamená to, že
z tohoto důvodu
důvodem je
díky tomu, že
výjimkou je
výsledkem je
za podmínky, že
díky tomu
dodala
to znamená
pokračuje
tím pádem
to neznamená, že
to se týká
v takovém případě
ve skutečnosti
z toho důvodu, že
jinak řečeno
právě tak
to souvisí s
to vede k
v tomto případě
s tím, že
k tomu, aby
na druhé straně
prostě
za to, že
místo, aby
na druhé straně však
na to, aby
na druhou stranu
místo toho
stejně
popřípadě

a pokračuje⁷⁰
což znamená, že
do třetice
jak dodal
k tomu
kvůli tomu
od toho, aby
oproti tomu
pokračoval
tím spíš, že
to však neznamená, že
v tom případě
a kromě toho⁷¹
ale ve skutečnosti
argumentují tím, že
což neznamená, že
doplňuje
důsledkem je
jak je vidět
jinými slovy
k čemuž
kromě toho, že
navzdory tomu, že
podmínkou je, aby
přesněji řečeno
s odůvodněním, že
tím spíše
to ovšem neznamená, že
v neposlední řadě
v případě, kdyby
vedle toho
z toho důvodu
za což
za tímto účelem
ze dvou důvodů
zkrátka
#neg_raději⁷²
a pokračoval

⁷⁰ Od této značky k další mají všechny formy konektorů v PDT shodně 4 výskyty a jsou řazeny abecedně.

⁷¹ Od této značky k další mají všechny formy konektorů v PDT shodně 3 výskyty a jsou řazeny abecedně.

⁷² Od této značky k další mají všechny formy konektorů v PDT shodně 2 výskyty a jsou řazeny abecedně.

a z tohoto důvodu
ale na druhé straně
ale to neznámá, že
argumentovali tím, že
dodejme
doplnil
důvod spočívá v tom, že
důvodem je skutečnost, že
je to z toho důvodu, že
jen v případě, že
k tomuto účelu
krátce nato
krátce řečeno
kromě toho však
místo toho, aby
na druhé straně zase
na jedné straně_ na druhé straně však
nedosti na tom:
nehledě na to, že
nemluvě o tom, že
o to víc
o to více
pokračovala
pouze v případě, že
přesněji
rozdíl je v tom, že
s podmínkou, že
s rozdílem, že
s tím, aby
to ale neznámá, že
upřesňuje
v čemž
v opačném případě
v případech, kdy
vlastně
výsledkem je, že
z čehož
z jednoduchého důvodu
z těchto důvodů
z toho plyne, že
z toho vyplývá, že
zásluhou toho, že
ze stejného důvodu
zjednodušeně řečeno
– ostatně⁷³
– prostě
– vlastně
#neg_místo toho

#neg_rozhodně
#neg_to jenom
a bude pokračovat
a do třetice
a dodal jsem
a dodejme
a doplňuje
a chtělo by se dodat
a jako důvod uvedla, že
a ještě k tomu
a k tomu
a k tomu ještě
a kvůli tomu
a měly k tomu pochopitelné důvody
a na závěr
a naproti tomu
a pokračovali
a pokračovat bude
a stejně
a tím pádem
a to je důvod, proč
a to znamená, že
a v důsledku toho
a v tomto případě
a ve skutečnosti
a vlastně
a z toho důvodu
a za druhé
aby pokračovala
ač ve skutečnosti
ale místo toho
ale o to víc
ale stejně
anebo ještě přesněji
anebo přesněji
ani v tomto případě
argumentoval
argumentoval tím, že
argumentovalo tím, že
argumentuje, že
argumentují, že
avšak na druhé straně naopak
bude pokračovat
byla podmínkou v opačném případě
což znamenalo, že
dále dodal
dám příklad
dlužno dodat, že
dodávám
dodejme rovněž

⁷³ Od této značky k další mají všechny formy konektorů v PDT shodně 1 výskyt a jsou řazeny abecedně.

dospěl k závěru, že
došli zkrátka k závěru, že
důvod je jednoduchý
důvod je složitější
důvod je zřejmý
důvodem bylo to, že
důvodem bylo, že
důvodem je to, že
důvodem je, že
důvodů je mnoho
důvodů proč je více
důvody jsou dva
důvody jsou jiné
existuje řada důvodů, aby
existuje více důvodů, proč
existují pro to tři důvody
fakt neznámá, že
hlavní důvod, proč
hlavní příčinou je to, že
hlavním důvodem byla skutečnost, že
hned z několika důvodů
chtěli bychom dodat
i přes tato fakta
jak dodali
jak dodává
jak upřesňuje
jak_tak případně
jako poslední
jako příklad uvádí to, že
je to důsledek toho, že
je to důvod, proč
je to způsobeno tím, že
je třeba dodat
jednak a kromě toho
jednak na straně druhé zase
jednak_a za druhé
jedním z těchto důvodů je skutečnost, že
jednoduše řečeno
ještě k tomu
jinými podmínkami bylo, že
jmenovitě
jmenujme dva příklady
k tomu přistupuje skutečnost, že
k tomu třeba dodat
když ve skutečnosti
konkrétně
kromě toho ovšem i
kromě toho také
kvůli tomu, že
lépe řečeno

má pokračovat
má to své důvody
mají k tomu dobrý důvod
mají pokračovat
mezi tím
místo toho však
mohl bych pokračovat
mohli bychom pokračovat
na druhé straně ovšem
na jedné straně_a na druhé straně
na jedné straně_ale na druhé straně
na jedné straně_jenže na druhé straně
na jedné straně_na druhé straně
na jednu stranu
na první pohled však
na prvním místě
na rozdíl od toho
na rozdíl od toho, co
na straně druhé
na straně jedné_a na straně druhé
na základě čehož
na základě toho, že
následoval poté, co
následovat bude
navzdory tomu
navzdory tomu však
ne dosti na tom
nedostí na tom
nejen kromě toho
nemluvě o tom, kdy
není to z důvodu, že
neznámá to, že
nutno ovšem dodat
o což
o čemž
o důvod víc
o důvodech není třeba příliš hovořit
o to víc, že
o to více, že
od čehož
od okamžiku, kdy
ostatně
ovšem jen v případě, že
ovšem s tím, že
ovšem za podmínky, že
pak dodal
podmínka je to nevyhnutelná
podmínkou je
podmínkou je, že
podobně i

pokračovat by mělo
pokračují
pokud budou pokračovat
popř.
poslužme příkladem
pouze v případě, že nebo když
pouze v těch případech, kdy
pravděpodobnější verzí je, že
právě tak jako
pro případ, že
pro takový případ
pro tento účel
především vinou toho
přeloženo
přes tyto skutečnosti
příčin proč je několik
příčinu vidí v tom, že
příklady můžeme nalézt
připustil v této souvislosti, že
přispívá k tomu fakt, že
rozdílem je skutečnost, že
s jednou podmínkou
s tím
s tím kontrastuje skutečnost, že
s tím, že pak
s tím, že zároveň
s výjimkou případů, kdy
sice_ale na druhé straně
sice_ale ve skutečnosti
spolu s tím rovněž
srozumitelně řečeno
srozumitelněji řečeno
stejným dechem
stručně řečeno
tato skutečnost se stala jedním z důvodů,
kvůli kterým
tím argumentoval
tím spíše, že
to ale znamenalo, že
to byl důvod, proč
to byl hlavní důvod, proč
to byl jeden z hlavních důvodů, proč
to je důvod, proč
to je důvodem k tomu, aby
to je výklad možných příčin toho, že
to lze přičíst skutečnosti, že
to má svůj důvod
to má závažné důsledky
to může znamenat pouze jediné
to nebyl důvod, proč

to odůvodnil tím, že
to se dá přičíst tomu, že
to znamená jediné
to znamenalo, že
uvedeme příklad
uvedl příklad, kdy
uvedu příklad
v každém případě
v podstatě
v podstatě to znamená, že
v první řadě
v prvním případě_ve druhém případě
v případě
v případě toho, že
v případě, kdy
v případě, že a pokud
v souvislosti s tím, že
v této souvislosti nezbyvá dodat, že
v tom pokračují
v tom případě, že
v tom, že zatímco
v závěru
ve spojení s tím, že
vede k tomu skutečnost, že
vedle toho, že i
však dodat stejně tak
však na druhé straně
však navzdory těmto faktům
však nutno dodat
však ve skutečnosti
vyjma případu, že
vysvětlují si to dvěma důvody
vzhledem k těmto skutečnostem
vzhledem k tomu
vzhledem k tomu, co
z čehož lze odhadnout, že
z důvodu, že
z faktu nevyplývá, že
z jedné strany_ze strany druhé
z jednoduchých důvodů
z mnoha důvodů
z přiznání plyne, že
z různých důvodů, mezi nimiž
z toho důvodu, aby
z toho, když
z tohoto typu plyne
z uvedených důvodů
za druhé
za podmínky, kdyby
za tím účelem

zprvé_a za druhé
závěr je
závěr je tedy následující
zdůvodnil to tím, že
zdůvodňoval by to tím, že

ve stejném okamžiku
ze zcela jiných příčin totiž
znamená to toto
znamenalo to, že

Příloha 1.5: Frekvenční seznam forem primárních a sekundárních konektorů z Pražského závislostního korpusu s ohledem na vyjadřované významové vztahy

Disjunktivní alternativa: primární konektory

nebo
či
buď_nebo
anebo
jinak
buď_anebo
nebo alespoň
a nebo
a⁷⁴
buď_či
buďto_nebo
i
nebo ještě
nebo naopak

Disjunktivní alternativa: sekundární konektory

případně
popřípadě

⁷⁴ Od šedé značky dále mají v této příloze všechny formy konektorů v PDT shodně 1 výskyt a jsou řazeny abecedně.

Ekvivalence: primární konektory

tedy
čili
tak
tj.
a
–
takže
a tedy
tudíž
což
:
a tak
aneb
– tzn.
a tedy i
a tím
neboli
potom
proto

Ekvivalence: sekundární konektory

prostě
jinak řečeno
jinými slovy
srozumitelně řečeno
srozumitelněji řečeno
upřesnil
z toho vyplývá, že

Exemplifikace: primární konektory

například
např.
třeba
mimo jiné
mj.
a například
: např.
■
– např.
– například
:
: například
tak
tak např.
tak například
takže například
tedy

Exemplifikace: sekundární konektory

příkladem je
dám příklad
jako příklad uvádí to, že
jmenujme dva příklady
poslužme příkladem
uvedeme příklad
uvedl příklad, kdy
uvedu příklad

Explikace: primární konektory

totiž
:
–
vždyť
protože
když
neboť
: totiž
poté, co
přece
proto
proto, že
– totiž
přece
aby
nakonec totiž
takže

Explikace: sekundární konektory

to znamená, že
znamená to, že
argumentují tím, že
vzhledem k tomu, že
což znamená, že
s tím, že
– ostatně
argumentoval tím, že
argumentuje, že
argumentují, že
důvod je zřejmý
důvod spočívá v tom, že
jak je vidět
přeloženo
stejně
to se dá přičíst tomu, že
to znamená jediné
upřesnil
upřesňuje
ve skutečnosti
vzhledem k tomu, co
znamená to toto
znamenalo to, že

Generalizace: primární konektory

tedy
tak
takže
–
– což
– tedy
:
a tak
a tedy
čili
tj.
totiž

Generalizace: sekundární konektory

prostě
zkrátka
zjednodušeně řečeno
jak je vidět
jinak řečeno
krátce řečeno
– prostě
dospěl k závěru, že
došli zkrátka k závěru, že
jednoduše řečeno
jinými slovy
v každém případě
v podstatě
v podstatě to znamená, že
z toho plyne, že
závěr je
závěr je tedy následující
znamená to, že

Gradace: primární konektory

navíc	nejenže_ale	ještě
dokonce	nejen_ale rovněž	koneckonců
nejen_ale i	natož pak	nebo pouze
a dokonce	ale především	nejen, že_ale
a navíc	#neg jen_ale také	zejména
nejen ale	ba	nejen_ale ani
a ještě	a	nejen_ale naopak
však	– a to	nejen_ale přímo
#neg jen	– dokonce i	nejen_ale rovnou
přítom	#neg jen_ale též	nejen_ale též
a to	#neg jen_nýbrž	nejen_ale zároveň
nejen_ale také	#neg jenom	nejen_navíc i
a hlavně	#neg jenom_ale i	nejen_nýbrž i
nejen_ale dokonce	#neg pouze_ale	nejenom i
i	#neg pouze_ale i	nejenom_ale také
#neg jen_ale i	#neg pouze_i	nejenom_ale
nebo dokonce	#neg pouze_nýbrž	zejména
a i	#neg_ale i	nejenomže_ale
nejenže_ale i	#neg_ale také	nejenže_ale ani
nadto	#neg_i	nejenže_ale ještě
ale	:	nejenže_navíc
#neg jen_ale	a dokonce i	nejenže_rovněž
ani	a to i	pokud
nejen_ale především	a to jen	příčemž
ale i	a to ještě	příčemž ještě
nejenže	ale dokonce	příčemž navíc
a především	ale například i	přítom navíc
nejenom ale i	ale spíše	však i
#neg jen_ale především	ale třeba i	však navíc
ovšem	avšak	však nejen
#neg jen_i	ba co víc	však současně
– dokonce	ba i	však také
ale také	či	vždyť
ale ani	či spíše	zato
nejenže_ale také	dokonce ani	
dokonce i	dokonce aniž	
nejen	jednak_a hlavně	
natož	jednak_především však	

Gradace: sekundární konektory

kromě toho	kromě toho, že
nedosti na tom	ne dosti na tom
nehledě na to, že	nejen kromě toho
a ještě k tomu	nemluvě o tom, kdy
a k tomu	nemluvě o tom, že
a k tomu ještě	o to víc
k tomu	tím spíše, že
kromě toho ovšem i	

Konfrontace: primární konektory

zatímco	a ne	už jen
naopak	a přitom	však mezitím
však	a zrovna tak	však naopak
ale	ale naopak	zároveň naopak
zase	aniž	zároveň však
ovšem	i když	zas
a naopak	jakkoliv_tak	zase naopak
zato	jestliže_tak	zatímco naopak
a	kdežto naopak	zato ale
kdežto	nakonec	zato však
přitom	ne	
jenže	nikoli	
nicméně	ovšem také	
sice_ale	přestože	
avšak	příčemž	
když	sice však	
jestliže	sice zato	
-li	spíše však	
–	tak	
jinak	už	

Konfrontace: sekundární konektory

naproti tomu	na jedné straně_na druhé straně
na druhé straně	na jedné straně_na druhé straně však
oproti tomu	na první pohled však
na druhé straně však	na rozdíl od toho
na druhou stranu	na straně druhé
rozdíl je v tom, že	na straně jedné_a na straně druhé
ve skutečnosti	s rozdílem, že
a naproti tomu	s tím kontrastuje skutečnost, že
ale na druhé straně	sice_ale na druhé straně
avšak na druhé straně naopak	v opačném případě
místo toho	v prvním případě_ve druhém případě
na jedné straně_a na druhé straně	v tom, že zatímco
na jedné straně_ale na druhé straně	však na druhé straně
na jedné straně_jenže na druhé straně	

Konjunkce: primární konektory

a	stejně jako	jako
také	ještě	jednak_a dále
což	aby	jednak_a jednak
dále	ale také	jednak_také
rovněž	a tak	jestliže
přítom	jak tak	mimo jiné
příčemž	a přítom	nebo
pak	a taky	nejen ale i
i	mimoto	obdobně pak
navíc	jinak	ostatně i
zároveň	ovšem také	podobně i
a také	jednak_a	především však
a to	a navíc	při tom
–	však také	příčemž i
a i	dále pak	příčemž ještě
ani	tedy	příčemž zároveň
však	– a	přítom také
zase	mimo to	resp.
aniž	potom	sice
čímž	a zase	sice_a
současně	– a nakonec	stejně tak ale
:	a dále	stejně tak i
a ani	a dokonce i	tak
ale	a hned pak	tak tedy
stejně tak	a ještě	totiž
když	a nakonec	však i
též	a potom	vždyť také
ovšem	a přímo	zas
či	a případně i	zatímco
a stejně tak	a současně i	zčásti_zčásti
a rovněž	a též	
a pak	a zároveň	
dokonce	a_b	
taky	ale ani	
jednak_jednak	ale přítom	
ani_ani	dál	
čemuž	i ale	

Konjunkce: sekundární konektory

dodal
podobně
a dodal
dodává
kromě toho
s tím, že
a dodává
v této souvislosti
dodala
právě tak
a pokračuje
jak dodal
do třetice
pokračuje
a kromě toho
doplňuje
k čemuž
za což
na druhé straně
v poslední řadě
vedle toho
kromě toho
k tomu
kromě toho, že
dodejme
pokračoval
doplnil
a pokračoval
kromě toho však
stejně
a do třetice
a dodal jsem
a dodejme
a doplňuje
a chtělo by se dodat
a konečně
a pokračovali
a za druhé
dále dodal
dodávám

dodejme rovněž
chtěli bychom dodat
jak dodali
jak dodává
jak_tak případně
je třeba dodat
jednak_a kromě toho
jednak_a za druhé
jednak_na straně druhé zase
ještě k tomu
k tomu třeba dodat
kromě toho také
mohl bych pokračovat
mohli bychom pokračovat
na druhé straně zase
na druhou stranu
nemluvě o tom, že
o což
o čemž
od čehož
ostatně
pak dodal
pokračovala
pokud budou pokračovat
popř.
právě tak jako
případně
připustil v této souvislosti, že
s tím
s tím, že
spolu s tím rovněž
stejným dechem
upřesnil
v čemž
vedle toho, že i
z čehož
za druhé
zprvé_a za druhé

Konjunktivní alternativa: primární konektory

nebo
či
anebo
buďto
■
nebo i

Konjunktivní alternativa: sekundární konektory

případně
popřípadě

Nepravá podmínka: primární konektory

pokud
jestliže

-li

:

-li_tak

jestli_pak

Nepravá podmínka: sekundární konektory

v případě, že

Nepravá příčina a důsledek: primární konektory

totiž
tedy
proto
takže
tak
protože
:
a tedy
aby
neboť
proto, že
přece

Nepravá příčina a důsledek: sekundární konektory

kvůli tomu, že
o důvod víc
to znamená, že

Npravý kontrast: primární konektory

však

ale

ovšem

sice_ale

i když

nicméně

ať už

avšak

i když přesto

jenže

přitom

sice_ale zároveň

Npravý kontrast: sekundární konektory

Neobsahuje.

Nesoučasnost (předčasnost a následnost): primární konektory

pak	potom, co	mezitím
když	nejprve_potom	napřed_a teprve
poté, co	až nakonec	potom
a pak	už předtím	napřed_pak
potom	dále	nejdřív_pak
než	nejprve_a pak	nejdřív_potom
až	nejprve_a poté	nejdříve_potom
poté	po tom, co	nejprve
jakmile	nadále	nejprve záhy
nakonec	dříve, než	nejprve_dále
později	zprvu_později	nejprve_poté
předtím	a bezprostředně poté	nejprve_později
a potom	a konečně	nejprve_vzápětí
a poté	a naopak vzápětí	po tom, kdy
vzápětí	a posléze i	potom, kdy
dokud	a poté i	před tím
posléze	a předtím	tak
předtím, než	a teprve potom	tehdy, když
sotva	aby	teprve
a posléze	aby nakonec	teprve poté
a vzápětí	aby posléze	záhy poté
nejprve_pak	až když	
poté, kdy	až poté, kdy	
před tím, než	hned vzápětí	
kdykoliv	jak	
nejdříve_pak	jak tak	
následně	ještě předtím	
a později	již dříve	
a	kdykoli	
a nakonec	když předtím	

Nesoučasnost (předčasnost a následnost): sekundární konektory

pokračoval	mají pokračovat
pokračuje	následoval poté, co
krátce nato	následovat bude
a bude pokračovat	od okamžiku, kdy
a na závěr	pokračovala
a pokračovat bude	pokračovat by mělo
aby pokračovala	pokračují
bude pokračovat	s tím, že pak
jako poslední	v tom pokračují
má pokračovat	v závěru

Opozice: primární konektory

však	sice_zato	nikoliv
ale	však také	nikoliv ovšem
ovšem	ale přece	ovšem také
sice_ale	a přesto	pak
jenže	–	přece jenom
avšak	sice_přesto	přestože
nicméně	a zase	rovněž
sice_však	což	sice
přitom	však i	sice_ale aspoň
aniž	zároveň ale	sice_ale i tak
a	ale zato	sice_ale přesto
zato	současně však	sice_ale taky
a přitom	a současně	sice_jenže
zároveň však	– přitom	sice_jenže také
naopak	#neg jen_a	sice_nicméně
ale zároveň	#neg pouze	sice_nicméně však
sice_ovšem	#neg_ale	sice_přesto však
sice_avšak	#neg_jenom	tak ale
i když	: ovšem	tedy
jenomže	a to	už
nýbrž	ale ani	už ovšem
příčemž	ale i tak	však přece i
však zároveň	ale ne	však přesto
přesto	ale přece jen	zas
leč	což ale	zatímco
a ne	což ovšem	
a naopak	čímž	
ale také	i však	
přesto však	jenže také	
přitom však	jinak však	
sice_ale zato	leč přece jen	
ale i	naopak však	
ale přesto	nikoli však	

Opozice: sekundární konektory

na druhé straně	navzdory tomu však
na druhou stranu	nutno ovšem dodat
na druhé straně však	o to víc
ale místo toho	sice_ale ve skutečnosti
ale na druhé straně	současně ale
ale o to víc	v čemž
argumentovali tím, že	v této souvislosti
dlužno dodat, že	ve skutečnosti
na druhé straně ovšem	však nutno dodat
na druhé straně zase	však ve skutečnosti
na jednu stranu	z jedné strany_ze strany druhé
naproti tomu	a stejně

Podmínka: primární konektory

pokud	zejména pokud	kdyby_pak
-li	a	když už
když	pakliže	-li_potom
kdyby	tehdy_pokud	ovšem
jestliže	jestli_pak	pokud ovšem
pokud_pak	pokud_potom	pouze tehdy_pokud
pokud_tak	jestliže_tak	protože
dokud	jen když	protože pokud
-li_pak	když_tak	především pokud
jestli	pouze tehdy_-li	při tom
jakmile	■	tak
kdyby_tak	– leda	tehdy
jestliže_pak	– pokud	tehdy, až
pak	a pokud	tehdy, jestliže
tehdy_když	a také	teprve potom
jak	aniž	zdali
potom	ať už	zvláště když
tehdy_-li	ať už_-li	
i kdyby	až poté	
jestli_tak	až tehdy, až	
-li_tak	dokud tak	
až	i tehdy, kdy	
kdyby_pak	jako vždy, když	
ale	jen tehdy_-li	
a pak	jestliže_potom	
i když	jinak	
ať	kdyby_potom	

Podmínka: sekundární konektory

v případě, že	podmínkou je
za podmínky, že	podmínkou je, že
v takovém případě	pouze v případě, když
v tomto případě	pouze v případě, že
v tom případě	pouze v těch případech, kdy
podmínkou je, aby	s jednou podmínkou
s tím, že	stejně
v případě, kdyby	v opačném případě
v případech, kdy	v případě
s podmínkou, že	v případě toho, že
byla podmínkou_v opačném případě	v případě, kdy
jedním z těchto důvodů je skutečnost, že	v tom případě, že
jen v případě, že	ve spojení s tím, že
jinými podmínkami bylo, že	za podmínky, kdyby
podmínka je to nevyhnutelná	

Příčina a důsledek: primární konektory

protože	poněvadž	koneckonců
proto	: totiž	například
totiž	tím	neb
neboť	proto také	pak ovšem
tedy	právě proto	pak totiž
takže	přece	poté, co
a tak	a potažmo	poté, kdy
proto, že	čili	právě proto, že
a proto	-li	proto, aby
tak	– proto	protože_tak
:	– vždyť	přitom
–	a i proto	také proto
když	a také proto	takže nakonec
vždyť	a tím i	totiž navíc
a tudíž	a to	za to
jelikož	a to proto	
pak	ale	
a	což	
tím, že	čímž	
a tedy	dále	
tudíž	i	
aby	jestliže	
– totiž	ježto	
a tím	jinak	

Příčina a důsledek: sekundární konektory

vzhledem k tomu, že	ze stejného důvodu
za to, že	z jednoduchého důvodu
z tohoto důvodu	z těchto důvodů
díky tomu	je to z toho důvodu, že
díky tomu, že	výsledkem je, že
v této souvislosti	zásluhou toho, že
z toho důvodu, že	stejně
tím pádem	a jako důvod uvedla, že
s tím, že	a kvůli tomu
to znamená, že	a měly k tomu pochopitelné důvody
tím spíš, že	a tím pádem
kvůli tomu	a to je důvod, proč
na to, aby	a to znamená, že
k tomu, aby	a v důsledku toho
ze dvou důvodů	a v tomto případě
tím spíše	a vlastně
s odůvodněním, že	a z toho důvodu
upřesnil	argumentoval
z toho důvodu	argumentovali tím, že
důsledkem je	což znamená, že
a z tohoto důvodu	což znamenalo, že
důvodem je skutečnost, že	díky tomu, že

důvod je jednoduchý
důvod je složitější
důvod spočívá v tom, že
důvodem bylo to, že
důvodem bylo, že
důvodem je to, že
důvodem je, že
důvodů je mnoho
důvodů proč je více
důvody jsou dva
důvody jsou jiné
existuje řada důvodů, aby
existuje více důvodů, proč
existují pro to tři důvody
hlavní důvod, proč
hlavní příčinou je to, že
hlavním důvodem byla skutečnost, že
hned z několika důvodů
je to důsledek toho, že
je to důvod, proč
je to způsobeno tím, že
k tomu
k tomu přistupuje skutečnost, že
má to své důvody
mají k tomu dobrý důvod
na základě čehož
na základě toho, že
není to z důvodu, že
o důvodech není třeba příliš hovořit
o to víc, že
o to více
o to více, že
od toho, aby
pro takový případ
především vinou toho
příčin proč je několik
příčinu vidí v tom, že
přispívá k tomu fakt, že
s tím, aby
nemyslím si, že by to bylo z důvodu, že

tato skutečnost se stala jedním z
důvodů, kvůli kterým
tím argumentoval
to byl důvod, proč
to byl hlavní důvod, proč
to byl jeden z hlavních důvodů, proč
to je důvod, proč
to je důvodem k tomu, aby
to je výklad možných příčin toho, že
to lze přičíst skutečnosti, že
to má svůj důvod
to má závažné důsledky
to může znamenat pouze jediné
to nebyl důvod, proč
to odůvodnil tím, že
to znamenalo, že
v souvislosti s tím, že
ve skutečnosti
vede k tomu skutečnost, že
vysvětlují si to dvěma důvody
vzhledem k těmto skutečnostem
vzhledem k tomu
z čehož
z čehož lze odhadnout, že
z důvodu, že
z jednoduchých důvodů
z mnoha důvodů
z různých důvodů, mezi nimiž
z toho důvodu, aby
z toho plyne, že
z toho vyplývá, že
z toho, když
z tohoto plyne
z uvedených důvodů
za tímto účelem
zdůvodnil to tím, že
zdůvodňoval by to tím, že
ze zcela jiných příčin totiž
znamená to, že

Přípustka: primární konektory

i když	i tak	i kdyby_tak
přesto	však přesto	i přesto, že
přestože	a přitom	i tak ale
ale	sice_ovšem	i tak však
však	avšak	jestliže
ačkoli	přesto ale	jinak
přitom	sice však	kdyby
ačkoliv	a	poté, co
ač	ale i tak	přece jen
byť	sice_ale přesto	přitom zase
přesto, že	jenže	sice
ať	i přesto, že	sice_nicméně
nicméně	i přes to, že	současně však
i kdyby	a to	tedy
třebaže	– a přitom	tím ovšem
ať už	– přesto	zároveň však
přesto však	a současně	zatím ale
sice_ale	ale nakonec	
aniž	ale přece	
ovšem	ale přitom	
když	ale zároveň	
jakkoli	aniž tak	
a přece	ať již	
i přesto	což však	
a přesto	dokonce i když	
ale přesto	i	

Přípustka: sekundární konektory

to neznamená, že	fakt neznamená, že
to však neznamená, že	i přes tato fakta
to ovšem neznamená, že	když ve skutečnosti
což neznamená, že	na rozdíl od toho, co
navzdory tomu, že	navzdory tomu
to ale neznamená, že	neznamená to, že
ale to neznamená, že	o to více
a stejně	přes tyto skutečnosti
a ve skutečnosti	to ale znamenalo, že
ač ve skutečnosti	v této souvislosti nezbyvá dodat, že
ale stejně	však navzdory těmto faktům
ale ve skutečnosti	vzhledem k tomu, že
ani v tomto případě	z faktu nevyplývá, že
argumentovalo tím, že	

Rektifikace: primární konektory

#neg_ale	#neg_jen i	kdyby
#neg	#neg_a naopak	naprosto ne
#neg_nýbrž	nikoliv_nýbrž	ne pouze
#neg_jen	– tedy	nebo spíš
#neg_spíše	#neg –	nebo spíše
ale	#neg_a spíš	nikoli pouze
a ne	#neg_a spíše	nikoli tedy
#neg_pouze	#neg_a spíše jenom	nikoli však
ne_ale	#neg_ale daleko spíše	nikoliv ale
ne	#neg_ale pouze	nikoliv jenom
#neg_však	#neg_ale především	nikoliv však
naopak	#neg_ale spíš	pouze
nýbrž	#neg_ale zato	pouze_nikoli
nikoli	#neg_ani_jako spíš	resp.
#neg_a	#neg_ba naopak	respektive
spíše	#neg_jako	sice_ale
a nikoliv	#neg_mnohem spíše	teda
#neg_spíš	#neg_přítom	totiž
#neg_naopak	#neg_sice_však	zato
tedy	a naopak	
ne však	a nikoli	
#neg_ale jen	ale ne	
spíš	ani ne_ale spíše	
však	ani ne_jako spíše	
nikoli_ale	či	
#neg_ale spíše	i nebo spíše právě	
ne_nýbrž	jen	
#neg_jenom	ještě spíše	

Rektifikace: sekundární konektory

místo, aby	anebo přesněji
místo toho	lépe řečeno
ale ve skutečnosti	místo toho však
místo toho, aby	pravděpodobnější verzí je, že
#neg_raději	přesněji
– vlastně	přesněji řečeno
#neg_rozhodně	rozdílem je skutečnost, že
#neg_místo toho	s rozdílem, že
#neg_to jenom	ve skutečnosti
anebo ještě přesněji	

Restriktivní opozice: primární konektory

však	—	ovšem přeci jen
ale	a	přesto však
ovšem	a to	přitom jen
jen	ale jen	sice_ale pouze
jenže	ale však	sice_nicméně
pouze	alespoň	sice_ovšem
nicméně	aniž	sice_však
i když	aspoň	snad jen
jinak	avšak jen	však přeci jen
sice_ale	avšak pouze	
avšak	i když pouze	
přitom	jen však	
jenom	jenže i potom	
však pouze	když	
ledaže	ne	
než	ovšem pouze	

Restriktivní opozice: sekundární konektory

výjimkou je
jen v případě, že
na jedné straně_na druhé straně však
ovšem jen v případě, že
ovšem s tím, že
ovšem za podmínky, že
pouze v případě, že
s výjimkou případů, kdy
vyjma případu, že

Současnost: primární konektory

když
zároveň
zatímco
současně
mezitím
a zároveň
dokud
až
a současně
při tom
souběžně
přitom
a
kdykoliv
jakmile
– a zároveň
a při tom
až když
dokud tak
i tehdy, kdy
mezi tím
tehdy, když
zatím

Současnost: sekundární konektory

ve stejném okamžiku

Specifikace: primární konektory

:
—
když
a to
a
totiž
příčemž
přítom
tj.
tedy
a nadále
a přítom
a zejména
čili
neboť
ovšemže
především pak
tak
vždyť
zejména

Specifikace: sekundární konektory

upřesnil
s tím, že
přesněji řečeno
to znamená, že
jak upřesňuje
jmenovitě
konkrétně
na prvním místě
přesněji
stručně řečeno
upřesňuje
v první řadě

Účel: primární konektory

aby
proto, aby
pro to, aby
tak, aby
– aby
aby, tak
dokud
jestli
-li
zda

Účel: sekundární konektory

k tomu, aby
na to, aby
od toho, aby
za tímto účelem
k tomuto účelu
pro případ, že
pro tento účel
s tím, aby
s tím, že
za tím účelem

Příloha 2: Ukázka textu z Pražského závislostního korpusu s anotací významových diskurzních vztahů vyjádřených konektory

V této příloze představujeme ukázkou anotace konkrétního textu z Pražského závislostního korpusu (PDT). Nejprve uvádíme text jako celek pouze s označením diskurzních konektorů a poté znovu jednotlivé úseky textu s naznačením diskurzních vztahů (v PDT jsou diskurzní vztahy anotované na závislostních stromech, zde vztahy kvůli lepší přehlednosti značíme na lineárním textu).

Příkladový text *Referendum jako léčka* se skládá z 23 větných nebo souvětých celků (jedním celkem rozumíme úsek textu ohraničený počátečním velkým písmenem na jedné straně a na druhé straně koncovým interpunkčním znaménkem, nejčastěji tečkou). Tyto celky jsou rozděleny do 4 odstavců (odstavce jsou označeny vynecháním řádku).

V celém textu najdeme 20 diskurzních vztahů (11 vnitrovětných a 9 mezivětných) vyjádřených primárními nebo sekundárními konektory (vztahy vyjádřené primárními konektory jsou označeny modře, vztahy vyjádřené sekundárními konektory jsou označeny žlutě). Vztahy vyjádřené neuniverzálními volnými konektivními frázemi se v příkladovém textu nevyskytují. Implicitní diskurzní vztahy v textu (typu *Obloha se zatáhla. Bude pršet.*) nejsou (v souladu se současnou anotací diskurzních vztahů v PDT) anotovány.

Konektory jsou v anotovaném textu označeny velkými písmeny. Diskurzní konektor stojí mezi dvěma diskurzními argumenty (v tomto případě vždy mezi dvěma úseky textu)⁷⁵, mezi nimiž signalizuje určitý vztah (tj. diskurzní vztah si můžeme představit jako následující schéma: *argument 1 – konektor – argument 2*). První diskurzní argument je značen tučným písmem. Druhý diskurzní argument je značen podtržením.

Jak je vidět v textu, diskurzní argumenty mohou být různě rozsáhlé – argumentem může být například jedna vedlejší věta (jako je tomu např. v celku 15) nebo i několik souvětých celků, jako je tomu např. v případě vztahu výčtu⁷⁶ (tento vztah značíme v příkladovém textu zeleně) v celcích 5–9.

⁷⁵ Diskurzním argumentem může být i situace, tj. diskurzní argument nemusí být ve všech případech vyjádřen slovy.

⁷⁶ Vztah výčtu není v PDT anotován v rámci ostatních významových vztahů v textu vyjadřovaných konektory (viz Poláková a kol., 2013), ale je zpracováván samostatně (technicky odlišně). Od ostatních vztahů se totiž výčet liší – nenaplnuje klasické schéma *argument 1 – konektor – argument 2*. Vztah výčtu (seznamu) vychází zpravidla od nominální fráze a vztahuje se většinou k větším úsekům textu, jako je celý odstavec (srov. *dvě možnosti – 1., 2.; tři příklady – první, druhý, třetí* apod.).

Jeden diskurzivní argument může být zároveň součástí jiného argumentu (naopak konektory součástí argumentů nejsou⁷⁷). Dva různé argumenty mohou být dokonce součástí jednoho rozsáhlejšího argumentu – jako je tomu např. v celku 15 (mezi argumentem 1 *která by se jimi mohla vybavit* a argumentem 2 *naučit se je používat* je vztah gradace vyjádřený konektorem *natož*; oba tyto argumenty jsou pak v jiném vztahu součástí jednoho většího argumentu: argumentem 1 je v tomto případě úsek textu *západní mocnosti stihnou dovézt zbraně*, argumentem 2 pak úsek textu *nebude už existovat armáda, která by se jimi mohla vybavit, natož se je naučit používat*, mezi oběma argumenty je v tomto případě vztah nesoučasnosti vyjádřený konektorem *než*). V anotaci příkladového textu jsou diskurzivní vztahy řazeny za sebou od vztahů mezi nejnižšími textovými jednotkami, které se v dalších vztazích mohou stát součástí jednotek rozsáhlejších.

Od jednoho argumentu mohou vést zároveň dva různé významové vztahy: v příkladovém textu můžeme tento jev částečně pozorovat ve větě celku 10 (jsou zde v těsné blízkosti dva různé konektory, srov. *POKUD se ALE Západu ... podaří...* – jeden konektor vyjadřuje vztah podmínky, druhý vztah opozice). Dva různé významové vztahy mezi dvěma identickými argumenty můžeme pozorovat například v souvětí: *V sobotu přijde teplo, ale potom se ochladí* (mezi argumentem 1 *v sobotu přijde teplo* a argumentem 2 *ochladí se* je jednak vztah opozice vyjádřený konektorem *ale*, jednak vztah nesoučasnosti vyjádřený konektorem *potom*).

Na rozboru významových vztahů v ukázkovém textu dále vidíme, že jednotlivé sémantické kategorie jsou chápány poměrně široce (v některých případech můžeme hledat např. další významové odstíny mezi dvěma různými příklady příčiny), někdy v praxi dochází k překrývání jednotlivých kategorií (např. vztah mezi celky 20 a 21 je označen jako vztah příčiny a důsledku, ale můžeme v něm vidět i rysy vztahu explikace).

V rámci anotace explicitních diskurzivních vztahů nebyly v PDT značeny vztahy čistě přívlastkové (např. *sankcí, které ruší hospodářství srbsko-černohorské Jugoslávie*) a vztahy mezi větou řídicí a předmětnou (např. *Západu se podaří ukázat, že slova o zrušení embarga jsou míněna vážně.*) nebo podmětnou (např. *Je hezké, že přijdeš.*). Informace o existenci těchto vztahů je však v PDT zachycena v rámci anotace tzv. tektogramatické roviny (tj. roviny hloubkové syntaxe).

Jsme si vědomi toho, že na celkové koherenci textu se kromě vztahů vyjádřených konektory (a volnými konektivními frázemi) podílejí i vztahy vyjádřené jinými jazykovými

⁷⁷ Výjimku tvoří konektory typu *dodat, upřesnit* apod. – viz Rysová (2014).

jevy a prostředky – například elipsou, aktuálním členěním větným (například tematickými posloupnostmi) nebo anaforickými a koreferenčními řetězci. PDT obsahuje anotaci všech těchto zmíněných jevů. V rámci anotace diskurzu jsou však v současné době v PDT značeny pouze textové vztahy vyjádřené explicitními konektivními prostředky typu *ale*, *protože*, *z tohoto důvodu*, *z toho vyplývá*.

Příkladový text

(1) *Referendum jako léčka*

(2) *Petr Dudek*

(3) *Bosenští Srbové učinili, co předem slibovali.*

(4) *V referendu odmítli navrhovanou mapu rozdělující území Bosny a Hercegoviny.*

(5) *Před kontaktní skupinou, spojující čtveřici západních velmocí a Rusko, tím pádem stojí dvě možnosti.*

(6) *Jednak může mezinárodní společenství vyjít vstříc požadavkům Radovana Karadžiče a jeho nacionalistické kliky a začít navrhovanou mapu překreslovat tak, aby více vyhovovala poměrům nastoleným vojenskou silou.*

(7) *Jinými slovy, odměnit bosenské Srby, kteří nyní ovládají sedmdesát procent území, za jejich agresí.*

(8) *Pokud k něčemu takovému dojde, což není vyloučeno, bude to, řečeno slovy českého vyslance v OSN Karla Kovandy, debakl západoevropské a severoamerické politiky.*

(9) *Druhou možností je to, že západní mocnosti se Spojenými státy v čele odmítnou požadavky Srbů v Bosně a prosadí zrušení zbrojního embarga proti bosenským Muslimům.*

(10) *Pokud se ale Západu – s Moskvou nelze v tomto směru počítat – podaří ukázat, že slova o zrušení embarga jsou míněna vážně, nebudou bosenští Srbové nejspíše čekat, až se situace v poli změní v jejich neprospěch.*

(11) *Místo toho zahájí Karadžičovy jednotky ofenzívu proti Muslimům.*

(12) *S jasným cílem: rozhodnout vleklou bosenskou válku jednou provždy.*

(13) *Nikdo nemůže s jistotou předvídat, zda se jim to podaří.*

(14) *Jisté je pouze to, že odhodlání bojovat bosenským Srbům nechybí a vidina lépe vyzbrojených Muslimů jejich rozhodnost pouze zvyšuje.*

(15) *Než západní mocnosti stihnou dovézt zbraně, nebude už existovat armáda, která by se jimi mohla vybavit, natož se je naučit používat, pohrozil nedávno jeden z bosenskoserbských vojenských velitelů.*

(16) *Bosenští Srbové se svým referendem – o jehož výsledku Karadžič nemohl pochybovat – pokusili zdramatizovat situaci v naději, že západní mocnosti a Rusko navzájem zablokují vliv, který na situaci v Bosně mají.*

(17) *Jejich taktiku posílil i Slobodan Milošević.*

(18) *Bělehrad ucítil příležitost využít bosenského dilematu a domoci se zrušení mezinárodních sankcí, které ruínují hospodářství srbsko-černohorské Jugoslávie.*

(19) *Neváhá kvůli tomu vystupovat i proti svým tradičním spojencům v Bosně.*

(20) *Podle reakcí ruského ministra zahraničí Kozyreva soudě, není Miloševićova nabídka Moskvě nesympatická.*

(21) *Tím spíše nesmí Západ výsledky referenda přecenit.*

(22) *Nátlak na bosenské Srby musí pokračovat.*

(23) *Nemůže být ale uplatňován přes Slobodana Miloševiče, i když jeho prsty v Bosně sahají velmi daleko.*

Anotace příkladového textu

1. vztah příčina a důsledek

(4) *V referendu odmítli navrhovanou mapu rozdělující území Bosny a Hercegoviny.*

(5) *Před kontaktní skupinou, spojující čtveřici západních velmocí a Rusko, TÍM PÁDEM stojí dvě možnosti.*

2. vztah účel⁷⁸

(6) *[může] začít navrhovanou mapu překreslovat tak, ABY více vyhovovala poměrům nastoleným vojenskou silou.*

3. vztah konjunkce

(6) *může mezinárodní společenství vyjít vstříc požadavkům Radovana Karadžiče a jeho nacionalistické kliky A začít navrhovanou mapu překreslovat tak, aby více vyhovovala poměrům nastoleným vojenskou silou.*

4. vztah ekvivalence

(6) *může mezinárodní společenství vyjít vstříc požadavkům Radovana Karadžiče a jeho nacionalistické kliky a začít navrhovanou mapu překreslovat tak, aby více vyhovovala poměrům nastoleným vojenskou silou.*

(7) *JINÝMI SLOVY, odměnit bosenské Srby, kteří nyní ovládají sedmdesát procent území, za jejich agresi.*

5. vztah konjunkce

(8) *k něčemu takovému dojde, COŽ není vyloučeno*

6. vztah podmínka

(8) *POKUD k něčemu takovému dojde, což není vyloučeno, bude to, řečeno slovy českého vyslance v OSN Karla Kovandy, debakl západoevropské a severoamerické politiky.*

7. vztah konjunkce

(9) *západní mocnosti se Spojenými státy v čele odmítnou požadavky Srbů v Bosně A prosadí zrušení zbrojního embarga proti bosenským Muslimům.*

8. vztah nesoučasnost

(10) *nebudou bosenští Srbové nejspíše čekat, AŽ se situace v poli změní v jejich neprospěch.*

9. vztah podmínka

(10) *POKUD se Západu – s Moskvou nelze v tomto směru počítat – podaří ukázat, že slova o zrušení embarga jsou míněna vážně, nebudou bosenští Srbové nejspíše čekat, až se situace v poli změní v jejich neprospěch.*

10. vztah opozice

(9) *západní mocnosti se Spojenými státy v čele odmítnou požadavky Srbů v Bosně a prosadí zrušení zbrojního embarga proti bosenským Muslimům.*

⁷⁸ Vztah vyjádřený konektorem *aby* se v rámci klasifikace diskurzních vztahů v PDT nejvíce blíží vztahu účelu, nicméně v tomto případě nese spíše rysy vztahu způsobu (ten se ovšem v klasifikaci diskurzních vztahů v PDT neobjevuje).

(10) Pokud se ALE Západu – s Moskvou nelze v tomto směru počítat – podaří ukázat, že slova o zrušení embarga jsou míněna vážně, nebudou bosenští Srbové nejspíše čekat, až se situace v poli změní v jejich neprospěch.

11. vztah rektifikace

(10) nebudou bosenští Srbové nejspíše čekat, až se situace v poli změní v jejich neprospěch.

(11) MÍSTO TOHO zahájí Karadžičovy jednotky ofenzívu proti Muslimům.

12. vztah účel

(11) zahájí Karadžičovy jednotky ofenzívu proti Muslimům.

(12) S JASNÝM CÍLEM: rozhodnout vleklou bosenskou válku jednou provždy.

13. vztah konjunkce

(14) Jisté je pouze to, že odhodlání bojovat bosenským Srbům nechybí A vidina lépe vyzbrojených Muslimů jejich rozhodnost pouze zvyšuje.

14. vztah nesoučasnost

(15) NEŽ západní mocnosti stihnou dovézt zbraně, nebude už existovat armáda, která by se jimi mohla vybavit, natož se je naučit používat

15. vztah gradace

(15) kteřá by se jimi mohla vybavit, NATOŽ se je naučit používat

16. vztah výčet (seznam)

(5) dvě možnosti

1. položka seznamu:

(6) JEDNAK může mezinárodní společenství vyjít vstříc požadavkům Radovana Karadžiče a jeho nacionalistické kliky a začít navrhovanou mapu překreslovat tak, aby více vyhovovala poměrům nastoleným vojenskou silou.

(7) Jinými slovy, odměnit bosenské Srby, kteří nyní ovládají sedmdesát procent území, za jejich agresí.

(8) Pokud k něčemu takovému dojde, což není vyloučeno, bude to, řečeno slovy českého vyslance v OSN Karla Kovandy, debakl západoevropské a severoamerické politiky.

2. položka seznamu

(9) DRUHOU MOŽNOSTÍ JE TO, ŽE západní mocnosti se Spojenými státy v čele odmítnou požadavky Srbů v Bosně a prosadí zrušení zbrojního embarga proti bosenským Muslimům.

(10) Pokud se ale Západu – s Moskvou nelze v tomto směru počítat – podaří ukázat, že slova o zrušení embarga jsou míněna vážně, nebudou bosenští Srbové nejspíše čekat, až se situace v poli změní v jejich neprospěch.

(11) Místo toho zahájí Karadžičovy jednotky ofenzívu proti Muslimům.

(12) S jasným cílem: rozhodnout vleklou bosenskou válku jednou provždy.

(13) Nikdo nemůže s jistotou předvídat, zda se jim to podaří.

(14) Jisté je pouze to, že odhodlání bojovat bosenským Srbům nechybí a vidina lépe vyzbrojených Muslimů jejich rozhodnost pouze zvyšuje.

(15) Než západní mocnosti stihnou dovézt zbraně, nebude už existovat armáda, která by se jimi mohla vybavit, natož se je naučit používat, pohrozil nedávno jeden z bosenskosrbských vojenských velitelů.

17. vztah příčina a důsledek

(18) *Bělehrad ucítil příležitost využít bosenského dilematu a domoci se zrušení mezinárodních sankcí, které ruínují hospodářství srbsko-černohorské Jugoslávie.*

(19) *Neváhá KVŮLI TOMU vystupovat i proti svým tradičním spojencům v Bosně.*

18. vztah příčina a důsledek

(20) *Podle reakcí ruského ministra zahraničí Kozyreva soudě, není Miloševićova nabídka Moskvě nesympatická.*

(21) *TÍM SPÍŠE nesmí Západ výsledky referenda přecenit.*

19. vztah přípustka

(23) *Nemůže být uplatňován přes Slobodana Miloševiče, I KDYŽ jeho prsty v Bosně sahají velmi daleko.*

20. vztah restriktivní opozice

(22) *Nátlak na bosenské Srby musí pokračovat.*

(23) *Nemůže být ALE uplatňován přes Slobodana Miloševiče, i když jeho prsty v Bosně sahají velmi daleko.*