



DISKURZNÍ KONEKTORY V ČEŠTINĚ

Od centra k periférii

Magdaléna Rysová



ÚSTAV FORMÁLNÍ
A APLIKOVANÉ LINGVISTIKY

**STUDIES IN COMPUTATIONAL
AND THEORETICAL LINGUISTICS**

Magdaléna Rysová

**DISKURZNÍ KONEKTORY V ČEŠTINĚ
Od centra k periferii**

Published by the Institute of Formal and Applied Linguistics
as the 17th publication in the series
Studies in Computational and Theoretical Linguistics. First edition, Prague 2018.

Editor-in-chief: Jan Hajič

Editorial board: Nicoletta Calzolari, Mirjam Fried, Eva Hajičová, Petr Karlík,
Joakim Nivre, Jarmila Panevová, Patrice Pognan, Pavel Straňák,
and Hans Uszkoreit

Reviewers: prof. PhDr. Eva Hajičová, DrSc.
doc. PhDr. Markéta Malá, Ph.D.
doc. PhDr. Ondřej Pešek, Ph.D.

This book has been printed with the support of the Czech Science Foundation project no. GA17-06123S (*Anaphoricity in Connectives: Lexical Description and Bilingual Corpus Analysis*) and of the institutional funds of Charles University.

The book has utilized the language resources distributed by the LINDAT/CLARIN project of the Ministry of Education, Youth and Sports of the Czech Republic (project LM2015071).

Printed by Printo, spol. s r. o.

Copyright © Institute of Formal and Applied Linguistics, 2018

ISBN 978-80-88132-05-9

Obsah

Abstrakt	3
Abstract	5
1 Úvodem	7
2 Cíle práce	9
3 Teoretické východisko	13
3.1 Kořeny textové lingvistiky ve světové literatuře	13
3.2 Kořeny textové lingvistiky v české literatuře	15
3.3 Diskuze o nadvětném útvaru – vymezení pojmu diskurz	17
3.3.1 Strukturní přístup k diskurzu	19
3.3.2 Funkční přístup k diskurzu	20
3.3.3 Diskurz jako fenomén přesahující využití jazyka v praxi	21
3.4 Co mohou spojovat diskurzní konektory	22
3.4.1 Teorie koherence	22
3.4.2 Teorie relevance	23
3.4.3 Diskurzní konektory z pohledu teorie koherence a teorie relevance	24
3.5 Vymezení jednotky diskurzu	25
3.5.1 Diskurzní argumenty v Pražském diskurzním korpusu	26
3.5.2 Povrchová realizace diskurzních argumentů – verbální vs. nominální argumenty	26
3.6 Diskurzní vztahy	28
3.6.1 Diskurzní vztahy implicitní a explicitní	28
3.6.1.1 Zpracování implicitních diskurzních vztahů ve velkých korpusech	29
3.6.2 Významová klasifikace diskurzních vztahů	30
3.6.2.1 Významové vztahy v textu v Pražském diskurzním korpusu	32

3.7	Diskurzní konektory	37
3.7.1	Diskurzní konektory ve světové literatuře	37
3.7.2	Diskurzní konektory v české literatuře	41
3.7.3	Základní vymezení diskurzních konektorů pro tuto práci	44
3.8	Zpracování diskurzních vztahů a konektorů ve velkých korpusech	45
3.8.1	RST Discourse Treebank	45
3.8.2	Discourse Graphbank	46
3.8.3	Penn Discourse Treebank a navazující projekty	47
3.8.4	Další významné diskurzní korpusy	47
3.8.5	Zpracování víceslovných konektivních struktur ve velkých korpusech	48
4	Jazykový materiál	51
4.1	Pražský závislostní korpus	51
4.2	Pražský diskurzní korpus	53
4.2.1	Pražský diskurzní korpus 1.0	53
4.2.2	Pražský diskurzní korpus 2.0	55
4.3	Korpus DIALOG	55
4.4	Český národní korpus	56
5	Metody a postup práce	57
6	Výsledky práce	59
6.1	Definice a dělení diskurzních konektorů	59
6.1.1	Vymezení diskurzních konektorů z funkčního hlediska	60
6.1.2	Primární a sekundární konektory	61
6.1.2.1	Slovnědruhovú charakteristika primárních konektorů	62
6.1.2.2	Charakteristika sekundárních konektorů	63
6.1.2.3	Variabilita sekundárních konektorů: lexikální realizace a gramatické varianty	66
6.1.2.4	Primární a sekundární konektory a povaha diskurzních argumentů	67
6.1.2.5	Vzájemná nahraditelnost primárních a sekundárních konektorů	69
6.1.2.6	Prostupnost hranic mezi primárními a sekundárními konektory	71
6.1.2.7	Úvahy ke vzniku primárních konektorů	72
6.1.2.8	K otázce plnovýznamovosti	74

6.1.2.9	Volné konektivní fráze	76
6.1.3	Diskurzní konektory a reference	76
6.1.4	Souvýskyt diskurzních konektorů – složené konektory	77
6.1.4.1	Modifikované konektory	79
6.2	Anotace sekundárních konektorů v Pražském diskurzním korpusu 2.0 . . .	80
6.2.1	Dvouúrovňová anotace sekundárních konektorů	82
6.2.2	Evaluační anotace	82
6.3	Vyhodnocení anotace sekundárních konektorů v PDiT	86
6.3.1	Diskurzní konektory podle frekvence	86
6.3.1.1	Rozložení primárních konektorů, sekundárních konektorů a volných konektivních frází v PDiT	86
6.3.1.2	Nejčastěji používané diskurzní konektory	87
6.3.1.3	Sekundární konektory v češtině a v angličtině	89
6.3.1.4	Vybrané sekundární konektory v psané a mluvené češtině	90
6.3.2	Sémantické typy diskurzních vztahů vyjádřené primárními a sekundárními konektory	93
6.3.2.1	Sekundární konektory a syntaktická realizace argumentů	97
6.3.2.2	Rozšíření klasifikace diskurzních vztahů	99
6.3.3	Diskurzní vztahy mezivětné a vnitrovětné	103
6.4	Faktory ovlivňující výběr diskurzních konektorů	109
6.4.1	Testování kontextové (ne)závislosti diskurzních konektorů	109
6.4.1.1	Hranice mezi sekundárními konektory a volnými konektivními frázemi	110
6.4.1.2	Kontextová (ne)závislost vybraných primárních a sekundárních konektorů	113
6.4.2	Principy a tendence v užívání diskurzních konektorů	116
7	Závěr	123
	Summary	125
	Seznam obrázků	131
	Seznam tabulek	133

Seznam zkratk	135
A Přílohy	137
A.1 Seznamy konektorů	137
A.1.1 Abecední seznam forem primárních konektorů z Pražského diskurzního korpusu	139
A.1.2 Frekvenční seznam forem primárních konektorů z Pražského diskurzního korpusu	144
A.1.3 Abecední seznam forem sekundárních konektorů z Pražského diskurzního korpusu	149
A.1.4 Frekvenční seznam forem sekundárních konektorů z Pražského diskurzního korpusu	154
A.1.5 Frekvenční seznam forem primárních a sekundárních konektorů z Pražského diskurzního korpusu s ohledem na vyjadřované významové vztahy	159
Literatura	181
Zdroje	195
Rejstřík	197

Poděkování

Srdečně děkuji prof. PhDr. Evě Hajičové, DrSc., za cenné připomínky k textu, za ochotu a přátelský přístup. Velmi děkuji doc. PhDr. Markétě Malé, Ph.D., a doc. PhDr. Ondřeji Peškovi, Ph.D., za jejich přínosné odborné komentáře k rukopisu.

Zároveň děkuji RNDr. Jiřímu Mírovskému, Ph.D., za ochotnou technickou pomoc při zpracovávání dat *Pražského diskurzního korpusu* a za změření mezianotátorské shody. Za pomoc s anotací jazykových dat vřele děkuji Mgr. Pavlíně Synkové, Ph.D., a Veronice Scheller, M.A. Za technickou pomoc při přípravě knihy k tisku děkuji Mgr. Pavlu Straňákovi, Ph.D.

V neposlední řadě děkuji všem kolegům z Ústavu formální a aplikované lingvistiky za přátelskou atmosféru, kterou na pracovišti vytvářejí.

Magdaléna Rysová

Klíčová slova

diskurz, diskurzní konektory, diskurzní vztahy, Pražský diskurzní korpus, primární konektory, sekundární konektory, text

Key words

discourse, discourse connectives, discourse relations, Prague Discourse Treebank, primary connectives, secondary connectives, text

Abstrakt

V knize se zabýváme popisem a analýzou diskurzních (tj. textových) konektorů v češtině, tedy tím, jakými jazykovými prostředky je možné vyjadřovat v textu významové vztahy. Výzkum přitom neomezujeme na předem stanovenou skupinu výrazů (danou například příslušností k určitým slovním druhům, jako jsou souřadicí a podřadicí spojky), 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 tomu; 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 jejich 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 z anotovaných dat Pražského diskurzního korpusu (PDiT) obsahujícího psané publicistické texty. V práci představujeme principy a výsledky anotace sekundárních konektorů v PDiT. Zaměřujeme se především na srovnání primárních a sekundárních konektorů 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 PDiT a DIALOG).

Z výsledků analýzy diskurzních konektorů se snažíme vyvodit obecné teoretické závěry o zkoumaných jazykových prostředcích. 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.

mi (především na základě jejich lexikální, syntaktické a sémantické charakteristiky). V závěru práce představujeme faktory ovlivňující výběr diskurzních konektorů v textu, resp. snažíme se nalézt a popsat principy a tendence v užívání konektorů.

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.

Abstract

The book focuses on description and analysis of discourse connectives in Czech, i.e. on the issue by which language means it is possible to express sense relations within a text. The book 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 a coherent complex. The book investigates discourse connectives in Czech with respect to so-called **secondary connectives** (i.e. mainly multiword phrases like *to je důvod, proč* – *that is the reason why*; *kvůli tomu* – *because of this*; *z těchto důvodů* – *for these reasons* 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 book 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 on the annotated data of the Prague Discourse Treebank (PDiT) containing written newspaper texts. In the book, we introduce the principles and results of the annotation of secondary connectives in the PDiT. We focus mainly on the comparison of primary 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 with 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 the PDiT and in the corpus DIALOG).

On the basis of the 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). Finally, we present factors influencing selection of discourse connectives in authentic texts – we try to find and describe the principles and tendencies in using connectives.

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

1

Úvodem

Tato práce se zaměřuje na to, jaké jazykové prostředky v češtině mají schopnost vyjadřovat vztahy mezi jednotlivými částmi textu a pomáhat tak čtenáři lépe a rychleji textu porozumět.

Podle mnohých autorů (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 obsahově 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. *proto* v příkladu 1) nebo implicitně, tj. daný diskurzní vztah není vyjádřen jazykovým výrazem, ale pouze vyplývá z významu příslušných úseků textu, srov. příklad 2.

- (1) *Divadelní prostředí mě oslovilo. Proto jsem hned po gymnáziu začala studovat divadelní vědu.*
- (2) *Divadelní prostředí mě oslovilo. Hned po gymnáziu jsem začala studovat divadelní vědu.*

Některé jazykové studie (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. Těmto jazykovým prostředkům, nazývaným nejčastěji diskurzní konektory, je tedy zapotřebí v rámci výzkumu koherence textu věnovat pozornost.

Diskurzní konektory bývají mnohdy chápány jako synsémantické výrazy (srov. Halliday a Hasanová, 1976; Martin, 1992; Knott, 1996; Stede a Neumann, 2014), tj. často bývají omezeny na určité slovní druhy, především na spojky či částice strukturující text. Vzájemné diskurzní vztahy mezi jednotlivými úseky textu ovšem nemusí být vyjádřeny pouze spojkami či některými druhy částic. Ukazuje se, že škála výrazů s konektivní funkcí je podstatně bohatší (srov. Hoffmannová, 1983, 1984; Prasadová a kol., 2010; Rysová, 2012, 2013; Rysová a Rysová, 2014). Například důsledkový vztah nemusí být vyjádřen pouze výrazy typu *proto* nebo *tedy*, ale lze jej vyjádřit i víceslovnými spojeními typu *z tohoto důvodu*; *to je důvod*, *proč*; *důsledkem je* atd. Obdobnými frázemi mohou být vyjadřovány i další diskurzní vztahy, srov. spojení *podmínkou je*

pro významový vztah podmínky či *to znamená* pro vztah explikace. Předmětem naší práce je tyto konektivní prostředky podrobně popsat (na základě anotovaných korpusových dat). Jedná se většinou o víceslovné struktury, které často nebývají chápány jako diskurzní konektory, ale které spojovací funkci v rámci diskurzu zcela jistě mají a množinu diskurzních prostředků podstatně obohacují.

Problematika diskurzu a výrazů vyjadřujících diskurzní vztahy je v současné době aktuálním tématem na poli světové lingvistiky. Mnohé studie (srov. Sanders a Noordman, 2000; Blakemoreová, 2002; Prasadová a kol., 2010; Johannsen a Søgaaard, 2013, Danlosová a kol., 2018 apod.) se zabývají 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). Naše práce se snaží na tyto studie navázat a obohatit dosavadní výzkum o analýzu anotovaných korpusových dat.

2

Cíle práce

Jak jsme naznačili výše, obecně můžeme vyčlenit dva základní typy diskurzních významových vztahů – vztahy implicitní, nevyjádřené konkrétními jazykovými prostředky, ale pouze vyplývající z obsahu daných výpovědí či vět (srov. příklad 3i),¹ a vztahy explicitní, vyjádřené konkrétními jazykovými prostředky. Škála těchto prostředků je přitom značně široká, viz příklady 3ii–x, ve kterých je diskurzní vztah důsledku vyjádřený pokaždé jiným jazykovým výrazem.

- (3) *Petr je nemocný a nemůže s námi jet.*
- i) *Výlet se ruší.*
 - ii) ***Proto** se výlet ruší.*
 - iii) *Výlet se **tedy** ruší.*
 - iv) *Výlet se **tak** ruší.*
 - v) *Výlet se **tím** ruší.*
 - vi) *Výlet se **tímto** ruší.*
 - vii) ***Na základě toho** se výlet ruší.*
 - viii) ***Z tohoto důvodu** se výlet ruší.*
 - ix) ***Kvůli tomu** se výlet ruší.*
 - x) ***Kvůli jeho nemoci** se výlet ruší.*

Všechny uvedené výrazy (*proto, tedy, tak, tím, tímto, na základě toho, z tohoto důvodu, kvůli tomu, kvůli jeho nemoci*) signalizují důsledkový vztah v textu explicitně. Ne u všech je ale míra explicitnosti stejná, resp. liší se míra konkrétnosti vyjádření (srov. *proto* vs. *kvůli jeho nemoci*). V naší práci se zaměřujeme na to, které z uvedených výrazů (resp. které typy výrazů obecně) 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.

Zabýváme se přitom výrazy morfologicky, lexikálně i syntakticky neomezenými typy *důsledkem je, příčinou bylo, to je v rozporu s tím* (nazýváme je sekundární konektory), které v mnohých pojetích nejsou mezi diskurzní konektory řazeny (srov. Stede a Neumann, 2014; Halliday a Hasanová, 1976; Hrbáček, 1994 atd.).

¹ Příklady uváděné v této práci pocházejí především z Pražského diskurzního korpusu, dále z Českého národního korpusu či jsou nalezené internetovým vyhledávačem. V některých případech jsou tyto příklady pro lepší přehlednost upraveny, zejména kráceny.

Výzkum provádíme na ručně anotovaných datech Pražského diskurzního korpusu 2.0 (PDiT; Rysová a kol., 2016), který v současné době obsahuje rozšířenou anotaci diskurzních vztahů, zahrnující mj. vztahy vyjádřené právě sekundárními konektory (srov. *podmínkou je; to je důvod, proč* aj.). Dílčím cílem je tuto anotaci krátce představit a srovnat její výsledky s výsledky dřívější diskurzní anotace zaměřené především na jednoslovné konektory typu *ale, proto, zatímco* (srov. Poláková a kol., 2012a). 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.

K dosažení těchto cílů budeme v práci postupovat v následujících krocích. Nejprve se budeme zabývat 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 skupin. Diskurzní spojovací prostředky 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*. V práci budeme vycházet z českých i zahraničních studií, které se ve vymezení diskurzních konektorů značně liší. Někteří autoři (srov. 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 (*protože, i, nebo, také, poté* atd.), jiní (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 naší práci se pokusíme na základě anotovaných korpusových dat přispět k této diskusi novými poznatky, podle kterých by bylo možné jazykové prostředky vyjadřující významové vztahy v textu lépe vymezit. Pokusíme se tedy najít kritéria, podle kterých lze například stanovit škálu diskurzních konektorů na základě míry jejich gramatikalizace, vhodnosti jejich užití v různých kontextech apod.

V práci dále ukážeme, jakým způsobem je možné zachycovat sekundární konektory ve velkých korpusech. Krátce představíme obecné anotační principy, které jsme sestavili pro anotaci těchto výrazů v Pražském diskurzním korpusu. Přiblížíme také možnosti měření mezianotátorské shody, jejíž výsledky slouží jako důležitá zpětná vazba a mohou být využity pro vylepšování anotace do konečné podoby. Obdobná anotace sekundárních konektorů v češtině nebyla dosud provedena, a pokud je nám známo, nebyla provedena ani jinde v zahraničí – tyto výrazy do jisté míry zachycuje pensylvánský korpus Penn Discourse Treebank, ovšem ne v takovém měřítku jako Pražský diskurzní korpus (srov. kapitola 3.8.5).

V další části práce popíšeme výsledky vlastního výzkumu diskurzních konektorů v češtině. Pokusíme se obecně vymezit kategorii diskurzních konektorů a popsat rozdíly mezi konektory primárními (*proto, ale, tedy...*) a sekundárními (*z tohoto důvo-*

du, kvůli tomu, to znamená...). Dále sestavíme charakteristiku sekundárních konektorů po stránce lexikální, syntaktické a sémantické a popíšeme jejich vlastnosti a chování v textech. V práci také představíme podrobnou analýzu diskurzních konektorů z hlediska jejich frekvence v textu – zaměříme se především na rozdíl v četnosti užívání primárních a sekundárních konektorů a na to, které konkrétní konektory se v textech PDiT objevují nejčastěji.

Výsledky anotace sekundárních konektorů v češtině srovnáme s anotací obdobných jazykových prostředků v angličtině (na datech korpusu Penn Discourse Treebank – PDTB). Korpusy PDiT a PDTB obsahují publicistické texty, počet anotovaných vět je také srovnatelný (PDiT obsahuje přibližně 50 000 vět, PDTB 49 000 vět). Oba korpusy tak nabízejí vhodný jazykový materiál pro srovnávací studii. Pokusíme se také srovnat výskyt, charakteristiku a chování vybraných sekundárních konektorů v psané a mluvené češtině (na základě dat PDiT a korpusu mluvené češtiny DIALOG).

Zaměříme se i na to, které významové vztahy v textu diskurzní konektory vyjadřují – konkrétně které typy vztahů se vyskytují nejčastěji, které méně často, zda sekundární konektory vyjadřují stejné významové vztahy jako konektory primární či zda existují významové vztahy častěji vyjadřované sekundárními konektory než konektory primárními.

V práci dále popíšeme, v jaké míře vyjadřují primární a 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).

Na závěr v práci představíme experiment provedený na vybraných konektorech (primárních i sekundárních), které budeme testovat ve velkém množství různých kontextů a zjišťovat tak míru jejich kontextové (ne)závislosti. Obecně se tedy zaměříme na to, jaké faktory ovlivňují výběr vhodného konektivního prostředku pro určitý kontext, tj. podle jakých kritérií mohou mluvčí volit příslušný výraz, jaká existují možná omezení tohoto výběru apod.

Naše 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 (kapitoly 3.1 a 3.2). Blíže se zaměříme na vztah mezi pojmy text a diskurz² (kapitola 3.3) 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, kapitoly 3.4 a 3.5), klasifikací diskurzních významových vztahů (kapitola 3.6), vlastním vymezením diskurzních konektorů (kapitola 3.7) a jejich zachycením v různých korpusech (kapitola 3.8).

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

² 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 termín analýza diskurzu (*discourse analysis*).

Velkou pozornost věnuje van Dijk studiu tzv. makrostruktur, tj. hloubkovým obsahovým strukturám textu – na začátku osmdesátých letech 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. *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),³ intencionalita, akceptabilita, informativnost, situativnost, intertextualita.

Významnou prací věnující se textu a textové lingvistice je dále kniha *Cohesion in English* (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. kládou 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 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 rozlišují kohezi gramatickou a lexikální. V rámci gramatické koheze se věnují popisu reference (osobní, ukazovací a srovnávací), substituce (nominální, verbální, větné), elipsy (nominální, verbální, větné) a spojovacích prostředků (v naší terminologii konektorů). V rámci lexikální koheze vydělují opakování (tj. opa-

³ 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.

kování významu použitím stejného slova, synonyma, slova nadřazeného, podřazeného či obecného) a kolokace.

Dalším významným světovým autorem zabývajícím se textovou lingvistikou je Michael Hoey (srov. 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 3.3).

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. „sdružení 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. Hrbáček, 1994; viz níže).

Výzkum nadvětných jevů v českém prostředí se následně ubírá různými směry. Někteří lingvisté se zaměřují na analýzu literárních textů, jiní se zabývají popisem specifických pravidel výstavby textu (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 Jan Mukařovský (*Kapitoly z české poetiky*, 1948), Roman Jakobson (*Poetická funkce*, 1995), Lubomír Doležel (*O stylu moderní prózy*, 1960, či *Narativní způsoby v české literatuře*, 1993), Miroslav Červenka (*Významová výstavba literárního díla*, 1992) nebo 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 jevů nadvětných. Ve svých pracích mu věnuje pozornost především Vilém Mathesius (srov. 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 (např. 1962, 1978, 1992) a Daneš (např. 1957, 1968) zvolili pro svůj výzkum aktuálního členění tradiční přístup k jazyku, Sgall (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. 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 text či diskurz). Zkoumáním nadvětné syntaxe se zabývají např. Bohumil Palek (srov. *Cross-reference. A study from hyper-syntax*, 1968), Karel Hausenblas (*Gramatická syntax textu*, 1979), Jana Hoffmannová (srov. 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*, 1977, 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 (*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. 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 a zdůrazňuje nutnost osvojování zásad, pravidel a norem v komunikaci. V poslední době se komunikaci a kultuře vyjadřování věnovali např. Světlá Čmejrková (1999, 2000 aj.) či Petr Kaderka (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 (např. 1983, 1984), dále František Daneš (1985), Josef Hrbáček (1994), Přemysl Adamec (1995), Marek Nekula (1996) nebo Ondřej Pešek (2011). Více k problematice diskurzních konektorů a jejich popisu v českém a světovém kontextu viz kapitola 3.7.

3.3 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 odborné 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. Zároveň 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 nebo písně).

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 termínu promluva).

S termínem promluva (ve smyslu nejvyšší jazyková jednotka či jednotka nad větou) pracuje ve třicátých letech Skalička (1937) (ačkoli v pozdějších letech nad promluvu staví sdružení promluv, které nazývá diskurzem). Podle Hausenblase je nevýhodou termínu 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 výtvor, 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 termín používá již 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 Šgall (1973). Nevýhodou termínu diskurz je ovšem podle Hausenblase to, že na rozdíl např. od textu nemá v českém prostředí tradici.

Hausenblasův článek 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 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 Tárníkové, která se zabývá systematickým popisem základních termínů textové lingvistiky v evropském i americkém kontextu v knize *From Text to Texture* (2002).

Tárníková 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. Tárníková,

2002: 20). Na začátku každé práce, která tyto termíny používá, je tedy nutné je nejprve významově vymezit. V některých přístupech je podle Tárnyikové text spojován s psaným jazykem, diskurz s mluveným, v jiných je text chápán jako abstraktní výtvor (funkčně analogický k termínu *langue*), zatímco diskurz jako jeho konkrétní manifestace (tj. na úrovni *parole*). Zároveň může být text vykládán jako pojem statický, tj. výtvor, 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 monologické i dialogické (srov. Tárnyiková, 2002: 19, 20).

Podle Tárnyikové je termín diskurz častěji používán 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 Tárnyiková používá text jako zastřešující termín pro text a diskurz zároveň (srov. Tárnyiková, 2002: 21).

Vzhledem k tomu, že slovo diskurz a diskurzní analýza se začínají častěji objevovat i v současné české lingvistice (např. Poláková a kol., 2012a, Zikánová a kol., 2015a) – patrně v reakci na lingvistiku světovou – pokusíme se v následující části práce různé významy termínu diskurz popsat a vysvětlit.

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.

3.3.1 Strukturní přístup k diskurzu

Z hlediska strukturního přístupu je diskurz vnímán v zásadě jako synonymum k textu – jako manifestace jazyka přesahující rámec věty. Podstatou lingvistického zkoumání je v tomto případě organizace jazyka nad úrovní věty – 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 Sabbettai Harris (srov. *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 (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. Halliday a Hasanová, 1976, či 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.

3.3.2 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 Ralpha 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 (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 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. 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. 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: „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 diskurzním korpusu se zaměřuje na zachycení vztahů mezi jednotlivými úseky textu, tj. přiřklání se k tomu, že je diskurz dělitelný na menší jednotky, které jsou provázány vzájemnými významovými vztahy (tato skutečnost je podle Blakemoreové, 2002, či Schiffrinové, 1994, podstatou strukturního přístupu), tyto vztahy ale zároveň hledá v autentických textech, tj. produktech lidského chování a myšlení.

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

3.3.3 Diskurz jako fenomén přesahující využití jazyka v praxi

V tomto pojetí se autoři (např. Jaworski a Coupland, 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 re-

flecting social order but also language shaping social order, and shaping individuals' interaction with society." (Jaworski a Coupland, 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 Jaworského a Couplanda 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. Jaworski a Coupland 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.“ (Jaworski a Coupland, 2006: 11)

V naší práci budeme termín diskurz (diskurzni) 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 3.5) s tím, že podstatou diskurzni analýzy je zpracování významových vztahů mezi jednotkami diskurzu a zkoumání jazykových prostředků, které mají tyto vztahy schopnost vyjádřit (tj. především diskurzni konektory).

3.4 Co mohou spojovat diskurzni konektory

Jak již bylo řečeno výše, podle některých přístupů je podstatou diskurzni analýzy hledání a interpretace diskurzni 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 diskurzni konektorům s ohledem na to, mezi jakými jednotkami vyjadřují příslušný diskurzni vztah – prvním je tzv. teorie koherence (v originále *coherence theory*), druhým teorie relevance (v originále *relevance theory*).

3.4.1 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 diskurzni konektorů je pomoci čtenáři tyto vztahy lépe identifikovat (srov. Mann a Thompsonová, 1986; Fraser, 1990; Sanders, Spooren a Noordman, 1992; Hovy, 1990, či Knott a Dale, 1994).

3.4.2 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 diskurzní konektory.

Pro teorii relevance je významnou opozicí dvojice pojmový (koncepční) 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 Blassová, 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 (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 *říci jinými slovy* (v originále *to put in other words*) a přispívat tak k celkovému pojmovému obsahu dané výpovědi (srov. Blakemoreová, 2002).

3.4.3 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 diskurzní konektory spojují – 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 výpovědí 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):

- (4) (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 spojoovat 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 výrokem a situačním kontextem, tj. Peter přišel z práce dříve, než Mary očekávala. Tato situace ji zaskočila. Daným výrokem 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 diskurzním korpusu (PDiT) jsou diskurzní vztahy zpracovány spíše z hlediska 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), PDiT nepomíjí. Všechny tyto výskyty jsou v anotaci důsledně značeny. V datech PDiT se vyskytuje 277 podobných případů⁴ – srov. například nadpisy článků *A nejistota zůstává* nebo *Kdo je tedy lhář?* Počet diskurzních konektorů v PDiT, které stojí mezi dvěma úseky textu, je 21 416 (viz kapitola 6.3). Je tedy zřejmé, že korpus psaných novinových textů nepod-

⁴ Tj. výskytů diskurzních konektorů (primárních a sekundárních), které mají pouze jeden (vyjádřený) diskurzní argument – k vymezení pojmu diskurzní argument viz kapitola 3.5.1.

poruje 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 souviset také s tím, zda se jedná o text mluvený⁵ či psaný, resp. s tím, zda jsou autor i adresát sdělení přímými účastníky komunikační situace. 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.

3.5 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 (Redekeroová, 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) 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

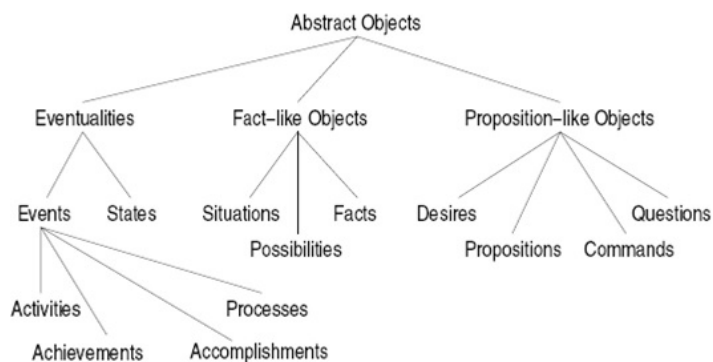
⁵ V mluvených projevech se často setkáváme např. s tím, že mluvčí má nutkání začínat svou promluvu konektorem – zřejmě má potřebu na něco navázat, tj. zasazuje svoji promluvu do situačního kontextu, srov. *Takže dobrý den / Tak začneme*.

2002: 152) Podstatou diskurzní analýzy je pak hledat mezi danými úseky významové vztahy a prostředky, které je vyjadřují.

3.5.1 Diskurzní argumenty v Pražském diskurzním korpusu

Vzhledem k tomu, že pro náš výzkum diskurzních konektorů používáme data Pražského diskurzního korpusu, přebíráme i terminologii, která je pro popis těchto dat užívána (srov. např. Poláková a kol., 2012a). V návaznosti na Penn Discourse Treebank je zde pro jednotky diskurzu používán termín diskurzní argumenty.

Diskurzní argumenty jsou definovány jako abstraktní objekty (srov. Asher, 1993), které mohou být hierarchizovány (srov. obrázek 3.1).



Obrázek 3.1: Hierarchie abstraktních objektů podle Ashera (1993: 57)

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

Diskurzní argumenty jsou dále vymezovány 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 diskurzním korpusu (na rozdíl od korpusu Penn Discourse Treebank) specificky vyznačeny.

3.5.2 Povrchová realizace diskurzních argumentů - verbální vs. nominální argumenty

V první verzi Pražského diskurzního korpusu (PDiT 1.0, srov. Poláková a kol., 2012b) byly nejprve zachyceny pouze takové diskurzní vztahy, které byly vyjádřeny konekto-

ry v užším smyslu, tedy hlavně spojkami, příslovci a částicemi, spojujícími dva verbální argumenty, které obsahují sloveso v určitém tvaru, a to vnitrovětně (viz příklad 5 s konektorem *a tak*) i mezivětně (viz příklad 6 s konektorem *proto*).

- (5) *Největším nedostatkem slovenského celku byla nedůraznost, a tak Švédům stačil k vítězství jen poloviční výkon.*
- (6) *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 následném výzkumu víceslovných konektivních struktur (především typu *důsledkem je, důvodem je, podmínkou je* apod.) jsme se ovšem neomezovali na určitou 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 7:

- (7) *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 7 vyjadřuje konektivní struktura *důsledkem je* diskurzní vztah 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 textech chovají jinak než „klasické“ konektory typu *proto, ale* apod., a tudíž na ně nemůžeme důsledně aplikovat stejné anotační zásady.

Z tohoto důvodu jsme se rozhodli neomezovat v naší práci diskurzní argumenty jejich syntaktickou povahou, ale vymezovat je čistě sémanticky, tj. diskurzní 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 posloupnostmi několika vět i celým odstavcem) nebo právě nominálními frázemi. Z hlediska syntaxe tak mají diskurzní 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).

Dosud jsme se zabývali diskurzními argumenty realizovanými textově. V některých případech mohou být ovšem 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 3.4.3. Vzhledem k povaze dat, na kterých výzkum diskurzních konektorů provádíme (tj. vzhledem k tomu, že PDiT obsahuje psané novinové texty), necháváme nicméně situační argumenty, typické zejména pro mluvené texty, dále stranou.

3.6 Diskurzní vztahy

Mezi diskurzními argumenty můžeme v textu sledovat různé typy diskurzních vztahů. V obecné rovině můžeme tyto vztahy dělit na implicitní a explicitní (viz kapitola 3.6.1). Stěžejní otázkou při analýze diskurzních vztahů je ovšem jejich významová klasifikace (viz kapitola 3.6.2), tj. to, jaké sémantické typy (či jejich podtypy) můžeme mezi diskurzními argumenty zachycovat.

3.6.1 Diskurzní vztahy implicitní a explicitní

Jak jsme popsali výše, jednotlivé úseky textu (argumenty) jsou vzájemně provázány diskurzními významovými vztahy. Tyto vztahy mohou být v textu vyjádřeny explicitně (tj. pomocí konkrétních jazykových prostředků) nebo implicitně (tj. příslušný diskurzní vztah lze vyvodit pouze z obsahů obou diskurzních argumentů), srov. příklady 8 a 9.

- (8) *Žehličkám s funkcí anticalc neublíží voda z vodovodu. **Proto** jim není třeba kupovat destilovanou.*
- (9) *Žehličkám s funkcí anticalc neublíží voda z vodovodu. **Není jim třeba kupovat destilovanou.***

V příkladech 8 a 9 je diskurzní vztah důsledku mezi argumenty *Žehličkám s funkcí anticalc neublíží voda z vodovodu* a *Není jim třeba kupovat destilovanou*. V příkladu 8 je tento vztah explicitně vyjádřen konektorem *proto*, v příkladu 9 pouze vyplývá z obsahů příslušných argumentů (tj. v textu není žádný jazykový prostředek, který by čtenáři pomohl vztah důsledku interpretovat). Jak dokazují některé studie (např. Irwinová a Pulverová, 1984), čtenář/posluchač rozumí lépe textům, ve kterých jsou diskurzní vztahy vyjádřeny explicitně. Diskurzní konektory tedy působí jako jistá vodítka pro čtenáře, kteří díky nim dokáží textu snáze a rychleji porozumět. Přítomnost diskurzních konektorů může zároveň zabránit víceznačnosti (srov. Dolková, 2010).

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

U konkrétních textů pak můžeme mluvit o různém stupni explicitnosti. Podle Milana 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 diskurzních vztahů implicitně či explicitně natolik výrazný, že je možné hovořit dokonce o implicitním a explicitním stylu.

Někteří autoři ovšem existenci čistě implicitních vztahů zpochybňují – Maite Taiboová (2009) vyslovuje hypotézu, že neexistují diskurzní vztahy, které by nebyly

vyjádřeny žádnými explicitními jazykovými signály. Kromě diskurzních konektorů považuje za tyto signály, které se podílejí (v konkrétních případech) na výstavbě textu, např. také slovesný čas, způsob, užití určitého, resp. neurčitého slovesného tvaru, užití vztahné 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 například i mluvnickou kategorií. Hlavní myšlenkou Taboadové tedy je, že diskurzní signály je zapotřebí hledat na různých jazykových rovinách.

3.6.1.1 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), Potsdam Commentary Corpus (Stede a Neumann, 2014), korpus DisFrEn (Cribble a kol., 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 3.8.3). 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 na rozpoznání diskurzních argumentů (85,1 %) i na dosazení explicitního konektoru (72 %), 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 konek-

tivního prostředku pomáhá příslušný typ vztahu snáze interpretovat a může zabránit víceznačností), srov. příklady 10, 11 a 12 z *Příruční mluvnice češtiny* (1995: 749).

- (10) *Bude pršet, zůstaneme doma.*
- (11) *Bude pršet, **proto** / z tohoto důvodu zůstaneme doma.*
- (12) *Jestli bude pršet, **tak** zůstaneme doma.*

Příklad 10 je ilustrací implicitního diskurzního vztahu. Je ovšem otázkou, o jaký významový typ vztahu se jedná (viz např. vztah důsledku nebo vztah podmínky). Tuto ambiguitu diskurzního vztahu může vyřešit právě přítomnost explicitního diskurzního konektoru (v příkladu 11 *proto* či z *tohoto důvodu*, v příkladu 12 *jestli_tak*).

3.6.2 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 blíže představíme v této kapitole.

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 diskurzní 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 diskurzních 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 (např. Prasadová a kol., 2008) vymezují diskurzní 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 elabore, vysvětlení, příklad, účel, pokračování, důsledek, zobecnění a kontrast). Podobně

Hobbs počet typů diskurzních 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 diskurzních 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 nejrozumnějších 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* (*však*) 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 klasifikací diskurzních významových 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, adverbativní (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 Hoffmannové je v textu výrazná především síť vztahů časových, prostorové vztahy jsou zastoupeny méně často. Všechny vyjmenované vztahy se uplatňují především mezi jednotlivými větami. Hoffmannová 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.

3.6.2.1 Významové vztahy v textu v Pražském diskurzním korpusu

Vzhledem k tomu, že v této práci vycházíme z anotovaných dat Pražského diskurzního korpusu, přebíráme klasifikaci jednotlivých diskurzních vztahů z anotačního manuálu pro tento korpus (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 diskurzním korpusu (PDiT) chápány významy diskurzních vztahů hierarchicky. V klasifikaci diskurzních vztahů pro PDiT 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 PDiT 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 PDiT je např. navíc vztah explikace či gradace. Pro rozlišení konkrétních typů diskurzních vztahů je spíše využita klasifikace vztahů představená v *Mluvnici češtiny 3* (1987) (i když PDiT nezahrnuje všechny vztahy, které jsou v mluvnici uvedené).

V následující části práce představíme všechny typy významových vztahů v textu, které jsou v Pražském diskurzním 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 (Poláková a kol., 2012a) a článku shrnujícího anotaci v PDiT 1.0 (Zikánová a kol., 2015b). 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 13.

- (13) *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 14.

- (14) *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 diskurzním korpusu chápán poměrně široce (nerozlišuje například mezi příčinou a důvodem). Příčina a důsledek jsou zde pojaty jako jeden typ vztahu (liší se pouze směrem, kterým je vztah veden). V tomto ohledu se klasifikace diskurzních vztahů Pražského diskurzního 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 diskurzním 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 15:

- (15) *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 16.

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

Příklad 16 můžeme interpretovat následujícím způsobem – na základě toho, že se svítí v kuchyni, usuzují, že je Petr doma. Světlo v kuchyni tedy není vlastní příčinou toho, že je Petr doma.

e) Podmínka a výsledek podmínky (COND) – srov. příklad 17:

(17) *Pokud nezískáme dlužné peníze zpět, hrozí nám bankrot.*

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 18.

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

g) Účel (PURP)

(19) *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. Prasadová a kol., 2008). Důvodem může být to, že například angličtina (na rozdíl od češtiny) nemá pro vztah explikace typický konektor – srov. klasické české výrazy pro vyjádření explikace *totiž*, *vždyť*, *přece* 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 20), přičemž obsah prvního argumentu je obecnější, druhého konkrétnější. Významným prvkem explikace bývá také subjektivnost (prvek obhajování – obhajujeme jisté tvrzení uvedením detailu, který naše tvrzení vysvětluje).

(20) *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.

i) 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 21.

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

j) Opozice (OPP) – vztah opozice vyjadřuje, že obsahy daných diskurzních argumentů si odporují, srov. příklad 22.

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

k) 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 23 (ve kterém se jedná o vztah mezi dvěma implikacemi).

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

l) Restriktivní opozice (RESTR) – restriktivní opozice vyjadřuje vztah výjimky, tj. platnost prvního argumentu je omezena obsahem druhého argumentu, srov. příklad 24.

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

m) Přípustka (CONC) – ve vztahu přípustky implikuje obsah prvního argumentu určité očekávání spojené s následujícím argumentem (srov. příklad 25 – 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).

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

n) Rektifikace (CORR) – u vztahu rektifikace druhý argument opravuje či modifikuje obsah prvního argumentu, srov. příklad 26.

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

o) Gradace (GRAD) – u vztahu gradace dochází ke stupňování obsahu prvního argumentu, obsah druhého argumentu převyšuje očekávání spojené s obsahem prvního argumentu, srov. příklad 27.

- (27) *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 (např. generalizace) se vyskytují spíše mezi většími úseky textu (např. celými odstavci).

p) Konjunkce (CONJ) – vztah konjunkce vyjadřuje prosté navazování na předchozí kontext, srov. příklad 28.

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

q) Exemplifikace (EXEMPL) – srov. příklad 29:

- (29) *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.*

r) 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 30.

- (30) *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 „vyvoléknout“.*

s) Ekvivalence (EQUIV) – u vztahu ekvivalence dochází k opakování téhož obsahu u obou diskurzních argumentů (druhý argument říká jinými slovy totéž co první), srov. příklad 31.

- (31) *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.*

t) 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 32.

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

u) Konjunktivní alternativa (CONJALT) – konjunktivní alternativa označuje vztah dvou diskurzních argumentů vyjadřujících alternativy, které je možné chápat ve slučovacím i vylučovacím vztahu, srov. příklad 33.

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

v) Disjunktivní alternativa (DISJALT) – disjunktivní alternativa označuje vylučovací vztah mezi dvěma diskurzními argumenty, srov. příklad 34.

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

Jak již bylo zmíněno výše, klasifikace významových vztahů v textu byla v rámci jejich zpracování v první verzi Pražského diskurzního korpusu sestavena pro konektory v užším smyslu (tj. především synsémantika typu *ale, i, proto* atd.). Vzhledem k tomu, že náš výzkum proběhl na datech PDiT, 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.

3.7 Diskurzní konektory

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

3.7.1 Diskurzní konektory ve světové literatuře

Pojetí diskurzních 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 (diskurzní)

konektory (z anglického *discourse connectives*, Blakemoreová, 1987, 1992, 2002; Rouchota, 1996; Celleová a Huartová, 2007), diskurzní částice (*discourse particles*, Schourup, 1985; Fischerová, 2006), diskurzní ukazatele (*discourse markers*, Schiffrinová, 1987; Urgellesová-Collová, 2010), diskurzní 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.

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

Slovnědruhová příslušnost

Diskurzní konektory bývají nejčastěji chápány jako výrazy patřící mezi podřadící a souřadící spojky, předložkové fráze a adverbia (srov. Fraser, 1999; Prasadová a kol., 2008, 2010 aj.).⁶ Někteří autoři (Schiffrinová, 1987; Hansen, 1998; Aijmerová, 2002) mezi konektory řadí i částice či jmenné fráze. Podle Frase (1999) jsou diskurzními 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 – např. podle Urgellesové-Collové (2010) mohou být diskurzními 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 diskurzních konektorů, a to vzhledem k překrývání hranic mezi některými slovními druhy (srov. např. obtížně definovatelné hranice mezi částicemi a spojkami – viz kapitola 6.1.2.8).

Pojmový vs. procedurální význam

Někteří autoři vymezují diskurzní konektory na základě toho, zda nesou pojmový či procedurální význam (srov. kapitola 3.4.2), 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

⁶ 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.

(1975, 1989; rozdíl mezi říkáním a implikováním).⁷ Zatímco někteří autoři (srov. Rouchota, 1996; Fraser, 1999) vymezují diskurzní konektory jako výrazy s významem procedurálním, jiní (např. Urgellesová-Collová, 2010) argumentují, že některé konektory obsahují i výrazy s významem pojmovým.

Jednou z nejvýznamnějších autorek zabývajících se konektory z hlediska pojmového a procedurálního významu je D. Blakemoreová, která se dané problematice věnuje v několika svých dílech, a její názory se v tomto ohledu významně vyvíjejí. Zatímco ve svých dřívějších pracích (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):

- (35) *Myslím, že to, co potřebujeme, je místnost se stolem, to znamená stolem, okolo kterého by studenti mohli sedět.* (Blakemoreová, 2007: 328)

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*.

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):

- (36) *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, např. Fraser (1990, 1996) či Rieber (1997). Podle Urgellesové-Collové (2010) se na pravdivostních podmínkách nepodílí většina konektorů, některé ovšem ano (např. *therefore* – *proto*).

⁷ 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.

Další autoři nepovažují přispívání k pravdivostním podmínkám za zásadní kritérium pro vymezování 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. Schiffrinová (1987) či Knott a Dale (1994).

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. Zwicky, 1985; Schiffrinová, 1987; Fischerová, 2006 aj.).

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

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

Schiffrinová (1987) definuje diskurzní ukazatele (používá pro ně výraz *discourse markers*) 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.

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.

Urgellesová-Collová (2010) 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.

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 ostatní jazykové prostředky s diskurzní spojovací funkcí (typu *příkladem je*) používá termín alternativní vyjádření diskurzních konektorů.

Většina autorů se shoduje v základní charakteristice typických diskurzních konektorů (tj. výrazů jako *a*, *ale*, *proto*, *však*). Liší se ovšem v tom, kde je možné stanovit hranici konektorů – někteří autoři vymezují konektory v užším smyslu, například

Shloushová (1998), Hakulinen (1998) a Maschlerová (2000) omezují konektory pouze na synsémantická slova, jiní v širším smyslu, tj. jako konektory definují například i slovesné nebo jmenné fráze – srov. Schiffrinová (1987), Hansen (1998) a 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, například Fischero-ová (2006) analyzuje „the most typical members“ – nejtýpčtější 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* či *a*) 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 vymezování 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 6.1).

3.7.2 Diskurzní konektory v české literatuře

V českém prostředí se diskurzními konektory (a výrazy podílejícími se na struktura- ci textu obecně) zabývali především Karel Hausenblas (1962), Karel Svoboda (1968, 1972), 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 češ- tině přináší ve svých studiích Jana Hoffmannová (srov. např. *Sémantické a pragmatické aspekty koherence textu*, 1983; *Typen der Konnektoren und deren Anteil an der Organisation 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 například na slovnědruhovou příslušnost daných výrazů, ale vymezuje je především z funkčního, sémanticko-pragmatického hlediska. Upozorňuje totiž na skutečnost, že konektory nesou jak prvek sémantický (jedná se o vztahovou sémanti- ku v rámci textu), tak pragmatický (tj. zejména postoj autora k textu, autora k adresá- tovi).

Ve svých studiích Hoffmannová představuje podrobnou 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 – například *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 například 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. *tehdy, 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 nebo shrnutí.

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*), otevřená skupina sekundárních předložek (často ve spojení s ukazovacím zájmenem – např. *vedle toho, v tomto okruhu, v této souvislosti, v tom případě*) či ustálené fráze a idiomy (*lépe řečeno, přesněji řečeno*) (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.

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 výpověďmi (pro výrazy signalizující vztahy uvnitř souvětí souřadného 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 (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 vymezování 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 nebo textové.

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 slovnědruhové 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 ani k jedné skupině 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 obecněmu vyme-

zení 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 diskurzní 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. *musíme se zamyslet*, *chceme zdůraznit*, *musíme podtrhnout*.

3.7.3 Základní vymezení diskurzních konektorů pro tuto práci

Z analýzy vybraných předních studií o diskurzních 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 větnečlenské platnosti – srov. zejm. Hrbáček, 1994), jiní pojetí širší (vymezující konektory především ze sémanticko-pragmatického hlediska – srov. zejm. Hoffmanová, 1983, 1984).

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

Tato práce je výsledkem dlouhodobého výzkumu diskurzních konektorů v češtině. Navazujeme v ní na své předchozí dílčí studie, resp. shrnujeme zde hlavní myšlenky předešlého výzkumu konektorů, které dále rozšiřujeme. Především navazujeme na vymezení konektorů z roku 2014 a 2015 (srov. Rysová a Rysová, 2014, 2015), ve kterém jsme zavedli dvě nové kategorie konektorů – konektory primární (základní spojovací prostředky typu *a*, *ale*, *proto*, *nebo*, *protože*, *totiž*) a konektory sekundární (především víceslovné struktury typu *z tohoto důvodu*, *to znamená*, *příčinou bylo*, *za této podmínky*). Jako opozici k diskurzním konektorům (primárním i sekundárním) jsme ve zmíněných studiích zavedli tzv. volné konektivní fráze (mezi ně řadíme silně kontextově závislé struktury typu *díky tomuto počasí*, *vzhledem k tomuto navýšení*). V této knize na-

še původní vymezení konektorů zpřesňujeme, rozšiřujeme a obohacujeme o výsledky nového výzkumu.

Ve výsledcích této práce (srov. kapitola 6) tedy postupně představíme svoji definici diskurzních konektorů, obecnou charakteristiku konektorů primárních a sekundárních a posléze se zaměříme na zpracování a detailní analýzu sekundárních konektorů v Pražském diskurzním korpusu (PDiT).

3.8 Zpracování diskurzních vztahů a konektorů ve velkých korpusech

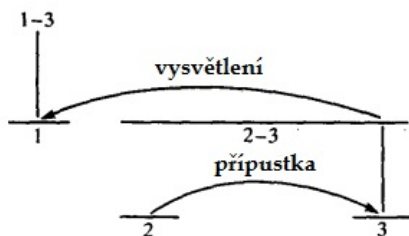
Pražský diskurzní korpus je zatím jediným korpusem češtiny, který obsahuje anotaci diskurzních vztahů (v současné době pouze vyjádřených explicitně, tj. primárními a sekundárními konektory). Pro ostatní jazyky ovšem podobných korpusů existuje více. V následující kapitole představíme hlavní z nich. Zaměříme se přitom na různé způsoby, jakými je možné diskurzní vztahy ve velkých korpusech zpracovávat. Jednotlivé přístupy 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, tedy konektory – srov. teorii RST), či zda jejich zpracování vztahů v textu začíná „zdola“, tj. vychází právě od identifikace a anotace diskurzních konektorů jako signalizátorů textových vztahů (srov. Penn Discourse Treebank). Mezi další rozdíly patří například to, zda jsou diskurzní vztahy anotovány na lineárních textech či ve stromových strukturách, jaké typy vztahů jsou značeny (tj. implicitní, explicitní apod.) či jak jsou definovány diskurzní konektory.

3.8.1 RST Discourse Treebank

Jedním z prvních velkých korpusů, který obsahuje anotaci diskurzních vztahů, je RST Discourse Treebank (Carlson a kol., 2003). Tento korpus obsahuje 385 textů (anglických článků z časopisu Wall Street Journal). Diskurzní vztahy jsou zde anotovány podle teorie Rhetorical Structure Theory (RST, srov. Mann a Thompsonová, 1988) chápající text jako síť vztahů mezi 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 diskurzního korpusu) tedy spočívá ve vytváření spojitě struktury zachycující 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 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 3.2. 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 3.2: Ukázka anotace textů podle teorie RST (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).

3.8.2 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 diskurzních 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 37.

- (37) *Společná práce u Stavební geologie byla ze začátku zajímavá, ale pak jsem začal pociťovat jistou rutinu.*

V daném příkladu jsou dva diskurzní vztahy zároveň – 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 diskurzních vztahů v korpusu Discourse Graphbank naopak překrývá ní určitých typů významových vztahů dovoluje (podobně jako PDiT). Textové vztahy

v Discourse Graphbanku nejsou zachycovány ve stromových strukturách, ale ve „volnějších“ grafech (srov. Wolf a Gibson, 2005).

3.8.3 Penn Discourse Treebank a navazující projekty

Dalším významným projektem (představujícím odlišné pojetí zpracování diskurzních 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 korpus Penn Treebank. Obsahuje přibližně 49 000 vět, texty pocházejí z časopisu Wall Street Journal. V korpusu Penn Discourse Treebank nejsou diskurzní vztahy zachyceny ve stromových strukturách či grafech, ale na lineárních textech. Zpracovány jsou přitom diskurzní vztahy vyjádřené explicitními konektory, alternativními lexikálními vyjádřeními diskurzních konektorů (zkráceně altlexy), diskurzní vztahy vyjádřené implicitně, dále tzv. nulové vztahy (tj. případy, ve kterých není mezi dvěma sousedními jednotkami významový vztah, tedy například místa, kde začíná nový odstavec) a vztahy mezi jmennými entitami. Diskurzní konektory jsou zde chápány jako výrazy otevírající pozice pro dva diskurzní argumenty (diskurzní 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 diskurzních vztahů v různých jazycích – srov. diskurzní 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 (Tonelliová a kol., 2010) či na anglických biomedicínských textech (Prasadová a kol., 2011) a navazuje na něj také zpracování diskurzních vztahů v Pražském diskurzním korpusu.

3.8.4 Další významné diskurzní korpusy

V posledních letech vznikají další diskurzní projekty, které se snaží zpracovat diskurzní vztahy v textech různých žánrů a stylů, často ve spojení s anotací dalších diskurzních jevů (jako je například 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 a kol., 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. Existuje také korpus LSFB (French Belgian Sign Language; Meurantová a Sinteová, 2013), francouzsko-belgický korpus znakového jazyka, který se zaměřu-

je (mimo jiné) na 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.

3.8.5 Zpracování víceslovných konektivních struktur ve velkých korpusech

Všechny zmíněné korpusy obsahují zpracování diskurzních vztahů – jak jsme již ale naznačili, jednotlivá pojetí se vzájemně 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 (tj. struktur typu *za této podmínky, z těchto důvodů*), 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 diskurzní konektory zahrnuje v daných pojetích také konektory sekundární či nikoli. Patrně nejbližším termínem k sekundárním konektorům je pojem alternativní lexikální vyjádření diskurzních konektorů (zkráceně altlexy) používaný v pensylvánském korpusu Penn Discourse Treebank (PDTB). Vzhledem k tomu, že je tento termín ve světové literatuře poměrně rozšířený (projekt PDTB má mnoho pokračovatelů), pojem diskurzní konektory bude v mnohých případech spíše zahrnovat pouze konektory primární (tj. především jednoslovné výrazy typu *protože, totiž, ale*).

Zatímco tedy v naší práci používáme termíny primární a sekundární konektory, u většiny zmíněných projektů by odpovídaly nejspíše termínům diskurzní konektory a jejich alternativní lexikální vyjádření, tj. altlexy. Vymezení konektorů a altlexů podle korpusu PDTB je ale v některých ohledech jiné než naše vymezení primárních a sekundárních konektorů (viz dále kapitola 6.3.1.3). Termín alternativní lexikální vyjádření diskurzních konektorů tedy není totožný s termínem sekundární konektory. V projektu PDTB jsou altlexy pojímány jako výrazy s konektivní diskurzní funkcí, které nejsou na seznamu diskurzních konektorů sestaveném pro anotaci diskurzních vztahů v PDTB (srov. Prasadová a kol., 2010: 1025–1026). Vymezení diskurzních 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 diskurznímu 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 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 kontextově závislé a nezávislé 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 tedy v PDTB označen jakýkoli výraz signalizující příslušný diskurzní 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 přibližně 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. Označení alternativní lexikální vyjádření konektorů není zcela přesné ani terminologicky – existují totiž případy (např. *v tomto ohledu, na závěr* – viz kapitola 6.3.2.2), které mezi (primárními) konektory nemají přímý ekvivalent. Ne vždy je tedy sekundární konektor alternativou k primárnímu.

Anotace alternativních lexikálních vyjádření konektorů není v PDTB plně systematická a konzistentní (anotovány byly například pouze případy, ve kterých altlexy vyjadřovaly vztah mezi dvěma sousedními argumenty). 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,⁸ než ukazují současná anotace (ke srovnání výsledků anotace altlexů v angličtině v PDTB a sekundárních konektorů v češtině v PDiT viz kapitola 6.3.1.3).

Částečná anotace altlexů se objevuje i v některých diskurzních projektech navazujících na Penn Discourse Treebank – např. v projektu Hindi Discourse Relation

⁸ 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)

Bank pro hindštinu. Altlexy zde dosáhly 6,14 % z celkového počtu diskurzních vztahů (tj. v rámci všech typů vztahů anotovaných v daném korpusu, vyjádřených 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 převzato 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* (která se v textu mohou objevit v množném čísle a v různých dalších obměnách) do anotace diskurzních vztahů nezahrnují. Je tedy zřejmé, že přední zahraniční projekty zaměřené na zpracování diskurzních vztahů ve velkých korpusech systematickou anotaci sekundárních konektorů neobsahují. Nejvíce prostoru věnuje těmto výrazům projekt Penn Discourse Treebank, ovšem ani zde není jejich anotace zcela konzistentní a nevychází z hlubšího teoretického základu.

4

Jazykový materiál

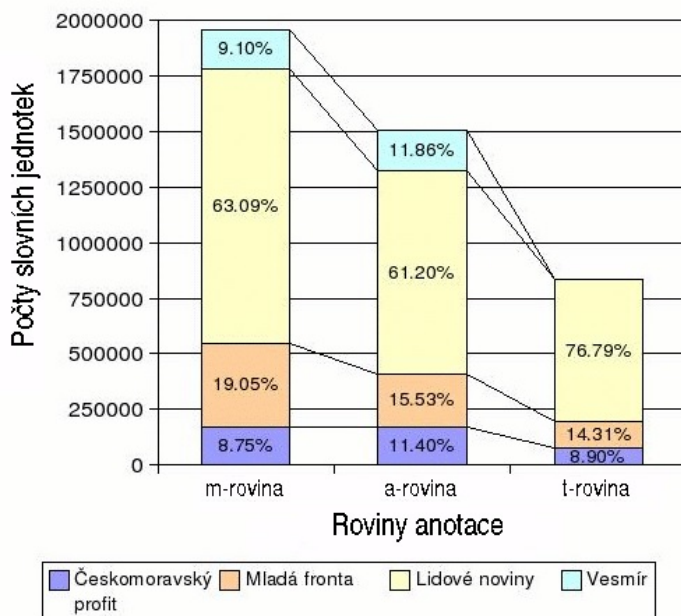
Výzkum diskurzních konektorů jsme prováděli na datech Pražského diskurzního korpusu (který vznikl v návaznosti na Pražský závislostní korpus), okrajově na datech Českého národního korpusu (SYN2013PUB). V další části práce jsme pak využili i data korpusu DIALOG (pro srovnání sekundárních konektorů v psané 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). PDT zpracovává věty na několika jazykových úrovních zároveň – jedná se o rovinu morfologickou, analytickou (tj. povrchovou syntax) a tektogramatickou (tj. rovinu hloubkovou, syntakticko-sémantickou). PDT obsahuje na každé ze tří anotovaných rovin různé množství textů – srov. obrázek 4.1 zobrazující rozvržení počtu textů (resp. počtu slovních jednotek) z příslušných novin a časopisů na jednotlivých rovinách. Obrázek je převzatý z dokumentace k PDT 2.0 (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 4.1 zobrazující konkrétní počet dokumentů, vět a slovních jednotek v jednotlivých množinách tektogramatické roviny Pražského závislostního korpusu (převzata ze zadávací dokumentace k PDT 2.0 – Hajič a kol., 2006).

Anotace jazykových dat Pražského závislostního korpusu je neustále rozšiřována a obohacována, průběžně tedy vycházejí nové, aktualizované verze korpusu (srov. nejnovější verzi PDT 3.5, Hajič a kol., 2018). Na datech PDT byla provedena také anotace nadvětných jevů, tj. koreference a asociační anafory a diskurzních vztahů, vydaná pod názvem Pražský diskurzní korpus (PDiT).



Obrázek 4.1: Rozvržení počtu slovních jednotek na jednotlivých anotačních rovinách v Pražském závislostním korpusu (viz Hajič a kol., 2006)

tamw	train	dtest	etest	celkem
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 4.1: Množství dat v Pražském závislostním korpusu (viz Hajič a kol., 2006)

4.2 Pražský diskurzní korpus

4.2.1 Pražský diskurzní korpus 1.0

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 vydána jako Pražský diskurzní korpus 1.0 (Poláková a kol., 2012b). Anotace byla omezena na diskurzní vztahy (vnitrovětné i mezivětné) vyjádřené explicitními konektory, vymezenými jako spojovací výrazy patřící do následujících tříd: spojky souřadící a podřadící, částicové výrazy, příslovce, některá užití zájmen (srov. *místo toho*), ustálené víceslovné prostředky (srov. *stručně řečeno, jinými slovy*), písmenné nebo číselné prvky pro vyjádření vztahu výčtu (srov. 1., 2.), interpunkční znaménka – dvojtečka a pomlčka (více viz Poláková a kol., 2012a).

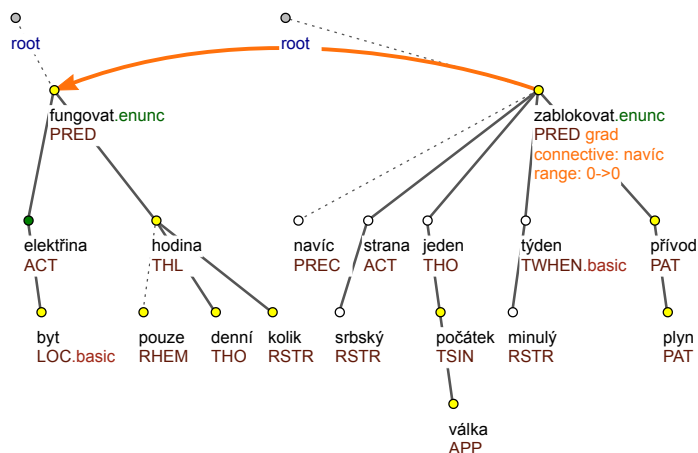
Diskurzní vztahy byly anotovány pouze mezi argumenty obsahujícími sloveso v určitém tvaru (označovanými jako slovesné argumenty) – srov. příklad 38 z PDiT s konektorem *a tak* signalizujícím vnitrovětný vztah důsledku a příklad 39 z PDiT s konektorem *navíc* vyjadřujícím mezivětný vztah gradace.

- (38) *Tyto potraviny a bavlna se v České republice nepěstují, a tak jejich dovoz naše zemědělcé neohrozí.*
- (39) *Elektrina v bytech funguje pouze několik hodin denně. Poprvé od počátku války minulý týden srbská strana navíc zablokovala přívod plynu.*

Příklad 39 je zároveň zachycený na obrázku 4.2. Obrázek ukazuje dané argumenty ve stromové struktuře PDiT. 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 – konektorem *navíc*. Diskurzní vztahy jsou v PDiT 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 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 diskurzní šipka, a větou, ke které šipka směřuje). Diskurzní vztahy ovšem mohou být i mezi většími úseky textu, např. mezi větou a odstavcem nebo mezi dvěma odstavci.⁹

⁹ Například značka „range: 2→6“ by znamenala, že do argumentu, od kterého vede diskurzní vztah, patří celkem tři věty (věta, kde diskurzní šipka začíná, plus dvě věty po ní následující), a do argumentu, ke kterému diskurzní vztah vede, patří sedm vět (věta, kde diskurzní šipka končí, plus šest vět po ní následujících).



Obrázek 4.2: Ukázka anotace konektorů v Pražském diskurzním korpusu 1.0, viz příklad 39: *Elektřina v bytech funguje pouze několik hodin denně. Poprvé od počátku války minulý týden srbská strana navíc zablokovala přívod plynu.*

Diskurzni argumenty jsou v PDiT definovány podle Asherovy koncepce abstraktních objektů (1993) a principu minimality (Prasadová a kol., 2007), tj. do diskurzniho argumentu patří pouze výpovědi (nebo jejich části), které nesou informace nezbytně nutné pro interpretaci příslušného diskurzniho vztahu (viz kapitola 3.5.1). V textech mohou být diskurzni argumenty realizovány například větou, souvětím, samostatnou vedlejší větou, posloupností několika vět i celým odstavcem.

Diskurzni vztahy jsou v PDiT značeny mezi argumenty, které jsou v textu bezprostředně vedle sebe (viz příklad 39), 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.

V průběhu anotace neměli anotátoři k dispozici předem sestavený seznam explicitních konektorů, podle kterého by určovali, který výraz mezi konektory patří a který ne (na základě seznamu konektorů byla anotace diskurzni 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 PDiT byli na začátku seznámeni s obecnou definicí konektorů a sami je posléze v textech hledali a rozpoznávali.

První anotace diskurzni vztahů v PDiT, jak již bylo zmíněno výše, zahrnuje zpracování mezivětných i vnitrovětných vztahů. Mezivětné vztahy byly anotovány manuálně, vnitrovětné částečně automaticky (resp. automatická anotace vnitrovět-

ných vztahů byla následně zkontrolována a manuálně upravena anotátory, viz Jínová a kol., 2012).

4.2.2 Pražský diskurzní korpus 2.0

Cílem této práce je výzkum a popis konektivních struktur typu z *tohoto důvodu*, v naší terminologii tzv. sekundárních konektorů. Vzhledem k tomu, že žádný český korpus dosud systematické zpracování těchto výrazů neobsahoval, museli jsme příslušnou anotaci nejprve sami provést (tedy stanovit anotační principy a poté je aplikovat na konkrétní data), abychom získali jazykový materiál pro vlastní výzkum. Tato nová diskurzní anotace byla provedena jako pokračování Pražského diskurzního korpusu 1.0 a byla již publikována – pod názvem Pražský diskurzní korpus 2.0 (Rysová a kol., 2016).

PDiT 2.0 tedy reflektuje naše dělení konektorů na primární a sekundární. Obě skupiny konektorů tak bylo možné po dokončení anotace porovnat, resp. srovnat primární a sekundární konektory z hlediska jejich výskytu a chování v psaných textech.

PDiT 2.0 obsahuje revizi prvotní anotace z PDiT 1.0 (některé typy explicitních konektorů uvedené v kapitole 4.2.1, jako např. předložkové fráze či ustálené víceslovné prostředky, byly přeznačeny na konektory sekundární) a především obsahuje systematické zpracování ostatních (dosud nijak neanotovaných) sekundárních konektorů.

Výsledky našeho výzkumu tedy vycházejí z jazykových dat korpusu PDiT 2.0, obohaceného o dosud nepublikované zpracování diskurzních vztahů mezi nominálními diskurzními argumenty a tzv. volnými konektivními frázemi, tj. neustálenými, kontextově závislými prostředky typu *kvůli tomuto počasí*, *vzhledem k tomuto zvýšení*, více viz kapitola 6. Průběh anotace sekundárních konektorů (včetně její evaluace) je podrobně popsán v kapitole 6.2.

4.3 Korpus DIALOG

Korpus DIALOG je multimediální 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 diskurzním korpusem 2.0 (srov. 932 373 textových slov v DIALOGu vs. 833 195 v PDiT). Značná část pořadů, ze kterých nahrávky pocházejí, je zároveň z oblasti publicistiky (srov. 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 lematizovaný (anotace proběhla automaticky). Anotace diskurzních vztahů není v korpusu obsažena.

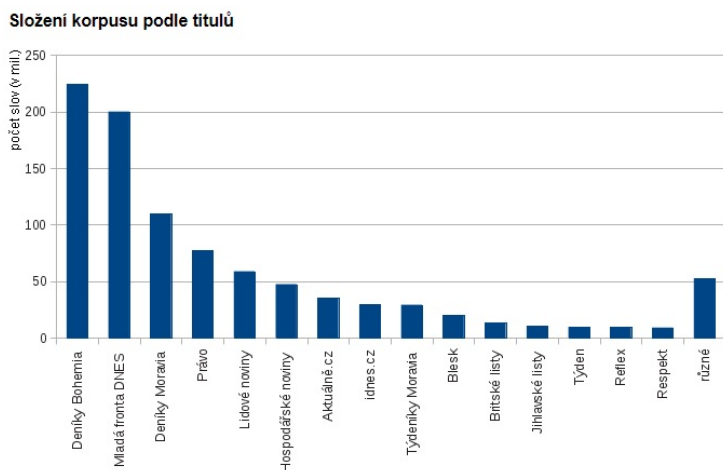
Data z korpusu DIALOG 1.1 jsme použili pro jednu dílčí studii, a to pro porovnání vybraných sekundárních konektorů v psané a mluvené češtině (viz kapitola 6.3.1.4).

4.4 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 jednotlivá slova i spojení slov v kontextu.

V rámci projektu ČNK byla vytvořena řada korpusů, jedná se především o korpusy psané češtiny (např. SYN2010, 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 (Křen a kol., 2013) 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 4.3.¹⁰



Obrázek 4.3: Složení korpusu SYN2013PUB podle titulů

¹⁰ Převzato z <http://ucnk.korpus.cz/syn2013pub.php>.

5

Metody a postup práce

Nejprve jsme provedli anotaci zkušebního vzorku dat PDiT (zaměřenou na primární a sekundární konektory). Poté jsme vymezili diskurzní konektory jako skupinu a blíže specifikovali konektory primární a sekundární, resp. formulovali jsme rozdíly mezi nimi. Výsledky tohoto prvního kroku, tj. bližší vymezení diskurzních konektorů a specifikace konektorů primárních a sekundárních, jsou popsány v kapitole 6.1.

Následně jsme vytvořili anotační principy pro zpracování sekundárních konektorů v PDiT, protože anotaci těchto jazykových prostředků dosud žádný korpus češtiny neobsahoval. Při vytváření anotačních principů jsme vycházeli z prvotní anotace konektorů v PDiT 1.0 (viz kapitola 4.2.1). Tyto principy ovšem nebylo možné pro sekundární konektory aplikovat zcela. Již během prvotní, zkušební anotace se ukázaly obecné jazykové odlišnosti mezi konektory sekundárními a primárními. Tyto odlišnosti bylo tedy nutné při vytváření anotačních principů pro sekundární konektory zohlednit. Při formulaci základních principů jsme řešili různé teoretické problémy a otázky týkající se např. rozsahu sekundárních konektorů, tj. zda je konektor celé spojení *to je důvod, proč* nebo jenom část *to je důvod* atd. Finální podobu anotačních principů pro sekundární konektory (včetně evaluace anotace), problematičtější teoretické otázky a návrh jejich řešení rozpracováváme v kapitole 6.2.

Výsledky nově provedené anotace sekundárních konektorů včetně jednotlivých dílčích metod použitých během našeho dalšího výzkumu (např. jak jsme vyhodnocovali dokončenou anotaci nebo jak a s jakými výsledky jsme testovali kontextovou nezávislost konektorů) jsou představeny v kapitolách 6.3 a 6.4.

6

Výsledky práce

V následující části práce představíme hlavní výsledky našeho výzkumu diskurzních konektorů v češtině. Nejdříve (kapitola 6.1) se zaměříme na teoretický popis konektorů, tj. na jejich definici a dělení, dále na obecnou charakteristiku primárních konektorů, sekundárních konektorů a volných konektivních frází (na základě zkušenosti s jejich výskytem a chováním v textech). Posléze představíme (kapitola 6.2) a vyhodnotíme (kapitola 6.3) vlastní anotaci sekundárních konektorů v PDiT, především srovnáme výsledky této anotace s obdobnou anotací primárních konektorů (viz Poláková a kol., 2012b). Na závěr (kapitola 6.4) představíme experiment na ověření naší definice diskurzních konektorů, popíšeme možnosti testování a vyhodnocení míry závislosti konektorů na kontextu a posléze představíme obecné principy a tendence v užívání diskurzních konektorů.

6.1 Definice a dělení diskurzních konektorů

Jak jsme popsali v teoretické části práce (viz kapitola 3.7), pojetí diskurzních konektorů se u jednotlivých autorů liší. Autoři se většinou shodují na základní charakteristice typických konektorů (typu *a*, *ale*, *nebo*), 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. konektory (např. *proto*, *však*, *naopak*) a 2. tzv. alternativní lexikální vyjádření konektorů (zkráceně altlexy – srov. *příkladem je, z tohoto důvodu* – viz Prasadová a kol., 2010), jiní (např. Hoffmannová, 1983; Schiffrinová, 1987) definují konektory široce, tj. jako konektory označují i zmíněná „alternativní vyjádření“, někteří (srov. Hrbáček, 1994) vymezují konektory v užším smyslu (tj. „alternativní vyjádření konektorů“ mezi konektory vůbec neřadí).

V následující části 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 kapitole 3.7, nabízí se zde dělení na „typické“ a „méně typické“ konektory, tj. na centrum a periferii diskurzních konektorů).

Jak již bylo naznačeno v úvodu práce, schopnost vyjadřovat významové vztahy mezi dvěma úseky textu mají jazykové prostředky různých slovních druhů, různé lexikální i syntaktické povahy – srov. příklady uvedené již v úvodu této práce:

- (40) *Petr je nemocný a nemůže s námi jet. **Proto** se výlet ruší. / Výlet se **tedy** ruší. / Výlet se **tak** ruší. / Výlet se **tím** ruší. / Výlet se **tímto** ruší. / **Na základě toho** se výlet ruší. / **Z tohoto důvodu** se výlet ruší. / **Kvůli tomu** se výlet ruší. / **Kvůli jeho nemoci** se výlet ruší.*

V uvedených příkladech je mezi danými diskurzními argumenty (*Petr je nemocný... a výlet se ruší*) vztah důsledku. Ve všech případech je tento vztah také explicitně vyjádřen pomocí konkrétních jazykových prostředků (*proto, tedy, tak, tím, tímto, na základě toho, z tohoto důvodu, kvůli tomu, kvůli jeho nemoci*). Otázkou ovšem zůstává, zda všechny dané prostředky můžeme označit jako diskurzní konektory. V následující části práce 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 jazykových dat Pražského diskurzního korpusu 2.0.

6.1.1 Vymezení diskurzních konektorů z funkčního hlediska

V naší práci vymezujeme diskurzní konektory z funkčního hlediska, tj. jako výrazy, které i) **vyjadřují mezi dvěma úseky textu jeden ze sémanticko-pragmatických vztahů** (jako je podmínka, přípustka, specifikace, příčina a důsledek apod.) a které zároveň ii) **mohou být pro vyjádření tohoto typu vztahu použity ve velkém množství různých kontextů, tj. které jsou do jisté míry kontextově nezávislé.**

Různými kontexty přitom myslíme kontexty s příslušným typem diskurzního vztahu, který je daný konektor schopný vyjadřovat, např. tedy konektor *proto* může být použit v různých kontextech s diskurzním vztahem důsledku (*Je mi horko, sundám si proto svetr. / Do hotelu je to ještě daleko, vezmeme si proto taxi.*), nikoli ale samozřejmě v kontextech se vztahem opozice, přípustky apod.

Naše vymezení diskurzních konektorů můžeme demonstrovat na následujících kontextech z PDiT.

- (41) i) *V dnešní době je zpravidla nepoužitelné všechno, co je starší než čtrnáct dní. **Proto** / **z tohoto důvodu** je dobré vrátet se ke zdrojům, jejichž kvalitu potvrdil čas.*

ii) *Italská vláda je rozhodnuta udělat rázná opatření proti nárůstu rozpočtového deficitu. **Proto** / **Z tohoto důvodu** uvažuje o urychleném zavedení vyšší věkové hranice pro odchod do důchodu.*

iii) *Klienti, kteří si kupují určité katalogy, mnohdy neví, jak s informacemi v nich pracovat. **Proto** / **Z tohoto důvodu** by asociace měla mimo jiné učit, jak k informacím přistupovat, jaké jsou informační zdroje, jak se dělí a jak je využívat.*

Všechny tři kontexty (41) obsahují mezivětný vztah důsledku – pro všechny je zároveň možné tento vztah vyjádřit výrazy *proto* či *z tohoto důvodu*. Oba tyto jazykové prostředky tedy považujeme za diskurzní konektory.

Nyní se vrátíme k příkladům 40 uvedeným v úvodu kapitoly 6.1 (tj. *Petr je nemocný a nemůže s námi jet. Proto / Tedy / Na základě toho / Z tohoto důvodu / Kvůli tomu / Kvůli jeho nemoci se výlet ruší.* či *Výlet se tak / tím / tímto ruší.*). Podle našeho vymezení můžeme jako diskurzní konektory vyhodnotit výrazy *proto, tedy, tak, tím, tímto, na základě toho, z tohoto důvodu a kvůli tomu*.

Spojení *kvůli jeho nemoci* za konektor nepovažujeme. Toto spojení je kontextově závislé a pro vyjádření diskurzního vztahu důsledku nemůže být použito v textech různého tematického zaměření.

Je zřejmé, že naše vymezení diskurzních konektorů splňují značně různorodé spojovací prostředky (v rámci vztahu důsledku jsou to jednoslovné výrazy jako *proto, tudíž, tedy* i víceslovná spojení typu *z tohoto důvodu; to je důvod, proč* apod.). Z tohoto důvodu vyčleňujeme v rámci kategorie diskurzních konektorů dvě skupiny, konektory primární a sekundární.

6.1.2 Primární a sekundární konektory

Primární konektory vymezujeme jako jazykové výrazy se spojovací funkcí, které jsou gramatikalizované, většinou jednoslovné (srov. *a, ale, proto, když, protože, nebo, i*), zřídka víceslovné (tvořící korelativní páry typu *bud'nebo* či sousloví typu *i když*). Primární konektory jsou z lexikálního hlediska ustálené, ve větě neplní funkci větných členů, jsou neohebné a většinou nedovolují modifikaci.

Sekundární konektory jsou (dosud) ne plně gramatikalizované jazykové struktury, které v textu plní spojovací funkci. Na rozdíl od primárních konektorů jsou méně ustálené – jejich variabilita se přitom týká lexikálního obsazení (srov. *z tohoto / toho / uvedeného důvodu*), morfologických tvarů (*důvod je* vs. *důvodem je*) či slovosledu (*to je důvod, proč* vs. *je to důvod, proč*). Sekundární konektory ve větě plní funkci větných členů (*kvůli tomu*) či větných modifikátorů (*stručně řečeno*), zpravidla podléhají flexi (*za této podmínky – za těchto podmínek*) a dovolují modifikaci (*hlavní / jedinou / další podmínkou je*).

Sekundární konektory zároveň obsahují tzv. jádrová slova, tj. sémanticky „silná“ slova tvořící základ dané struktury (jádrová slova zpravidla odkazují k určitému sémantickému typu diskurzního vztahu). Jádrová slova mohou být např. podstatná jména (jako *podmínka, důvod, příčina*), předložky (jako *kvůli, díky, navzdory, na základě*) či slovesa (jako *shrnout, vysvětlit, předcházet*). Sekundární konektory jsou na rozdíl od primárních většinou víceslovné a v některých případech bývá obtížné určit jejich rozsah: například 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. Platnost konektoru nabývá až celá struktura typu *z tohoto důvodu* či *důvodem je*. Např. tedy slovo *důvod* není samo o sobě konektorem – nemá spojovací funkci ze své podstaty jako podřadicí a souřadicí spojky tvořící jádro primárních konektorů – také proto pro

tato konektivní spojení užíváme označení sekundární konektory. Slovo *důvod* navíc nemusí být nutně součástí konektivní struktury, a v některých případech se tedy na strukturaci diskurzu jako konektivní prostředek nepodílí – srov. příklad 42.

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

V jazyce existuje mnoho „prototypických“ primárních a sekundárních konektorů, které zcela odpovídají výše zmíněné charakteristice, srov. primární konektory *a, ale, nebo* a sekundární konektory *to je důvod, proč; podmínkou je či dojít k závěru*. Zároveň ovšem existují konektory, které se v některých bodech odlišují, např. primární konektory *aby, kdyby* podléhají flexi (*abych, kdybys...*), primární konektor *proto* může být modifikovaný (*hlavně / právě / jedině proto*) atd. Z tohoto hlediska nechápeme primární a sekundární konektory jako striktně uzavřené a oddělené skupiny výrazů, ale jako škálu jazykových prostředků v různém stupni gramatikalizace. To znamená, že některé sekundární konektory se v budoucnu mohou stát primárními – díky přirozeným jazykovým změnám a případné gramatikalizaci. Tento proces můžeme pozorovat na současných primárních konektorech (resp. na jejich historickém původu), které vznikaly ze struktur podobných současným sekundárním konektorům (srov. konektory *proto, přesto, přitom, potom, zato* původně vzniklé z předložkových frází). Více k prostupnosti hranic mezi primárními a sekundárními konektory a k úvahám ke vzniku primárních konektorů viz kapitoly 6.1.2.6 a 6.1.2.7.

Primární konektory bývají v různých studiích (srov. Rouchota, 1996; Fischerová, 2006; Urgellesová-Collová, 2010 aj.) užívány jako typičtí zástupci diskurzních konektorů, jejich frekvence ve vyjadřování diskurzních vztahů je mnohem vyšší než frekvence konektorů sekundárních (srov. kapitola 6.3.1). Z tohoto důvodu si můžeme primární konektory představit jako centrum množiny diskurzních konektorů a sekundární konektory jako její periferii.

6.1.2.1 Slovnědruhá charakteristika primárních konektorů

Ze slovnědruhového hlediska označujeme jako primární konektory především výrazy patřící mezi podřadící (*protože, když, pokud, aby...*) a souřadící (*a, ale, nebo, i...*) spojky a výrazy, které jsou na přechodu ke spojkám. Jedná se např. o výrazy *totiž, přitom, navíc, pak, potom, nakonec, například* či *proto*. Proces přechodu v jiný slovní druh nemusí být u těchto výrazů plně dokončen, a může tak docházet k nesouladu při určování jejich slovnědruhé příslušnosti. Obdobné výrazy jsou někdy chápány jako příslovce (např. *přitom, například* ve *Slovníku spisovného jazyka českého*, 1989), částice strukturující text (*totiž, přitom, například* v *Mluvnici češtiny 2*, 1986) či spojky (*totiž* ve *Slovníku spisovné češtiny*, 2003).

Vzhledem ke zmíněnému překrývání v rámci slovních druhů (především spojek, částic a příslovců) a ke skutečnosti, že u mnohých výrazů se spojovací funkcí dochází

k přechodu ve spojky, budeme chápat primární konektory především jako synsémantika, tj. výrazy bez větněčlenské platnosti.¹¹

6.1.2.2 Charakteristika sekundárních konektorů

Sekundární konektory tvoří širokou a rozmanitou škálu jazykových prostředků, a to po stránce lexikální, syntaktické i sémantické. Blíže můžeme tuto skupinu výrazů charakterizovat následujícím způsobem.

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 (například podstatné jméno jako *důvod*, *příčina*, *podmínka*, *vysvětlení* či sloveso typu *odůvodnit*, *zapříčinit*, *vysvětlit*), které signalizuje určitý typ diskurzního vztahu a které vstupuje do různých kombinací – srov. *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*).

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. *obecně* / *krátce* / *zjednodušeně řečeno*).¹² Tyto sekundární konektory jsou obvykle struktury, jejichž podstatou je jistá nepravidelnost či anomálie, např. omezená kombinovatelnost nebo nemožnost paradigmatické záměny – 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 Čermáka, 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. *stručně řečeno* – Čermák, 2007: 308, 442).

¹¹ 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 slouží k vyjádření diskurzního vztahu).

¹² 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 (např. z *tohoto důvodu*). Více k dané problematice viz např. Hoffmannová (1983).

Syntaktická charakteristika sekundárních konektorů

• Syntaktické fráze

Ze syntaktického hlediska můžeme sekundární konektory analyzovat podle toho, jaké typy syntaktických frází tvoří. Nejpočetnější jsou v PDiT verbální fráze (jejichž jádrem je sloveso, např. *odůvodnil to, to znamená, vysvětlila to*). Verbální fráze tvoří celkem 37 % v rámci 1 161 výskytů sekundárních konektorů v PDiT; k frekvenci konektorů v PDiT více viz kapitola 6.3.1. Dále se v PDiT s poměrně vysokou frekvencí objevují předložkové fráze obsahující anaforický výraz (25 %), např. *díky tomu; vzhledem k tomu; na základě toho; v důsledku toho; z tohoto důvodu*; předložkové fráze obsahující kataforický výraz (15 %), např. *kromě toho, že; z toho důvodu, aby; vzhledem k tomu, že*. Zbylých 23 % tvoří především ustrnulé výrazy typu *stručně řečeno, přeloženo, rozumějme* a fráze zakládající větné struktury – skládající se ze spojení slovesa s oslabeným lexikálním významem (např. *být, uvést, sloužit, tvořit*) a podstatného jména, které nese jádro lexikálního významu dané fráze (srov. *důvod je jednoduchý, příčinou je, uvést jako příklad, jako příklad slouží, výjimku tvoří*).

Ze syntaktického hlediska je zajímavé, že se některé sekundární konektory (zakládající větné struktury) v textu objevují ve formě hlavních či (nepravých) vedlejších vět – srov. příklady 43 a 44.

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

V příkladě 43 se vyskytuje sekundární konektor *to je důvod, proč*, který vyjadřuje vztah 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 výraz *proč*). Druhý diskurzní argument se tedy dostává do pozice věty závislé na konektoru. Jinými slovy, sekundární konektor je v tomto případě syntakticky výše než diskurzní argument.

Sekundární konektor *což znamená, že* v příkladě 44 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 diskurzní argumenty (z hlediska diskurzu rovnocenné úseky textu), nemusejí být na stejné úrovni ani syntakticky.

Některé sekundární konektory mohou mít také formu samostatné věty (srov. *Důvod je jednoduchý / jasný / zřejmý.*), srov. příklad 45.

- (45) *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. V tomto případě může být nahrazen i modifikovaným 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 diskurzní argumenty (úseky textu), které spojují. Tyto sekundární konektory tedy vykazují vysokou míru nezávislosti, protože tvoří syntakticky i sémanticky úplné úseky textu.

• Větněčlenská platnost sekundárních konektorů

Většina sekundárních konektorů či jejich komponentů (87 % v rámci všech výskytů¹³ sekundárních konektorů v PDiT) mají ve větě roli větných členů – např. předložkové fráze typu *kvůli tomu, z tohoto důvodu* 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 (13 %) tvoří konstrukce typu *stručně řečeno* nebo *jak je vidět*, které slouží jako větné modifikátory (vztahující se k obsahu celé věty či ke způsobu jeho vyjádření, tj. komentující obsah či styl celé výpovědi). V postupném procesu gramatikalizace nicméně větněčlenská platnost sekundárních konektorů slábne (viz kapitola 6.1.2.8).

• Rozvíjení sekundárních konektorů: modifikovatelnost

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

Rozvíjeny mohou být i sekundární konektory v podobě předložkových frází – srov. příklad 46 z PDiT se sekundárním konektorem *vinou toho* (signalizující diskurzní vztah důsledku), který je v tomto případě rozvit výrazem *především*:

- (46) *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.*

¹³ Do analýzy jsme nezahrnuli případy typu *Důvod je jednoduchý.*, tj. sekundární konektory tvořící samostatné věty, které stojí syntakticky mimo diskurzní argumenty, a neplní v nich tedy funkci větných členů.

V tomto ohledu se sekundární konektory liší od konektorů primárních, které být rozvíjeny v takovém rozsahu většinou nemohou (nelze například říci **významně ale*). Existují ovšem výjimky, jako je např. konektor *proto* – srov. *hlavně / jediné / právě proto*. Modifikovatelnost je v tomto případě ovšem omezená – srov. **významně proto*. Z tohoto hlediska může autor v konkrétním kontextu použít sekundární konektory právě *proto*, že mu umožňují přesněji vyjádřit jeho komunikační záměr (pokud chce například sdělit, že důvod k něčemu byl významný, nemůže použít primární konektor, který modifikovatelnost tímto slovem neumožňuje).

Sémantická charakteristika sekundárních konektorů

Sekundární konektory většinou obsahují alespoň jedno autosémantické slovo, na rozdíl od synsémantických konektorů primárních. Autosémantická bývají jádra jednotlivých konektorů – srov. podstatná jména typu *vysvětlení, příčina, podmínka* či slovesa typu *zdůvodnit, zapříčinit, znamenat, upřesnit*.¹⁴ U předložkových frází je jádrem předložka (typu *díky, kvůli, na rozdíl od, v důsledku, vinou*), tedy synsémantikum. Nicméně aby se příslušná předložka mohla stát součástí sekundárního konektoru, musí se spojit s plnovýznamovým (anaforickým či kataforickým) slovem – jako sekundární konektor fungují v textu fráze typu *díky tomu; v důsledku toho; vinou toho; kromě toho, že*. Jako konektor tedy nechápeme samotné *díky, kvůli*, ale celou předložkovou frázi *díky tomu, kvůli tomu* (můžeme zde vidět analogii k některým současným primárním konektorům, které vznikly z podobných předložkových frází, např. *přesto* ze spojení předložky *přes* a zájmena *to*). Při rozhodování o tom, co je a co není součástí konektoru, někdy můžeme využít také jinojazyčné ekvivalenty. Například němčina má pro české *kvůli tomu* v konektivním významu plně lexikalizovaný protějšek *deswegen* skládající se rovněž z předložky a výrazu s koreferenční funkcí.

6.1.2.3 Variabilita sekundárních konektorů: lexikální realizace a gramatické varianty

Vzhledem k tomu, že sekundární konektory jsou dosud ne plně gramatikalizované struktury, vykazují poměrně vysoký stupeň variability (oproti konektorům primárním) a mnohdy je obtížné určit, zda dvě podobné struktury (např. *za této podmínky* vs. *za té podmínky*, *z tohoto důvodu* vs. *z uvedeného důvodu*, *to je důsledkem* vs. *je to důsledkem*, *příčina je* vs. *příčinou je*) jsou dva různé sekundární konektory či pouze varianty téhož konektoru. Proto je zapotřebí sekundární konektory dále hierarchizovat.

Jak jsme uvedli v úvodu kapitoly 6.1.2, sekundární konektory obsahují tzv. jádrová slova (tvořící základ daných struktur). Můžeme pozorovat, že mnohé konektory se stejným jádrovým slovem jsou tvořeny podle stejného obecného schématu. Např. struktury *za této podmínky*, *za té podmínky*, *za dané podmínky* patří ke stejnému schématu „*za_Pron.anaph / Adj.anaph_podmínka.2*“ (jde o předložkovou frázi, ve které se před-

¹⁴ Více ke slovesným konektivním strukturám viz Rysová (2014).

ložka *za* objevuje ve spojení s anaforickým zájmenem či přídavným jménem a slovem *podmínka* v genitivu). Tyto struktury označujeme jako lexikální realizace a vymezujeme je jako formy sekundárních konektorů se stejným jádrovým slovem i obecným schématem, ale do jisté míry různým lexikálním obsazením (srov. *této / té / dané*).

Jednotlivé lexikální realizace se zároveň mohou v textu objevovat v různých gramatických variantách, které se liší v oblasti morfologie (*za této podmínky* vs. *za těchto podmínek*) či slovosledu (*to je příčinou* vs. *je to příčinou*).

Sekundární konektory tedy mohou být blíže popisovány podle následující hierarchie: i) obecné schéma sekundárních konektorů („*za_Pron.anaph / Adj.anaph_podmínka.2*“), ii) jeho lexikální realizace (*za této podmínky*, *za té podmínky*, *za dané podmínky*), iii) gramatické varianty těchto realizací (*za této podmínky* vs. *za těchto podmínek*).

Nyní se vrátíme k otázce, které struktury můžeme považovat za samostatné konektory a které pouze za varianty konektoru. V naší práci řadíme všechny lexikální realizace téhož obecného schématu (se stejným jádrovým slovem) k jednomu sekundárnímu konektoru, tj. nepovažujeme výrazy *za této podmínky*, *za té podmínky*, *za dané podmínky* za tři samostatné sekundární konektory, ale za realizace téhož konektoru. Jako základní formu daného konektoru označujeme jeho nejfrekventovanější realizaci (v základním gramatickém tvaru), tj. v našem případě strukturu *za této podmínky* (měřeno na datech PDiT 2.0).

Základní formou sekundárního konektoru se schématem „*za_Pron.anaph / Adj.anaph_podmínka.2*“ je tedy výraz *za této podmínky*, jeho dalšími (méně frekventovanými) formami jsou dále *za té podmínky* či *za dané podmínky*. Všechny tyto realizace se zároveň mohou vyskytovat v různých gramatických variantách (srov. *za těchto podmínek*, *za těch podmínek* či *za daných podmínek*). Na druhou stranu výrazy patřící k jinému schématu (např. *pod podmínkou*, *že*) chápeme jako jiné, samostatné konektory.

6.1.2.4 Primární a sekundární konektory a povaha diskurzních argumentů

V této kapitole se zaměříme na rozbor primárních a sekundárních konektorů z hlediska syntaktické povahy argumentů, které spojují. V některých pojetích (např. Anders a Søgaard, 2013) je diskurzní konektor vymezován mj. právě na základě syntaktické realizace argumentů – je zde 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*).

Takovéto vymezení konektorů je ovšem problematické, pokud do diskurzní analýzy zahrneme i sekundární konektory typu *příčinou je* nebo *odůvodnil to*. Zatímco některé sekundární konektory 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 dis-

kurzních argumentů, mezi kterými stojí, viz bod b).

a) Konektory mezi slovesnými argumenty

Většina diskurzních konektorů (primárních i sekundárních) obvykle spojuje dva úseky textu realizované slovesnými argumenty. V těchto případech bývají primární a sekundární konektory vzájemně nahraditelné – srov. příklad 47, ve kterém je spojení z *tohoto důvodu* zaměnitelné za primární konektor *proto* a význam zůstane zachován:

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

b) Konektory mezi nominálními argumenty

Některé konektivní struktury dovolují nominalizaci druhého argumentu (viz kapitola 3.5.2) – za konektorem nenásleduje celá věta, ale nominální fráze – srov. příklady 48 a 49.

- (48) *Privatizované mlékárny se však zatím mezi sebou nedokázaly domluvit. Důsledkem je nekompromisní konkurenční boj, který tlačí ceny výrobků až takřka k nulové rentabilitě zpracovatelů.*
- (49) *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).*

Nominalizace druhého argumentu je v těchto případech alternativou k vyjádření argumentu vedlejší větou – srov. *Důsledkem je, že mezi sebou nekompromisně konkurenčně bojují... a Důvodem je, že napadl ředitele soupeřova týmu...* Nominalizace se může vyskytovat i u prvního diskurzního argumentu, srov. příklad 50.

- (50) *Dalším tématem našich zpráv je zvýšení cen benzínu. Kvůli tomu si některé rodiny nebudou moct dovolit jezdit autem.*

Příklad můžeme parafrázovat následujícím způsobem: *Zvýšily se ceny benzínu, a proto si některé rodiny nebudou moct dovolit jezdit autem.* 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 textech PDiT se nicméně podobný příklad neobjevil.

Rozdíl mezi nominálním a verbálním vyjádřením argumentu je syntaktický, nikoli sémantický. Z hlediska diskurzu jsou tedy obě varianty rovnocenné.¹⁵ Z tohoto důvodu jsme se rozhodli nebrat syntaktickou povahu diskurzních argumentů jako jedno z kritérií pro vymezení konektorů. Za diskurzní konektor považujeme takové jazykové vyjádření, které spojuje dva diskurzní argumenty, ať už jsou, např. vlivem

¹⁵ 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).

struktury konektivního prostředku, na povrchu realizovány slovesnou či nominální frází. Více k výskytu nominálních argumentů v Pražském diskurzním korpusu viz kapitola 6.3.2.

6.1.2.5 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 nejčastěji o příklady, ve kterých se vyskytuje diskurzní vztah mezi dvěma slovesnými argumenty – srov. příklad 51 z PDiT, ve kterém je sekundární konektor z *tohoto důvodu* nahraditelný primárním konektorem *proto*.

- (51) *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 případě, že je druhý diskurzní argument nominální (příklad 52i z PDiT), náhrada sekundárního konektoru za primární si vynucuje syntaktické změny příslušného argumentu (příklad 52ii).

- (52) i) *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.*

ii) *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.*

V případě, že je sekundární 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 rozvoitími, která se objevují ve struktuře daného sekundárního konektoru – např. sekundární konektor *Důvod je jednoduchý* může být nahrazen modifikovaným primárním konektorem *jednoduše proto*, že, přičemž lexikální význam zůstane zachován. V případě komplexnějšího rozvoje je náhrada primárním konektorem pouze částečná, některé lexikální složky nemohou být při transformaci zachovány – srov. sekundární

konektor z PDiT *Další důvod je složitější*.¹⁶ nemůže být plně nahrazen modifikovaným 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 PDiT: *Důvod je jednoduchý; Důvod je složitý; Důvod je filozofický*. V tomto ohledu se nabízí otázka, zda v podobných případech nemůže být dílčím kritériem pro vymezení sekundárních konektorů právě nahraditelnost konektorem primárním – 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 rozvití (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 zároveň pouze *hlavně proto, že*, nikoli **zásadně / základně / podstatně proto, že* (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 pomocné), držíme se především principu kontextové (ne)závislosti. Tedy čím je rozvití v dané struktuře lexikálně obecnější, tím má větší šanci stát se v textu součástí diskurzního konektoru. Z tohoto hlediska strukturu *Důvod je filozofický*. 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 53 a 54 z PDiT.

- (53) *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.*
- (54) *Gyula Horn se vyslovil pro možné zavedení majetkové daně. Zdůvodnil to tím, že utahování opasků se nemůže vztahovat pouze na lidi žijící ze mzdy.*

¹⁶ 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.

V příkladě 53 je mezi danými argumenty diskurzni vztah důsledku, v příkladě 54 vztah příčiny, v obou případech je diskurzni vztah vyjádřený sekundárními konektory – *to je důvod, proč* a *zdůvodnil to tím, že*. Při ověřování rozsahu těchto konektorů (tedy hranice mezi sekundárním konektorem a diskurzni argumentem) nám může pomoci náhrada konektorem primárním – srov. příklady 55 a 56:

- (55) *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.*
- (56) *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 diskurzni argumenty daného vztahu (tj. v příkladě 55 *ještě nebyla podepsána smlouva s dodavatelem*, v příkladě 56 *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 53 a 54, 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 diskurzni 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*, jejichž součástí je zesilovací příklonka *-ž(e)* (srov. Holub a Kopečný, 1952; Rejzek, 2001).

6.1.2.6 Prostupnost hranic mezi primárními a sekundárními konektory

Mezi sekundárními konektory můžeme vysledovat různé podskupiny, které jsou svou podstatou a chováním blízké konektorům primárním. Velkou skupinu sekundárních konektorů tvoří předložkové fráze obsahující předložku (např. *díky, kvůli, vinou*) a anaforický výraz odkazující k prvnímu diskurzni argumentu. Jako sekundární konektory se tedy v textu objevují spojení typu *díky tomu, kvůli tomu, vinou toho*. Tyto sekundární konektory jsou svojí strukturou 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 analogic-

ké příklady typu *přesto, potom, předtím, zatím, přitom, zato, nato, beztoho, vtom, nadto*). 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.

6.1.2.7 Úvahy ke vzniku primárních konektorů

Jak jsme uvedli v úvodu kapitoly 6.1.2, primárními konektory rozumíme výrazy patřící především mezi podřadící a souřadící spojky (a některé další výrazy s konektivní funkcí, které jsou na přechodu ke spojkám). Na základě předchozích úvah o prostupnosti hranic mezi primárními a sekundárními konektory a postupném procesu ustalování konektorů jsme se dále zaměřili na vznik primárních konektorů obecně.

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

„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)

Bauer dále zdůrazňuje, že spolehlivé kritérium pro rozlišení slovnědruhové funkce spojovacích výrazů neexistuje.

„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 mluvnické 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, 1962b: 29)

V Bauerových člancích můžeme najít rovněž analýzu původu některých (v současné době typických) primárních konektorů – např. *a*, *ale*, *však*, *avšak*, *proto*, *přesto*, *tedy*, *tudíž*, *přece*. 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* spojilo s *však* ve spoju *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, 1962b: 30). *Tedy* má původ v časovém *tehdy*; *tudíž* v místním (a přeneseně způsobovém) *tudy*. *Přece* vzniklo ze spojení *před* s původním významem ‚dále kupředu‘ (Bauer, 1962b: 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 (více k oslabování lexikálního významu u sekundárních konektorů viz kapitola 6.1.2.8).

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

Z tohoto hlediska můžeme chápat sekundární konektory 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).¹⁷

Současné primární konektory tedy pravděpodobně vznikaly z podobných více-slovných frází či slovních druhů jako současné sekundární konektory. Tento jev můžeme sledovat i v jiných jazycích – např. anglický konektor *because* (*protože*) vznikl ze

¹⁷ Více ke gramatikalizaci konektorů viz Rysová (2017).

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

V procesu vývoje spojovacích prostředků můžeme pozorovat několik fází. Na začátku obvykle stojí plnovýznamová struktura, která nebývá gramaticky ani lexikálně omezená. Postupně u ní dochází k oslabování významu a větněčlenské platnosti, nabývá gramatických omezení. Poté se lexikalizuje jako jednotka s primární spojovací funkcí. V průběhu tohoto vývoje dochází k zániku některých variant dané jednotky, srov. staročeské *proto*, *protož*, *protožť*, *protožto*, *protože*, *protoť*. V dalších případech dochází k posunu významu. Například spojka *čili* se dnes již neužívá v dříve běžných spojeních typu *Je to pravda čili nic?* ve významu ‚Je to pravda nebo ne?‘. Zároveň v jazyce v průběhu vývoje nedochází jen ke vzniku konektorů, ale také k jejich postupnému zániku, srov. staročeské, dnes již neužívané konektory *anébrže*, *anež*, *neda*.¹⁸

6.1.2.8 K otázce plnovýznamovosti

V úvodu kapitoly 6.1.2 jsme jako jeden z rozdílů mezi primárními a sekundárními konektory uvedli skutečnost, že primární konektory jsou synsémantika, zatímco konektory sekundární obsahují autosémantickou složku. Nicméně i u sekundárních konektorů dochází postupně k ustalování formy i funkce a k vyprazdňování původního lexikálního významu.

Dichotomii plnovýznamovosti vs. neplnovýznamovosti problematizuje ve svých studiích např. J. Bauer. V článku *Spojky a příslovce* (1962b) upozorňuje na problematiku zařaditelné výrazy, které jsou od původu příslovce, ale které nemají plný lexikální význam:

„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 zájmenných příslovců vztahných je přechod ke spojkové platnosti velmi častý, neboť relativa jsou už ve své původní funkci slova spojovací.“ (Bauer, 1962b: 29)

Oslabování vlastního věcného významu je vysvětlitelné právě přechodem daných výrazů ke spojkám, tj. k funkci primárních konektorů. Tento jev ale nenastává pouze na přechodu mezi příslovci a spojkami, ale u všech výrazů, které jsou v procesu ustalování a nabývání konektivní funkce – tedy i sekundárních konektorů.

¹⁸ Srov. *Elektronický slovník staré češtiny* (2018).

U sekundárních konektorů 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 mohou plnit 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* (2003) 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* (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* nebo *v té věci*, tj. spojení, která podobně jako *v tom / tomto směru* přispívají ke strukturaci textu. V obou slovnících je tedy konektivní funkce struktury *v tom / tomto směru* zachycena.

Můžeme shrnout, že v kategorii sekundárních konektorů dochází k ustalování formy, fixování konektivní funkce a oslabování vlastního věcného významu jednotlivých složek dané struktury, která jako celek nabývá v textu platnosti diskurzního konektoru. Základní rozdíly mezi primárními a sekundárními konektory shrnuje tabulka 6.1.

Primární konektory	Sekundární konektory
Gramatikalizované	(Dosud) ne plně gramatikalizované
Synsémantika	Struktury s autosémantickou složkou
Většinou jednoslovné (+ korelativní páry typu <i>bud' nebo</i> a sousloví typu <i>i když</i>)	Většinou víceslovné
Bez větněčlenské platnosti	Větné členy, větné modifikátory či samostatné věty
Nemodifikovatelné	Modifikovatelné Jedinečnost některých struktur: a) syntakticky výše než druhý argument b) forma samostatné věty c) nominalizace druhého argumentu

Tabulka 6.1: Charakteristika primárních a sekundárních konektorů

6.1.2.9 Volné konektivní fráze

Jak jsme ukázali v úvodu kapitoly 6.1, diskurzní konektory (ať už primární, či sekundární) nejsou jediné jazykové prostředky, kterými je možné v textu vyjadřovat diskurzní vztahy (viz příklad *Petr je nemocný a nemůže s námi jet. Kvůli jeho nemoci se výlet ruší.*). Pro výrazy typu *kvůli jeho nemoci* používáme termín volná konektivní fráze (prvně použitý v článku Rysová a Rysové, 2014). Volné konektivní fráze se od diskurzních konektorů liší tím, že jsou silně závislé na kontextu, tj. tyto fráze mohou být použity pro vyjádření příslušného diskurzního vztahu pouze v několika specifických kontextech – srov. příklady 57i a 57ii.

(57) i) *V Torontu je silná bouřka a mlha. Proto / Kvůli tomu / Kvůli tomuto počasí budeme muset 20–30 minut kroužit nad letištěm.*

ii) *Slovenská elita byla zklamána politickou volbou Slovenska. Proto / Kvůli tomu / *Kvůli tomuto počasí většina kvalitních odborníků zůstala v Praze.*

Příklady 57i a 57ii obsahují mezivětný vztah důsledku, který může být v obou případech vyjádřený výrazy *proto* i *kvůli tomu*, oproti tomu spojení *kvůli tomuto počasí* může být použito pouze v příkladu 57i. Z tohoto důvodu považujeme za konektory pouze *proto* a *kvůli tomu*, zatímco *kvůli tomuto počasí* označujeme termínem volná konektivní fráze.

Volné konektivní fráze zároveň obsahují specifickou referenční složku (např. *tomuto počasí, této nemoci, tomuto zvýšení*), a vždy *proto* musí být chápány v rámci širšího kontextu.

6.1.3 Diskurzní konektory a reference

Odkazování do kontextu je významnou složkou diskurzních konektorů obecně. Jak jsme popsali v úvodu kapitoly 6.1.1, základní vlastností konektorů je, že spojují dva úseky textu (dva diskurzní argumenty). Pokud se tedy v textu objeví diskurzní konektor, můžeme předpokládat, že určitým způsobem odkazuje do předchozího a následujícího kontextu. Obecně řečeno, přítomnost konektoru implikuje přítomnost dvou diskurzních argumentů (viz např. Halliday a Hasanová, 1976). V tomto ohledu můžeme všechny konektory chápat jako implicitně referenční (např. konektory jako *ale*, *a* či *nebo* neobsahují explicitní referenční komponent, ale přesto jasně signalizují přítomnost diskurzních argumentů).

Mezi diskurzními konektory ale můžeme zároveň vysledovat užší skupinu výrazů obsahujících referenční komponent explicitně, srov. příklady sekundárních konektorů typu *kvůli tomu; z tohoto důvodu; příkladem toho je; za této podmínky*. I mezi těmito výrazy ovšem můžeme pozorovat rozdíly – některé z nich (např. *kvůli tomu*) obsahují anaforický prvek obligatorně, zatímco u jiných být vyjádřen může i nemusí (srov. *příkladem toho je* vs. *příkladem je*).

Z hlediska reference tedy můžeme diskurzní konektory dělit následujícím způsobem.

- 1) Konektory **bez inherentní reference** (např. *ale, nebo, a, tak, však*)
- 2) Konektory **s inherentní referencí**, která může být:
 - a. **fakultativní** (srov. *výsledkem je* vs. *výsledkem toho je*)
 - b. **obligatorní**:
 - i. konektory gramatikalizované (*proto, přitom, zato, protože*)
 - ii. konektory dosud negramatikalizované (*kvůli tomu, z tohoto důvodu*)

Z hlediska dosahu můžeme referenční konektory dále dělit na anaforické (*proto, potom, z tohoto důvodu, díky tomu*) a kataforické (*protože; kvůli tomu, aby; kromě toho, že*), srov. následující příklady – v prvním sekundární konektor *z toho důvodu* odkazuje do předchozího kontextu (tj. působí anaforicky), ve druhém *z toho důvodu*, že do následujícího kontextu (tj. působí kataforicky).¹⁹

- (58) *Pro otce bylo velmi těžké udržet si stálou praxi. Z toho důvodu_{anaf} jsme se často stěhovali z místa na místo.*
- (59) *Praktické měření vzdálenosti hvězd je svízelné z toho důvodu, že_{kataf} hvězdy s výjimkou Slunce jsou od nás velice daleko.*

6.1.4 Souvýskyt diskurzních konektorů - složené konektory

Další významnou vlastností diskurzních konektorů je, že se v textu často kombinují. Vyskytují se v textu vedle sebe a fungují jako jeden složený konektor (vyjadřující jeden diskurzní vztah) – srov. příklady 60, 61 a 62.

- (60) *Tajemnou propastí je člověk, i když v sobě tu propast nevidí.*
- (61) *Byl jsi mým nejvěrnějším služebníkem, a proto se budu řídit tvou radou.*
- (62) *Měl exkluzivní smlouvu na všechna spisovatelova díla, a díky tomu se s ním řadu let intenzivně stýkal.*

Ve všech třech příkladech se objevuje složený diskurzní konektor (*i když, a proto, a díky tomu*), tj. dva konektory se v textu vyskytují vedle sebe a dohromady vyjadřují jeden diskurzní vztah (*i když* signalizuje vztah přípustky, *a proto* a *a díky tomu* vztah důsledku).

Složené konektory se přitom liší v míře či pevnosti svého spojení, např. složky konektoru *i když* od sebe nemohou být odděleny jiným textem (nejde např. říci **Tajemnou propastí je člověk, i v sobě když tu propast nevidí.*), zatímco složky konektoru *a proto* ano

¹⁹ Více ke vztahu diskurzních konektorů a reference viz Rysová a Rysová (2018a, 2018b), Rysová a Mírovský (2014).

(*Byl jsi mým nejověrnějším služebníkem, a budu se proto řídit tvou radou*).²⁰ Jak jsme ukázali na příkladech 60, 61 a 62, složené konektory mohou tvořit konektory primární (*i když, a proto*) i sekundární (*a díky tomu*).²¹

Složené konektory se v textu mohou vyskytovat v rámci jednoho argumentu (příklady 60, 61 a 62) nebo mohou tvořit korelativní páry, tj. jedna složka konektoru se objevuje v argumentu 1, druhá v argumentu 2 (příklad 63).

(63) *Buď si změnila jméno, nebo se vdala.*

V naší práci tedy rozlišujeme (kromě konektorů primárních a sekundárních) také 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** se skládá z více konektivních výrazů²² (ať už jde o kombinaci dvou primárních konektorů, nebo kombinaci konektoru primárního a sekundárního), např. *a tak; i když; a proto; a z tohoto důvodu*. Složené konektory jsou vždy víceslovné, přičemž všechny jejich složky se podílejí na vyjádření jednoho typu diskurzního vztahu (např. celé spojení *i když* vyjadřuje vztah přípustky).

V některých případech může vyvstat otázka, zda je dané slovní spojení složeným konektorem nebo není. U složených konektorů (typu *a tak, i proto*) 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. složený konektor *a proto* vyjadřuje „silněji“ či zřetelněji vztah důsledku než vztah konjunkce, srov. pří-

²⁰ Konektor *i když* je zároveň charakteristický tím, že jeho jednotlivé složky (*i a když*) nevyjadřují samy o sobě stejný typ vztahu, jako když jsou užity vedle sebe (tj. vztah přípustky). Tím se liší např. od konektoru *a proto*, jehož obě složky (*proto* zřetelně, *a* v ojedinělých případech) mohou vyjádřit stejný typ vztahu, jako když jsou užity dohromady (tj. vztah důsledku), i samostatně.

²¹ Je tedy zjevné, že sekundární konektory v tomto ohledu vykazují podobné chování jako konektory primární, tj. mohou se v textu objevovat jako složené výrazy (nejčastěji se přitom kombinují se spojkou *a*). 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 – *ale* se skládá z výrazu *a* (které v době sloučení bylo pravděpodobně již spojkou či na přechodu ke spojce) a z komponentu *-le* (s významem omezovacího *jen*); *avšak* vzniklo obdobně z komponentů *a* a *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 (píší se zvlášť), ale které se v textu vyskytují s vysokou frekvencí a jsou v současné době v PDiT vnímány jako jeden konektor. Konektivní struktury typu *díky tomu* 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 PDiT jsou tyto případy, ve kterých byly příslušné vztahy signalizovány primárním a sekundárním konektorem (tj. příklady typu *a z tohoto důvodu*), chápány jako složený konektor vyjadřující jeden diskurzní vztah.

²² Ve většině případů fungují složky složeného konektoru jako konektory i samostatně (srov. *i když, a proto* apod.), v některých případech se konektivní výrazy samostatně nepoužívají, tj. fungují pouze jako část složeného konektoru (např. *buď* v rámci složeného konektoru *buď_nebo*).

klad 64). U daných konektorů 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.²³ 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 různé diskurzní konektory. U každého z nich označujeme daný významový typ a dosah diskurzního vztahu zvlášť, viz příklad 65 se dvěma samostatnými konektory *ale* a *potom* vyjadřujícími zcela zřetelně dva různé typy vztahů (*ale* vztah opozice, *potom* vztah nesoučasnosti – předčasnosti a následnosti).

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

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

6.1.4.1 Modifikované konektory

Jak jsme ukázali výše, složené konektory jsou vždy víceslovné. Nemůžeme ovšem říci, že víceslovné konektory jsou vždy složené. Např. sekundární konektory typu *z tohoto důvodu* nebo *za této podmínky* považujeme za jeden (negramatikalizovaný) konektor.

Obdobně se v jazyce vyskytují jiné víceslovné struktury typu *hlavně proto*; *právě kvůli tomu*; *především pak*, které rovněž nesplňují naše vymezení složených konektorů – tj. neskládají se ze dvou či více konektorů. Tyto výrazy označujeme jako modifikované konektory, srov. další příklady jako *z jednoduchých důvodů*; *vlastně proto*, *že*; *jediným důvodem bylo*; *s jasným cílem*; *možná díky tomu*; *právě proto* aj. obsahující určitý hodnotící, modální či intenzifikační výraz, který sám o sobě nemá konektivní povahu (tj. nevyjadřuje příslušný diskurzní vztah, ale pouze ho modifikuje, upřesňuje či vyjadřuje jeho intenzitu). Modifikace se (vzhledem k jejich lexikálně-syntaktické povaze) vyskytují častěji u konektorů sekundárních než primárních. Modifikační výrazy zároveň u příslušných konektorů snižují jejich kontextovou nezávislost – tj. např. konektor *proto* je vhodnější pro více kontextů než jeho modifikovaná varianta *hlavně proto*.

Jak jsme se snažili popsat v předchozích kapitolách, diskurzní konektory tvoří značně heterogenní skupinu výrazů, v textu se vzájemně kombinují, mohou se objevovat s různými modifikacemi apod. V naší práci se ovšem zaměřujeme především na opozici konektorů primárních a sekundárních. V následující části tedy dále představíme zpracování sekundárních konektorů v PDiT (kapitola 6.2) a posléze se zaměříme na vyhodnocení této anotace (kapitola 6.3 a 6.4).

²³ 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 potlačen.

6.2 Anotace sekundárních konektorů v Pražském diskurzním korpusu 2.0

Práce na anotaci sekundárních konektorů začaly v roce 2012, resp. prvotně začala být zpracovávána tzv. alternativní lexikální vyjádření konektorů. Podnět k jejich zpracování vzešel z prvotní anotace konektorů v korpusu PDiT 1.0, v průběhu které anotátoři dostali pokyn (kromě anotace konektorů vymezených v kapitole 4.2.1) označovat v poznámkách i takové jazykové prostředky, které podle nich slouží k vyjádření diskurzního vztahu, nejsou „klasickými“ konektory, ale mohly by být zařazeny do kategorie alternativních vyjádření konektorů v pojetí Prasadové a kol. (2008, 2010).

V rámci prvotní diskurzní anotace tedy celkově nešlo o systematické zpracování alternativních vyjádření konektorů, ale především o sběr jazykového materiálu k dalšímu výzkumu (dané poznámky byly pracovní). Celkem se v rámci této anotace objevilo 306 anotátorských poznámek upozorňujících v textu na výskyt alternativních vyjádření konektorů, jejich zpracování přitom nebylo zcela konzistentní (určité typy spojovacích prostředků nebyly například zachyceny ve všech svých výskytech, srov. Rysová, 2012, 2013).

Při výzkumu těchto výrazů se ukázalo, že množina alternativních vyjádření konektorů stanovená Prasadovou a kol. (2008, 2010) není zcela jasně vymezená a zahrnuje široké spektrum jazykových prostředků. Za altlexy jsou považovány např. kontextově závislé (*the increase was due mainly to* – *zvýšení bylo hlavně kvůli tomu*) i nezávislé fráze (*that is why* – *to je, proč*), viz Prasadová a kol. (2010: 1026). Není také jasná hranice mezi konektorem a jeho alternativním vyjádřením, srov. kapitola 3.8.5.

Z těchto důvodů došlo k novému vymezení pojmu diskurzní konektor a k rozdělení konektorů na primární a sekundární (Rysová a Rysová, 2014, 2015). V roce 2014 jsme tedy přestali používat termín alternativní vyjádření konektoru (altlex) a začali používat termín sekundární konektor.

Na začátku práce se sekundárními konektory stálo tedy 306 anotátorských poznámek upozorňujících na možný výskyt alternativního vyjádření konektoru. Tyto poznámky bylo třeba protřídit v souladu s novým vymezením konektorů. Poté bylo na jejich základě možné sestavit prvotní seznam sekundárních konektorů, resp. spíše jejich jádrových slov vyskytujících se v PDiT:

argumentovat, díky, dlužno dodat, do třetice, dodat/dodávat, doplnit/doplňovat, důsledek (důsledkem je; to je důsledek), důvod (z tohoto důvodu; z těchto důvodů; z uvedených důvodů), jak je vidět, jako poslední, kontrastovat, kvůli, na prvním místě, na rozdíl od, na základě, následovat, navzdory, ne dosti na tom, nehledě na, nemluvě o, o to více, odhadnout, odůvodnění (s odůvodněním, že), odůvodnit / odůvodňovat, plynout, podmínka (být podmínkou; s podmínkou, že), podobně, pokračovat, pravděpodobnější, právě tak, předcházet, případně, přeloženo, přes, přesněji, přičíst, příčina (být příčinou), příklad (být příkladem; příklad pochází z; jako příklad slouží; uvést jako příklad), případ (být opačným případem; neplatí to v případě; v (tomto) případě), přispívat, přistupovat, rozdíl (být rozdílem), rozpor (to je v rozporu), řečeno (stručně řečeno; jinak řečeno; zjednodušeně řečeno; jednoduše řečeno), souviset, stejným dechem, tím

spíš, týkat se, účel (za tím / tímto účelem; pro tento účel), upřesnit, směr (v tomto směru), v (této) souvislosti, ve skutečnosti, věst, vinou, výjimka (být výjimkou; tvořit výjimku), vyplývat, výsledek (být výsledkem), vysvětlit, vzhledem k, závěr (závěr je; závěr zní), zdůvodnit / zdůvodňovat, znamenat, způsobit.

Prvotní seznam zachycuje především jádrová slova jednotlivých sekundárních konektorů (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 PDiT (označené prvními anotátory diskurzních vztahů) – např. *výjimka: být výjimkou; tvořit výjimku*. Ustálené fráze jsou uvedeny celé, resp. v takové podobě, v jaké mají funkci sekundárních konektorů (např. *přeloženo*). Byli jsme si zároveň vědomi toho, že daný seznam sekundárních konektorů není konečný – vzhledem k rozmanitosti sekundárních konektorů jsme předpokládali, že se v textech budou objevovat další struktury s konektivní funkcí, které v anotaci původně nebyly zachyceny. Příslušný seznam sekundárních konektorů tak byl dále doplňován a upravován.

Na základě uvedeného seznamu jsme v PDiT automaticky vyhledávali jednotlivá jádrová slova (příp. celé ustálené fráze) – např. všechny výskyty slova *důvod*. Pro tato vyhledávání byla použita celá data PDiT, 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 výraz ve funkci sekundárního konektoru (pak byl daný výskyt anotátorem označen jako sekundární konektor) nebo o výskyt struktury sice shodné s určitým sekundárním konektorem či jeho jádrovým slovem, ale v nekonektivní funkci (pak daný výskyt jako sekundární konektor anotován nebyl).

Anotátoři zároveň určovali, jaký typ vztahu daný sekundární konektor vyjadřuje, a 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, tj. značeny byly vztahy mezi slovesnými i nominálními argumenty (na rozdíl od prvotní anotace konektorů, při níž se označovaly vztahy pouze mezi argumenty slovesnými).

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 PDiT 1.0, tj. v době, kdy ještě nebyly 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 dělení konektorů na primární a sekundární).

Jednalo se především o ustálené fráze typu *jinými slovy, na jedné straně* a předložkové fráze typu *kromě toho, naproti tomu*. Kompletní abecední a frekvenční seznam primárních a sekundárních konektorů zachycených v současné anotaci v PDiT je uveden v příloze.

Do anotace byly zároveň zahrnuty volné konektivní fráze (tj. silně kontextově závislé konektivní struktury typu *díky tomuto počasí*). Nicméně vzhledem k tomu, že se nejedná o diskurzní konektory, anotovali jsme je pouze poznámkou upozorňující na jejich výskyt v korpusu (tj. nezachycovali jsme např. typ diskurzního vztahu, který vyjadřují, nebo rozsah jejich argumentů).

Manuální anotaci sekundárních konektorů a volných konektivních frází prováděli tři různí anotátoři (všichni se zároveň podíleli na předchozí anotaci konektorů v PDiT 1.0, měli tedy s anotací diskurzu již zkušenosti). Po dokončení těchto prací byla celá anotace znovu zkontrolována.

6.2.1 Dvouúrovňová anotace sekundárních konektorů

Vzhledem ke značné variabilitě sekundárních konektorů a vzhledem k tomu, že většina z nich obsahuje tzv. jádro (často vyjadřující sémantický typ diskurzního vztahu – např. *důvod, příčina, kontrastovat, na rozdíl od, znamenat*), jsme se rozhodli pro dvouúrovňovou anotaci těchto konektorů. U každého konektoru anotátor označil jeho jádro a zároveň zapsal celou konektivní strukturu, a to v takové podobě, jak se vyskytovala v textu – srov. obrázek 6.1 z PDiT. Obrázek zachycuje stromovou strukturu následujícího souvětí:

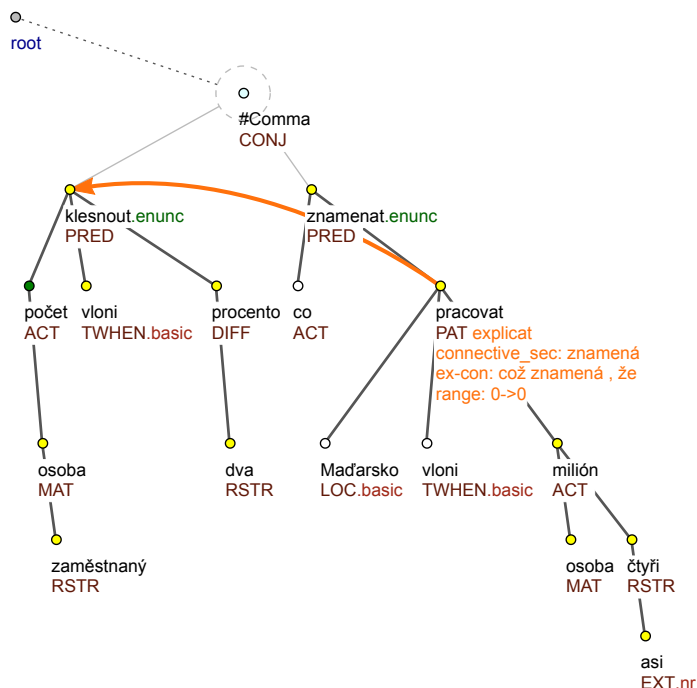
- (66) *Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.*

V daném příkladě je mezi diskurzními argumenty (*Počet zaměstnaných osob klesl loni o dvě procenta* a *v Maďarsku loni pracovaly asi čtyři milióny osob*) diskurzní vztah explikace (vysvětlení) znázorněný šipkou mezi uzly *klesnout* a *pracovat* reprezentujícími celé argumenty. Diskurzní vztah je explicitně vyjádřen sekundárním konektorem *což znamená, že*. Jádrem konektoru je v tomto případě sloveso *znamenat*.

Dvouúrovňová anotace sekundárních konektorů umožňuje (automaticky) vyhledávat jednotlivé konektory podle jejich základů či jader, tj. je možné např. vyhledat všechny sekundární konektory s jádrem *důvod* nebo *výjimka*. Můžeme tak např. vysledovat, v jakých konektivních strukturách se dané jádro objevuje (*příčina: příčinou je, vidět příčinu v, hledat příčinu v* atd.) či která realizace převažuje a začíná se tak ustalovat jako nejčastěji používaný sekundární konektor v rámci příslušných struktur.

6.2.2 Evaluace anotace

Anotace sekundárních konektorů v Pražském diskurzním korpusu začala v roce 2012, vydána byla v roce 2016 jako PDiT 2.0 (Rysová a kol., 2016). Anotaci prováděli tři ano-



Obrázek 6.1: Ukázka dvouúrovňové anotace sekundárních konektorů v Pražském diskurzním korpusu na souvětí: *Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.*

tátoři, kteří se podíleli již na anotaci konektorů v PDiT 1.0. Anotace zachycuje diskurzní vztahy mezivětné i vnitrovětné.

V průběhu anotace byla změřena mezanotátorská shoda, která slouží především k vyhodnocení úspěšnosti anotačních principů na konkrétních datech, k hledání případných problémových jevů nebo k odhalení nejčastějších anotátorských chyb. Měření mezanotátorské shody proběhlo na 500 větách anotovaných dvěma anotátory zároveň, zaměřili jsme se přitom na sedm různých aspektů.

1. Nejprve byla změřena celková shoda v rozpoznání existence vztahu (tj. do jaké míry se anotátoři shodují v tom, že na daném místě v textu existuje nějaký diskurzní vztah vyjádřený sekundárním konektorem);
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),²⁴

3. dále byla změřena shoda na odlišení sekundárního konektoru a volné konektivní fráze a

4. určení typu argumentu (tj. zda se jedná o sekundární konektor mezi verbálními či nominálními argumenty);

5. dále jsme sledovali, jaká je mezi anotátory shoda na sémantických typech diskurzních vztahů vyjádřených sekundárními konektory,

6. shoda na rozpoznání „jádra“ daného konektoru (např. slova *podmínka* ve spojení *za podmínky, že*) a

7. shoda na rozpoznání celkového rozsahu sekundárního konektoru (tj. zda anotátoři označili jako konektor např. celé spojení *to byl hlavní důvod, kvůli kterému* nebo pouze *to byl důvod*).

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).

Další hodnoty mezianotátorské shody (např. shoda na typech vztahů) 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 diskurzním korpusu jsou zobrazeny v tabulce 6.2.

Výsledky mezianotátorské shody při anotaci sekundárních konektorů jsme porovnali se shodou na primárních konektorech (převzatou z Polákové a kol., 2013). V rámci anotace konektorů v PDiT 1.0 byla změřena celková shoda na existenci diskurzního vztahu a shoda na určení sémantického typu vztahu, srov. tabulka 6.3.

Z tabulek 6.2 a 6.3 je patrné, že mezianotátorská 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.

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. *podmínka, příčina, důsledek*), které svým lexikálním významem odkazuje k odpov-

²⁴ 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, u 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.

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 konektorů vs. volných frází		100	
4. Shoda na rozpoznání typu argumentu (VP vs. NP)		100	
5. Shoda na sémantických typech vztahů:			
i) celkem		82	0,78
ii) pro verbální argumenty		83	0,76
iii) pro nominální argumenty		80	0,73
6. Shoda na rozpoznání jádra konektoru		78	
7. Shoda na rozsahu konektoru		81	

Tabulka 6.2: Mezianotátorská shoda pro anotaci sekundárních konektorů v PDiT 2.0

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

Tabulka 6.3: Srovnání mezianotátorské shody pro anotaci primárních a sekundárních konektorů v Pražském diskurzním korpusu

vídajícím diskurzním vztahům, např. vztahům podmínky, příčiny, důsledku (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. Například celková shoda pro anotaci na morfologické rovině Pražského závislostního korpusu byla 97 % (srov. Hajič a kol., 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 vypožorovat celkovou tendenci, tj. mezianotátorská shoda „klesá s tím, jak se zvyšuje rovina jazyka“.

6.3 Vyhodnocení anotace sekundárních konektorů v PDiT

V následující části práce popíšeme výsledky vlastní anotace diskurzních konektorů v aktuální verzi korpusu PDiT, zaměřenou na konektory sekundární. Nejprve (kapitola 6.3.1) představíme diskurzní konektory z hlediska frekvence jejich výskytu. Uvedeme rozložení jednotlivých typů konektivních prostředků v celém korpusu (tj. poměr mezi konektory primárními, sekundárními a volnými konektivními frázemi) a zaměříme se na konektory, které se v textech PDiT vyskytují s nejvyšší frekvencí. Posléze srovnáme výskyt sekundárních konektorů v PDiT s výskytem obdobných jazykových prostředků v angličtině v korpusu PDTB a na závěr srovnáme frekvenci vybraných sekundárních konektorů v psané a mluvené češtině (na datech korpusů PDiT a DIA-LOG).

Dále (kapitola 6.3.2) se zaměříme na sémanticko-pragmatické typy vztahů vyjadřované primárními a sekundárními konektory a na syntaktickou realizaci diskurzních argumentů (tj. na to, zda sekundární konektory preferují argumenty slovesné či nominální). Pozornost budeme věnovat také novým otázkám, které při anotaci sekundárních konektorů vyvstaly – jedná se především o začlenění „nových“ významových vztahů (které nejsou zahrnuty v původní klasifikaci v PDiT 1.0), začlenění vztahu juxtapozice, vztahů mezi implicitními argumenty a tematických předělů v textu.

Na závěr (kapitola 6.3.3) se zaměříme na srovnání vztahů mezivětných a vnitrovětných vyjadřovaných primárními a sekundárními konektory.

6.3.1 Diskurzní konektory podle frekvence

Výsledky analýzy nově provedené anotace sekundárních konektorů v Pražském diskurzním korpusu ukázaly, že se vztahy vyjadřované primárními a sekundárními konektory liší v první řadě z hlediska frekvence.

6.3.1.1 Rozložení primárních konektorů, sekundárních konektorů a volných konektivních frází v PDiT

Počet výskytů primárních a sekundárních konektorů v PDiT zachycuje tabulka 6.4.

Diskurzní konektory	Výskyt v PDiT	%
primární	20 255	94,6
sekundární	1 161	5,4
Celkem	21 416	100

Tabulka 6.4: Srovnání výsledků anotace primárních a sekundárních konektorů v PDiT

Z celkového počtu diskurzních vztahů vyjádřených explicitně (tj. pomocí konkrétních jazykových prostředků) tvoří primární konektory 94,6 %, sekundární 5,4 %. Z vý-

sledků je patrné, že označení primární a sekundární je v tomto ohledu příznačné – k vyjadřování diskurzních vztahů v textu slouží zejména (tj. z 94,6 %) primární konektory.

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

Výskyt konektorů v PDiT jsme dále porovnali s výskytem volných konektivních frází (srov. tabulka 6.5).

Typ konektivního prostředku	Výskyt v PDiT
Primární konektory	20 255
Sekundární konektory	1 161
Volné konektivní fráze	151
Celkem	21 567

Tabulka 6.5: Diskurzni konektory a volné konektivní fráze v PDiT

V PDiT bylo zaznamenáno 151 volných konektivních frází. Z výsledků je tedy patrné, že v rámci užívání konektivních prostředků jasně dominují gramatikalizované primární konektory, následují negramatikalizované sekundární konektory a kontextově silně závislé konektivní fráze se používají (alespoň podle dat PDiT) velmi zřídka.

Nová anotace sekundárních konektorů a volných konektivních frází může sloužit nejen k výzkumu konektivních prostředků, ale 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 a volných konektivních frází bude možné snadněji vyhledávat implicitní vztahy podle míst, ve kterých se nevyskytuje žádný explicitní konektivní prostředek, 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).

6.3.1.2 Nejčastěji používané diskurzni konektory

Diskurzni konektory s nejvyšším počtem výskytů v PDiT jsou uvedeny v tabulce 6.6. Tabulka zobrazuje dvacet nejfrekventovanějších konektorů (bez rozlišení na primární a sekundární) – kompletní abecední a frekvenční seznam primárních a sekundárních konektorů zachycených v současné anotaci v PDiT je uveden v příloze.

Z tabulky je patrné, že dvacet nejfrekventovanějších konektorů tvoří pouze primární konektory; výskyty jednotlivých forem primárních konektorů tedy značně převyšují výskyty konektorů sekundárních. Projevuje se zde patrně skutečnost, že pro auto-

Diskurzivní konektory podle četnosti	
<i>a</i>	5 765
<i>však</i>	1 521
<i>ale</i>	1 267
<i>když</i>	574
<i>protože</i>	525
<i>totiž</i>	460
<i>pokud</i>	403
:	395
<i>proto</i>	380
<i>tedy</i>	307
<i>aby</i>	305
<i>pak</i>	295
<i>ovšem</i>	285
<i>-li</i>	248
<i>také</i>	234
<i>neboť</i>	221
–	218
<i>zatímco</i>	204
<i>nebo</i>	191
<i>navíc</i>	182

Tabulka 6.6: Diskurzivní konektory s nejvyšším počtem výskytů v PDiT

ra je snazší používat formálně i funkčně ustálené primární konektory než rozvolněné a dosud ne zcela gramatikalizované konektory sekundární.

Variabilitu sekundárních konektorů můžeme demonstrovat na strukturách obsahujících jádrové slovo *důvod*. Sekundární konektory s tímto jádrem se v PDiT objevují celkem ve 119 výskytech (mezi verbálními i nominálními argumenty). V rámci sekundárních konektorů se zároveň jedná o jádro vyskytující se v nejrozmanitějších strukturách, resp. v nejvyšším počtu různých obecných schémat a jejich lexikálních realizacích (srov. struktury typu *z tohoto důvodu*; *z uvedeného důvodu*; *z toho důvodu*; *to je důvod*, *proč*; *důvodem je*; *z jednoduchého důvodu*; *z důvodu*, *že*; *z toho důvodu*, *aby*). Jednotlivé realizace se přitom v PDiT vyskytují poměrně řídky (maximálně v desítkách výskytů), např. nejčetnější spojení *z tohoto důvodu* se objevuje v 11 výskytech (více ke konektorům obsahujícím jádrové slovo *důvod* viz kapitola 6.1.2.3).

Výskyty některých jádrových slov sekundárních konektorů jsou tedy poměrně srovnatelné s výskytem vysoce frekventovaných primárních konektorů (srov. jádro *důvod* se 119 výskyty a dvacátý nejfrekventovanější konektor *navíc* se 182 výskyty), nicméně konkrétní realizace a varianty sekundárních konektorů samy o sobě již tak frekventované nejsou.

6.3.1.3 Sekundární konektory v češtině a v angličtině

Výsledky anotace sekundárních konektorů jsme dále porovnali s anotací tzv. alternativních lexikálních vyjádření diskurzních konektorů (altlexů) v angličtině, které se blíží sekundárním konektorům. Altlexy jsou zachyceny v korpusu Penn Discourse Treebank (PDTB) obsahujícím přibližně 49 000 vět, tedy srovnatelný počet, jako má Pražský diskurzní korpus. Oba korpusy zároveň obsahují publicistické texty. Cílem bylo zjistit, zda se sekundární konektory objevují v jazycích s podobnou frekvencí, tj. zda dva různé jazyky mají stejný či odlišný repertoár těchto prostředků a zda je využívají podobně často.

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 diskurzní korpus zatím neobsahuje. Relevantnější srovnání s výsledky anotace sekundárních konektorů v češtině tedy poskytne pouze podíl diskurzních konektorů a jejich alternativních vyjádření (altlexů) – srov. tabulka 6.7 (hodnoty pro PDTB byly převzaty z článku Prasadové a kol., 2014). Pojem altlexy se ovšem v PDTB používá pro širší množinu jazykových prostředků než pojem sekundární konektory – v rámci altlexů se nerozlišují kontextově závislé a nezávislé konektivní struktury. Pojem altlexy tedy v naší terminologii zahrnuje sekundární konektory a volné konektivní fráze zároveň.

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

Tabulka 6.7: Srovnání anotace diskurzních vztahů v češtině a angličtině

Penn Discourse Treebank obsahuje celkem 18 459 vztahů vyjádřených (primárními) diskurzními konektory a 624 vztahů vyjádřených altlexy, tj. altlexy tvoří 3,3 % všech explicitních konektivních prostředků, zatímco v Pražském diskurzním 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 diskurzním korpusu a Penn Discourse Treebanku velmi podobné.

6.3.1.4 Vybrané sekundární konektory v psané a mluvené češtině

V rámci studia frekvence sekundárních konektorů jsme dále provedli dílčí průzkum zaměřený na výskyt a chování těchto jazykových prostředků v psané a mluvené češtině. Pro srovnání s psanými daty Pražského diskurzního korpusu jsme využili data korpusu mluvené češtiny DIALOG. Oba korpusy obsahují srovnatelný počet textových slov (DIALOG 932 373, PDiT 833 195), jejich texty pocházejí z oblasti publicistiky.

Jako zástupce sekundárních konektorů jsme vybrali struktury vyjadřující vztah příčiny a důsledku (nejfrekventovanější typ vztahu u sekundárních konektorů v PDiT), které obsahují různorodá jádrová slova: *důvod*, *příčina*, *zdůvodnit* / *zdůvodňovat*, *souviset*.²⁵ 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 67 z korpusu DIALOG) a v kolika procentech bez konektivní platnosti (srov. příklad 68 z korpusu DIALOG).

- (67) *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.*
- (68) [Na Šumavě, M. R.] *opravdu hrozilo, že by žáci nemohli dojždět do jiné školy z důvodu povětrnostních podmínek.*

Z tabulky 6.8 je patrné, že vyšší počet zkoumaných struktur s konektivní funkcí se vyskytuje v PDiT, tj. v psaných textech (konektivní funkci plní 49 % z celkových 516 výskytů).

Struktury obsahující slovo	S konektivní funkcí		%		Celkem	
	PDiT	DIALOG	PDiT	DIALOG	PDiT	DIALOG
<i>důvod</i>	180	119	55 %	28 %	330	425
<i>příčina</i>	32	14	36 %	21 %	89	66
<i>zdůvodnit/-ovat</i>	20	5	51 %	15 %	39	34
<i>souviset</i>	20	42	34 %	28 %	58	148
Celkem	252	180	49 %	27 %	516	673

Tabulka 6.8: Struktury obsahující vybraná jádrová slova v PDiT a korpusu DIALOG

U všech vybraných slov vyskytujících se v konektivních strukturách jsme dále provedli podrobnější analýzu. Zkoumali jsme, zda mají dané struktury funkci sekundárního konektoru či volné konektivní fráze, a zároveň to, zda vyjadřují diskurzní vztah

²⁵ Korpus DIALOG neobsahuje anotaci diskurzních konektorů, vybrané struktury byly pro naše účely zpracovány manuálně.

mezi dvěma slovesnými argumenty (v tabulce jsou označeny jako „VP_VP“) či zda je druhý argument nominální (v tabulce „VP_NP“) – srov. tabulka 6.9.

Typ struktury	VP_VP		VP_NP		Celkem	
	PDiT	DIALOG	PDiT	DIALOG	PDiT	DIALOG
Sekundární konektory	84	122	72	51	156 62 %	173 96 %
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 6.9: Shrnující výsledky pro vybrané konektivní struktury v korpusech PDiT a DIALOG

Z porovnání syntaktické realizace diskurzních argumentů vyplývá, že korpus DIALOG obsahuje 69 % vybraných struktur vyjadřujících diskurzní vztah mezi dvěma slovesnými argumenty a 31 % vztahů s nominálním argumentem. V korpusu PDiT je situace vyrovnanější – mírně (v 54 %) dokonce převažují vztahy nominální. Z výsledků je tedy patrné, že v psané češtině je u daných struktur vyšší tendence k nominalizaci (či kondenzaci) diskurzních argumentů než v mluvené češtině.

Z analýzy dále vyplývá, že korpus DIALOG obsahuje (v rámci zkoumaných výrazů) téměř výhradně kontextově nezávislé konektivní struktury, tedy sekundární konektory (srov. 96 % sekundárních konektorů vs. 4 % volných konektivních frází). V korpusu PDiT je poměr zkoumaných struktur opět více vyvážený, i když i zde zřetelně převažují sekundární konektory nad volnými konektivními frázemi (srov. 62 % vs. 38 %). Výsledky tedy ukázaly, že v psané češtině se vyskytují struktury typu *důvodem pro propuštění bylo; důvodem poklesu je; toto počínání je důvodem, proč* častěji než v textech mluvené češtiny, ve kterých se na místě výrazu anaforicky odkazujícího k předchozímu kontextu typicky (u struktur se slovem *důvod* dokonce v 98 %, tj. ve 117 výskytech z celkových 119) užívá zájmeno *to* – srov. *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.

Lexikálně rozmanitější jsou v PDiT 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*, tj. struktury se slovesem *být*, PDiT obsahuje i struktury s jinými typy sloves – např. *vidět příčinu v tom, že; hledat příčinu v tom, že; označit za příčinu nebo příčiny tkví v tom, že*.

Vybrané konektivní struktury se v korpusech PDiT a DIALOG liší dále 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 PDiT vyjadřují pestřejší škálu významových vztahů.

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 6.8) v korpusu DIALOG se v této funkci objevilo 38 příkladů (tj. 90 %), v PDiT pouze 5 (tj. 25 % z celkového počtu 20 výskytů).

Zbylé výskyty (v PDiT 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. příklad 69 z PDiT:

- (69) *Za samozřejmost lze už považovat to, že většina vystavovatelů nabízela jak tuzemské, tak i zahraniční značky v nejrůznějším 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 uvedeném příkladě se objevuje sekundární konektor *souvisí to se* vyjadřující diskurzní vztah mezi prvním slovesným a druhým jmenným argumentem. Význam diskurzního vztahu se blíží vztahu příčiny (srov. parafrázi *většina vystavovatelů nabízela jak tuzemské, tak i zahraniční značky v nejrůznějším provedení..., protože se snaží získat všechny potencionální zákazníky...*).

Struktura *souvisí to s* tedy nutně nemusí sloužit k prostému navázání dvou různých témat (v dalších příkladech z PDiT se kromě příčiny vztahy blíží např. také explikaci). V PDiT je tento typ dokonce výrazně častější (tvoří 75 % z celkového počtu vztahů vyjádřených konektivními strukturami se základem *souviset*), v korpusu DIALOG se naopak vyskytuje zřídka (pouze v 10 %).

Na základě analýzy vybraných konektivních struktur v Pražském diskurzním korpusu (obsahujícím psané texty) a v korpusu mluvené češtiny DIALOG můžeme vyvodit následující obecnější závěry. Vyšší počet vybraných konektivních struktur (typu *z tohoto důvodu, příčinou bylo*) se vyskytuje v PDiT, tj. v psaných textech. Vysvětlení může spočívat v tom, že některé konektivní struktury typu *své chování odůvodnil tím, že* mohou působit formálněji (než například neutrální konektory, jako jsou *proto, protože* aj.), a proto se objevují častěji v psaných textech. V psané češtině je zároveň obecně vyšší tendence k nominalizaci (či kondenzaci) druhého diskurzního argumentu než v češtině mluvené.

V psané i mluvené češtině se častěji vyskytují sekundární konektory než volné konektivní fráze. V mluvené češtině je přitom rozdíl v počtu sekundárních konektorů a volných frází zcela jednoznačný, v psaných textech je poměr vyrovnanější. 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.

Také spektrum konektivních struktur je v psaných textech rozmanitější než v textech mluvených (srov. v psaných textech se objevující struktury typu *vidět příčinu v tom, že; hledat příčinu v tom, že; označit za příčinu*).

Sekundární konektory se v psané a mluvené češtině odlišují také z hlediska významových vztahů, které vyjadřují. Např. konektivní struktury obsahující jádro *souviset* (typu *s tím souvisí, to souvisí s*) se v mluvené češtině používají především k prostému, volnějšmu navazování dvou různých témat, přičemž význam daného diskurzního vztahu je spíše vágní, zamlžený (srov. *s tím souvisí i druhá věc, to souvisí s celou řadou dalších otázek, to souvisí ještě s jednou věcí, s tím souvisí třetí problém*).²⁶ Na druhou stranu v psané češtině se struktury typu *s tím souvisí* objevují většinou v „pevnějších“, významově zřetelnějších vztazích. Nevyjadřují tedy prosté navazování, ale mohou se významově blížit např. vztahům příčiny či explikace.

Podrobněji se významovým vztahům vyjadřovaným diskurzními konektory (primárními i sekundárními) věnujeme v následující kapitole.

6.3.2 Sémantické typy diskurzních vztahů vyjádřené primárními a sekundárními konektory

Srovnání toho, jaké sémantické typy diskurzních vztahů vyjadřují primární a sekundární konektory,²⁷ zobrazují tabulky 6.10 a 6.11. Tabulka 6.11 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 druhý argument nominální (v tabulce jako „VP_NP“).

Z tabulek je patrné, že rozložení jednotlivých typů diskurzních vztahů²⁸ je u primárních a sekundárních konektorů velice podobné. Diskurzní vztahy s vysokým počtem výskytů (u primárních i sekundárních konektorů) jsou vztahy konjunkce, pří-

²⁶ Někdy tedy nemusí být konektory užity pro vyjádření „pevnějšího“ významového vztahu, ale převládá u nich funkce pragmatická. Tato skutečnost souvisí s intenční vágností v přirozeném jazyce, viz k tomu poznámka J. Hoffmanové v *Mluvnici češtiny* 3 (1987: 706) o tom, že členící signály „spíše než vlastní funkci konektivní mají nezdědka funkci kontaktní nebo fungují jako výrazy „vycpávkové“ a umožňují podavateli promyslet si pokračování...“.

²⁷ Frekvenční seznam forem primárních a sekundárních konektorů z Pražského diskurzního korpusu s ohledem na vyjadřované významové vztahy je uveden v příloze.

²⁸ Pro jednotlivé typy diskurzních vztahů jsou v PDIT užity následující zkratky: *synchron* – 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* – explikace, *confr* – konfrontace, *opp* – opozice, *f_opp* – nepravý kontrast, *restr* – restriktivní opozice, *conc* – připustka, *corr* – rektifikace, *grad* – gradace, *conj* – konjunkce, *exempl* – exemplifikace, *spec* – specifikace, *equiv* – ekvivalence, *gener* – generalizace, *conjalt* – konjunktivní alternativa, *disjalt* – disjunktivní alternativa, *other* – jiný typ vztahu.

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
Explicace	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 6.10: Typy diskurzních vztahů vyjádřené primárními konektory v PDiT

činy a důsledku,²⁹ 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 nepravého kontrastu.

Primární a sekundární konektory se naopak liší ve vyjadřování vztahů opozice a explicace. 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 explicace je u sekundárních konektorů čtvrtým nejpočetnějším typem diskurz-

²⁹ Jak jsme popsali v kapitole 3.6.2.1, v PDiT 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.

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
Explikace***	67	52	78	15	22
Přípustka***	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 6.11: Typy diskurzních vztahů vyjádřené sekundárními konektory v PDiT; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p -hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$

ního vztahu, zatímco u konektorů primárních je až v druhé polovině tabulky (tj. mezi vztahy s nižším počtem výskytů).

Nižší čísla výskytů pragmatických vztahů podle našeho názoru neznamenaají, že by se tyto vztahy v textech vyskytovaly s tak nízkou frekvencí, ale spíše to, že jim v anotaci PDiT 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 – tj. postup 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* nebo *argumentovat* – srov. příklad 70:

- (70) *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 PDiT není u výrazů typu *jinak řečeno, jinými slovy, cynicky řečeno* 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*). V PDiT tedy v současné době převládá zpracování sémantických diskurzních vztahů nad vztahy pragmatickými. Pragmatické aspekty koherence textu bude zapotřebí v PDiT propracovat detailněji (a to v rámci primárních i sekundárních konektorů).

V rámci analýzy sémantických typů diskurzních vztahů v Pražském diskurzním korpusu jsme dále zjišťovali, jaké je u jednotlivých typů vztahů procentuální zastoupení primárních a sekundárních konektorů. Výsledky zachycuje tabulka 6.12.

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.³⁰

³⁰ Jako jiný typ vztahu byly zachyceny případy, ve kterých byl konektor (primární či sekundární) užít na nepatřičném místě, tj. významově se na dané místo nehodí – srov. např. *Je nebezpečné dělat rovnítko mezi bezdomovci a kriminálními živy, 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.*

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
Nepravý kontrast	50	100	0
Nepravá příčina a důsledek	44	93,2	6,8
Nepravá podmínka	17	94,1	5,9
Jiný vztah	3	66,7	33,3
Celkem	21 416	94,6	5,4

Tabulka 6.12: Typy diskurzních vztahů vyjádřené primárními a sekundárními konektory v Pražském diskurzním korpusu – procentuální vyjádření

6.3.2.1 Sekundární konektory a syntaktická realizace argumentů

V rámci sekundárních konektorů jsme dále zjišťovali, zda typ diskurzního vztahu souvisí se syntaktickou povahou diskurzních argumentů (tj. s tím, zda jsou argumenty vyjádřeny slovesně či jmenně).

Celkový počet výskytů sekundárních konektorů ve spojení s nominálním diskurzním argumentem je 237. Z hlediska syntaktické povahy diskurzních argumentů tedy celkově převažují u sekundárních konektorů vztahy mezi dvěma slovesnými argumenty, a to v 80 % výskytů (srov. 924 výskytů z celkových 1 161). Na výsledku se

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.

ovšem odráží skutečnost, že ne všechny konektivní struktury umožňují nominalizaci druhého argumentu (srov. předložkové fráze typu *díky tomu*, některá ustálená spojení typu *stručně / jednoduše / obecně... řečeno*). K nominalizaci druhého diskurzního argumentu tedy tíhnou pouze určité struktury.

Je zajímavé, že 136 případů nominálních argumentů z celkových 237 (tedy více než polovina) vyjadřuje vztah příčiny a důsledku. Počet výskytů tohoto typu vztahu ve spojení s nominálním argumentem tak značně převyšuje ostatní (druhým nejpočetnějším typem vztahu ve spojení s nominálním argumentem je exemplifikace s 21 výskyty). Vysvětlení může spočívat v tom, že vztah příčiny a důsledku je v PDiT chápán v širším smyslu, tj. nerozlišují se zde podtypy, jako je např. příčina, důvod nebo důsledek. 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).

Z výsledků dále vyplynulo (viz tabulka 6.11), že téměř u všech typů vztahů vyjádřených sekundárními konektory převažují slovesné argumenty nad nominálními – u vztahu konjunkce a vztahu podmínky (tj. dvou ze tří nejpočetnějších vztahů) slovesné argumenty převažují dokonce v 96 % a v 92 %.

Vztahy s nominálním argumentem převažují pouze ve dvou případech – ve vztahu restriktivní opozice (79 %) a exemplifikace (70 %). Nejčastějšími konektory pro tyto vztahy jsou struktury *výjimkou je* (pro restriktivní opozici) a *příkladem je* (pro vztah exemplifikace), srov. příklad 71 z PDiT:

- (71) *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.*

Syntaktická realizace druhého argumentu ovšem nemusí souviset ani tak s významovými typy diskurzních vztahů jako s konkrétními sekundárními konektory. Výskyt nominálního argumentu tedy nemusí být preferován konkrétním významovým vztahem, ale konkrétní konektivní strukturou, která ho vyjadřuje (např. *výjimkou je, příkladem je*).

Dále jsme proto v PDiT vyhledali další sekundární konektory (zachycené v současné anotaci) analogické ke konektorům *výjimkou je, příkladem je*, tj. 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 6.13.

Z tabulky³¹ je patrné, že zkoumané typy sekundárních konektorů většinou upřednostňují nominalizaci druhého diskurzního argumentu (tj. např. *důsledkem je stárnutí*

³¹ Tabulka zachycuje základní realizace daných konektorů, které se v PDiT objevily v různých gramatických variantách – např. *výjimkou je, výjimkou byl, výjimkou jsou*. Celkově jsou ovšem počty nalezených struktur v PDiT poměrně nízké.

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 6.13: Sekundární konektory typu *výjimkou je* v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p -hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$

obyvatelstva před důsledkem je, že obyvatelstvo stárne), a to celkově v 80 % výskytu. Tato malá sonda ukazuje, že struktura sekundárních konektorů může mít přímý vliv na syntaktickou realizaci druhého diskurzního argumentu.

6.3.2.2 Rozšíření klasifikace diskurzních vztahů

Jak jsme popsali v kapitole 3.6.2.1, klasifikace sémantických typů diskurzních vztahů pro sekundární konektory byla převzata z anotace primárních konektorů (vzhledem k tomu, že náš výzkum proběhl na datech Pražského diskurzního korpusu, který anotaci primárních konektorů již obsahoval). V poznámkách jsme zároveň značili „nové“ typy vztahů (tj. takové typy, které nebyly v nabídce původní klasifikace diskurzních vztahů pro anotaci primárních konektorů). Při zpracovávání textů jsme zároveň narazili i na další teoretické otázky, které je zapotřebí v rámci problematiky diskurzu dále řešit – jedná se o rozdíl, resp. vztah mezi konjunkcí, juxtapozicí a implicitním významovým vztahem, dále o začlenění diskurzních vztahů mezi nevyjádřenými (implicitními) argumenty a o vztahy mezi tematickými předěly.

Vztah vyplývání (či vyvození výsledků)

Prvním z dosud neznačených významových vztahů je vztah vyplývání (vyvození výsledků), který byl v datech PDiT vyjádřen následujícími sekundárními konektory: *z toho vyplývá, že; v každém případě; výsledkem je*, srov. příklad 72 z PDiT:

- (72) *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 do-*

běhne 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 nezcizitelného 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).

Vyvození závěru

Dalším typem nového vztahu bylo vyvození závěru – jednalo se o příklady se sekundárními konektory typu *závěr zněl, že; došel k závěru, že; jednoznačný závěr je, že; dospělo k tomuto závěru, srov. příklad 73.*

- (73) *Pokud zavoláte na zveřejněné telefonní číslo, operátorka vám vysvětlí, co je třeba udělat, aby od vás firma akcie skutečně odkoupila. Závěr je následující, pokud si vyberete k prodeji svých akcií nějakou firmu nabízející peníze rychle a hotově, je v tom obvykle ještě nějaký háček.*

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

Zřetel

Nově byl mezi typy diskurzních vztahů, které se objevují v PDiT, zařazen také vztah zřetele. Může být vyjádřen například sekundárními konektory *v tomto směru, se zřetelem k tomu, srov. příklad 74.*

- (74) *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 PDiT bude stávající klasifikace o dané tři typy významových vztahů, tj. vyplývání (či vyvození výsledku), vyvození závěru a zřetel, rozšířena. Nově zavedené vztahy nemusejí být zároveň nutně vyjadřovány pouze sekundárními konektory (vztah vyplývání může být jistě vyjádřen např. primárním konektorem *tak*). Z této skutečnosti je nicméně zřejmé, že lexikálně konkrétnější struktury sekundárních konektorů umí vyjadřovat/pojmenovávat nuance významových vztahů lépe, resp. zřetelněji než konektory primární (srov. obecnější vyjádření *tak* vs. konkrétnější *výsledkem je*). S ohledem na tuto skutečnost by v budoucnu mohla být zrevidována i anotace primárních konektorů.

Konjunkce, juxtapozice a implicitní vztahy

Při anotaci sekundárních konektorů v PDiT jsme se setkali s dalšími skutečnostmi, které ve stávající anotaci diskurzních vztahů zatím nebyly zohledněny a jejichž začlenění by v budoucnu mohlo sloužit k vylepšení anotace, resp. větší komplexnosti anotovaných dat – jednalo se o vztah juxtapozice (ve vztahu k implicitním vztahům), vztahy mezi implicitními argumenty a vztahy mezi tematickými předěly.

Při čtení textů jsme se často potýkali s otázkami, kde jsou hranice mezi implicitností a explicitností či ve kterých případech se v textu jedná o implicitní významový vztah a ve kterých o prosté kladení vět vedle sebe (tzv. juxtapozici). Jako nejproblematičtější se v tomto ohledu jevil vztah konjunkce, resp. odlišení konjunkce a juxtapozice.

P. Karlík a H. G. Škrabalová (2017) charakterizují konjunkci (slučovací vztah) jako „jeden ze základních významových (sémanticko-logických) vztahů mezi koordinovanými složkami. [...] Vyjadřuje, že koordinované složky jsou sémanticky rovnocenné a vykazují nějakou sounáležitost. [...] Je typicky realizován(a) koordinací s koordinací spojkou a [...] n. parataxí (asyndeton), která zvyšuje efekt rovnocennosti.“

Juxtapozice bývá chápána různě. J. Bauer a M. Grepl (1980) pod tímto pojmem rozumějí asyndetické spojení vět, jejichž obsahy nelze interpretovat jako vnitřně spolu související (jsou mezi sebou spojeny jen tematicky, patří do stejného jazykového projevu), srov. následující převzatý příklad.

(75) *Jděte mi k šípku, kdopak by do takové houpačky sedal, to je tak pro vás.*

Naproti tomu J. Hrbáček (1991) chápe juxtapozici širěji, nejen jako nulový vztah, ale obecně jako vztah nevyjádřený spojovacím prostředkem. Jako juxtapozici tedy označuje i implicitní významové vztahy.

„[...] Rozlišit, kdy v textu jde o vypuštění (elipsu) spojovacího prostředku, a tedy o asyndeton, a kdy jde o postavení vět vedle sebe s nulovým vztahem mezi nimi, tedy o juxtapozici, je velmi obtížné a v mnoha případech jednoznačně nemožné.“ (J. Hrbáček, 1991)

Vyvození či nevyvození konkrétního významového vztahu mezi propozicemi pak zůstává na čtenáři, srov. následující příklad převzatý od P. Karlíka (2017).

(76) *Vznikla světová socialistická soustava. Zanikly mnohé civilizace.*

Při anotaci diskurzních vztahů v korpusu navrhuje implicitní významové vztahy a juxtapozici (v užším smyslu) rozlišovat, i přesto, že jejich rozpoznání je v praxi obtížné a do jisté míry vždy závisí na interpretaci čtenáře.

Juxtapozici přitom rozumíme významově vyprázdněný vztah mezi dvěma propozicemi v rámci jednoho tematického celku. Ačkoliv mezi těmito diskurzními argumenty není bližší významový vztah, jejich řazení vedle sebe neporušuje koherenci textu. Koherence textu totiž není zajišťována pouze významovými mezipropozičními vztahy, ale i vztahy jiného typu, jako je koreference, anaforické a kataforické odkazování v rámci textu nebo vztahy akutálního členění, resp. informační struktury.

Základním kritériem pro rozlišení juxtapozice a implicitních sémanticko-pragmatických textových vztahů je přitom (ne)možnost vložení diskurzního konektoru (primárního či sekundárního) mezi diskurzní argumenty. Je-li možné mezi argumenty vložit diskurzní konektor, jedná se o implicitní vztah. V opačném případě jde o juxtapozici. Diskurzní konektory totiž ve své podstatě nejsou významově vyprázdněné (například *protože* vyjadřuje typicky příčinu, *ale* nejčastěji opozici). Z tohoto důvodu nelze konektory vložit tam, kde významový vztah mezi argumenty vyprázdněný (juxtapoziční) je. Problematická je v tomto ohledu spojka *a*. Její význam není silně vyhraněný – objevuje se mezi argumenty se vztahem konjunkce, příčiny a důsledku, specifikace, gradace, konfrontace, opozice, předčasnosti, následnosti i současnosti. Význam konektoru *a* je tak široký, že jej lze potenciálně vložit do velmi rozmanitého a rozsáhlého spektra kontextů. Tento konektor tedy chápeme jako spojovací můstek mezi implicitními významovými vztahy (zejména konjunkcí) a vztahy juxtapozice. Ostrá hranice mezi těmito dvěma typy vztahů neexistuje.

Vydělujeme tedy v obecné rovině následující typy mezipropozičních vztahů: vztahy explicitní (vyjádřené diskurzními konektory), vztahy implicitní (jejichž význam může být vyjádřen konektorem potenciálně) a vztahy juxtapozice, které jsou významově vyprázdněné do té míry, že je nevyjadřují diskurzní konektory. Zavedení těchto skupin vztahů do anotace diskurzního korpusu by mohlo přinést nové informace o tvorbě textu.

Implicitní vztahy vs. implicitní argumenty

Od problematiky implicitních vztahů se dostáváme k existenci implicitních diskurzních argumentů. Jak jsme popsali v kapitole 3.4, argumenty mohou být situační či textové. V PDiT jsou značeny argumenty textové, resp. vždy textově vyjádřené. Při rozboru dalších textů jsme si ovšem všimli, že diskurzní významový vztah může někdy vést i mezi dvěma argumenty, z nichž je jeden textově nevyjádřený (tj. je pouze implicitní), srov. příklad 77.

(77) *Kdy a kde bude další koncert? Abych včas nahlásila volno v práci.*

Příklad 77 můžeme interpretovat následujícím způsobem: *Kdy a kde bude další koncert?* Ptám se (impl. Arg), *abych včas nahlásila volno v práci.* Hlavní věta Arg1 (*ptám se*) není v příkladu explicitně vyjádřena. Určitá míra implicitnosti je přitom v přirozeném jazyce běžná, resp. dokonce nutná. Čtenář i posluchač jsou zvyklí si do textu vnášet informace, které v něm nejsou explicitně vyjádřeny, a proto implicitní struktura z příkladu 77 bez problémů rozumějí. Užívání implicitních argumentů tedy nenarušuje přirozenou plynulost a koherenci textu.

V naší práci tedy navrhujeme dělit argumenty na situační a textové a textové dále na argumenty explicitní a implicitní (výskyt jednotlivých typů argumentů přitom závisí na daném typu textu, situační a implicitní argumenty lze například očekávat

spíše v mluveném projevu) a všechny tyto typy zahrnout do komplexní anotace diskurzních vztahů.

Tematické předěly v textu

Při anotaci diskurzních vztahů v PDiT jsme dále nacházeli jazykové prostředky typu *dále, mimochodem, další věcí je, odjinud, jinak*, které nevyjadřovaly užší významový vztah mezi dvěma jednotkami textu, ale které signalizovaly změnu tématu v textu, srov. příklad 78.

- (78) *S diplomkou není zatím nic nového. **Jinak** já přijedu až v sedm, takže s večerí na mě nečekejte.*

Tematické předěly v textu mohou být i implicitní – mohou být naznačeny jen novým odstavcem nebo v mluveném projevu pomlkou.

Různé tematické celky mohou být navíc ještě propojeny dalšími spojovacími prostředky, které se tak podílejí na vyšší hierarchizaci textu, srov. příklad 79. V daném příkladě se v jedné větě vyskytují výrazy *jinak a také*; *jinak* signalizuje změnu (pod)tématu (mluvčí přestává hovořit o tom, že dílo F. Liszta existuje v mnoha verzích), *také* odkazuje k hlouběji zanořené otázce *Transkribujete i dílo Franze Liszta?* Diskurzní konektory tedy mohou mít dosah i přes jiné tematické celky (či jejich části).

- (79) *Zahrála jste také Patetickou sonátu od Beethovena, a to ve vaší úpravě na varhany. Transkribujete i dílo Franze Liszta? Liszt je známý tím, že existuje v mnoha verzích. Například Orfeus existuje jako symfonická báseň, ale Liszt to také napsal pro dva klavíry nebo varhany. Takže je to úplně legitimní Liszt, protože to dělal sám. **Jinak** já jsem **také** několik skladeb transkribovala. Jeho hudba je tak krásná a kompaktní, že se dá dobře ztvárnit, možná ještě lépe než klavírně.*

Vztahy juxtapozice (včetně významových vztahů implicitních), vztahy mezi implicitními argumenty a tematické předěly nejsou v textech PDiT zatím značeny. Jejich anotace by mohla být do jedné z příštích verzí tohoto korpusu zahrnuta, aby byla diskurzní anotace pokud možno co nejkomplexnější.

6.3.3 Diskurzní vztahy mezivětné a vnitrovětné

Jak již bylo řečeno výše, Pražský diskurzní korpus obsahuje anotaci diskurzních vztahů mezivětných i vnitrovětných (např. *Měl jsem jet na výlet, ale začalo pršet.* i *Měl jsem jet na výlet. Začalo však pršet.*).³²

V další fázi výzkumu jsme proto zkoumali, jaký je poměr mezivětných a vnitrovětných vztahů vyjádřených diskurzními konektory, a to s ohledem na sémantické typy diskurzních vztahů. Nejprve jsme se zaměřili na primární konektory, viz tabulka 6.14 (sémantické typy diskurzních vztahů jsou zde seřazeny podle četnosti).

³² Pod pojmy *mezivětný* a *vnitrovětný* rozumíme užití konektorů mezi větnými celky a uvnitř větných celků.

Z tabulky je patrné, že u diskurzních vztahů vyjádřených primárními konektory obecně převažují vztahy vnitrovětné nad mezivětnými (srov. 70 % vs. 30 %). U některých sémantických typů (srov. vztahy podmínky, účelu, nepravé 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	Mezivětné	%	Vnitrovětné	%	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
Expplikace*	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
Nepravý kontrast	27	54	23	46	50
Nepravá příčina a důsledek**	28	68	13	32	41
Nepravá podmínka***	1	6	15	94	16
Jiný vztah	1	50	1	50	2
Celkem***	6 060	30	14 195	70	20 255

Tabulka 6.14: Podíl mezivětných a vnitrovětných vztahů vyjádřených primárními konektory v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p -hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$

Při zkoumání mezivětných a vnitrovětných vztahů je nicméně zapotřebí zaměřit se blíže na konkrétní konektory, které dané vztahy vyjadřují. Např. vztah podmínky (který je podle PDiT téměř výhradně, z 99 %, vnitrovětný) je nejčastěji vyjadřován

konektory *pokud*, *-li* a *když*, tj. podřadicími spojky. Obdobně vztah účelu je téměř výhradně vyjadřován podřadicí spojkou *aby* (viz příloha). Na druhou stranu vztah generalizace (který je podle PDiT v 91 % mezivětný) je nejčastěji vyjadřován souřadícími konektory *tedy* a *tak*. Vztah opozice, který je v tomto ohledu méně vyhraněný (v 56 % výskytů preferuje mezivětnost), je nejčastěji signalizovaný souřadícími konektory *však* a *ale*.

Ukazuje se, že mezivětnost a vnitrovětnost nejsou vázány na sémantický typ daného vztahu, ale spíše na konkrétní konektory, které je vyjadřují, viz tabulka 6.15.³³

Nejfrekventovanější primární konektory	Celkem	Vnitrovětné užití	Mezivětné užití
	výskyt v PDiT	výskyt (%)	výskyt (%)
<i>a</i> ***	5 766	5 428 (94 %)	338 (6 %)
<i>však</i> ***	1 521	266 (17 %)	1 255 (83 %)
<i>ale</i> ***	1 267	847 (66 %)	420 (34 %)
<i>když</i> ***	574	574 (100 %)	0
<i>protože</i> ***	525	518 (98 %)	7 (2 %)
<i>totiž</i> ***	460	24 (5 %)	436 (95 %)
<i>pokud</i> ***	403	403 (100 %)	0
<i>proto</i> ***	380	41 (10 %)	339 (90 %)
<i>tedy</i> ***	307	33 (10 %)	274 (90 %)
<i>aby</i> ***	305	304 (100 %)	1 (0 %)

Tabulka 6.15: Mezivětnost vs. vnitrovětnost primárních konektorů v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$)

Tabulka 6.15 zachycuje deset nejfrekventovanějších primárních konektorů v PDiT.³⁴ Konektory ve formě podřadicích spojek nepřekvapivě preferují vnitrovětné užití, a to téměř ve 100 % výskytů (srov. *když*, *protože*, *pokud*, *aby*). Zajímavá je situace u konektorů v souřadných konstrukcích, některé z nich (*a*, *ale*) preferují rovněž vnitrovětné užití, jiné (*však*, *totiž*, *proto*, *tedy*) se častěji objevují mezivětně.

V následující části práce se proto blíže zaměříme na primární konektory vyskytující se v souřadných konstrukcích, resp. na to, zda se častěji užívají ve vnitrovětných či mezivětných diskurzních vztazích, srov. tabulka 6.16 a 6.17.

Z tabulek 6.16 a 6.17 vyplývá, že vnitrovětné užití preferují kromě konektorů ve formě podřadicích spojek (*protože*, *když*, *pokud*...) také některé konektory vyskytující se

³³ K vnitrovětnosti a mezivětnosti jednotlivých konektorů viz také Rysová a Rysová (2018).

³⁴ Mezivětným a vnitrovětným užitím konektorů v češtině se podrobně zabývá Jínová (2012) a Zikánová a kol. (2015a). Ve svých studiích docházejí k závěru, že podřadicí spojky jsou v češtině užívány převážně vnitrovětně, zatímco příslovce preferují užití mezivětné. Souřadící spojky nejsou podle jejich zjištění v tomto ohledu vyhraněné.

Primární konektory preferující vnitrovětné užití	Celkem výskyt v PDiT	Vnitrovětné užití výskyt (%)	Mezivětné užití výskyt (%)
<i>a</i> ***	5 766	5 428 (94 %)	338 (6 %)
<i>ale</i> ***	1 267	847 (66 %)	420 (34 %)
<i>neboť</i> ***	221	220 (99 %)	1 (1 %)
<i>nebo</i> ***	191	167 (87 %)	24 (13 %)
<i>sice_ale</i> ***	165	162 (98 %)	3 (2 %)
<i>a tak</i> ***	141	97 (68 %)	44 (32 %)
<i>a to</i> ***	118	97 (82 %)	21 (18 %)
<i>#neg_ale</i> ***	98	96 (97 %)	2 (3 %)
<i>a proto</i> ***	86	86 (100 %)	0
<i>či</i> ***	86	86 (100 %)	0
<i>avšak</i>	61	36 (59 %)	25 (41 %)
<i>a pak</i> ***	56	50 (89 %)	6 (11 %)
<i>a také</i> ***	51	46 (90 %)	5 (10 %)
<i>nýbrž</i> ***	12	12 (100 %)	0

Tabulka 6.16: Nejfrekventovanější primární konektory vyskytující se v souřadných konstrukcích v PDiT preferující vnitrovětné užití; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$)

v souřadných konstrukcích, a to ve formě spojek neadverbiálního původu (srov. *a*, *nebo*, *či*), příp. výrazů, které se s těmito spojkami pojí (srov. složené konektory typu *a tak*, *a to*, *a proto*, *a pak* či plně lexikalizované výrazy typu *ale*). Spojka neadverbiálního původu tedy může měnit adverbialní konektor z mezivětného na vnitrovětný (srov. převážně mezivětné *pak* a převážně vnitrovětné *a pak*). Vnitrovětně jsou dále užívány korelativní páry (*sice_ale*) či výrazy obsahující negaci (*nebo*, *neboť*, *#neg_ale*, *nýbrž*). Naopak mezivětné užití preferují konektory adverbialního původu (*avšak*, *proto*, *tedy*, *tak*).

Dále jsme se v našem výzkumu zaměřili na mezivětné a vnitrovětné vztahy vyjádřené sekundárními konektory (srov. tabulka 6.18 zachycující jednotlivé sémantické typy vztahů).

Z tabulky je zřejmé, že u sekundárních konektorů je situace na první pohled opačná než u konektorů primárních, tj. celkově u nich převažují mezivětně realizované diskurzivní vztahy nad vnitrovětnými (srov. 63 % vs. 37 %; primární konektory byly naopak v 70 % užity vnitrovětně, viz tabulka 6.14).

Dále jsme se podívali na konkrétní sekundární konektory, které jednotlivé vztahy vyjadřují. Například u třetího nejčastějšího vztahu vyjádřeného sekundárními konektory, u vztahu podmínky, dominuje (v 71 %) vnitrovětné užití. Tento vztah je nejčastěji vyjadřován podřadivými strukturami typu *v případě, že; za podmínky, že; v případě, kdy-*

Primární konektory preferující mezivětné užití	Celkem výskyt v PDiT	Vnitrovětné užití výskyt (%)	Mezivětné užití výskyt (%)
<i>však***</i>	1 521	266 (17 %)	1 255 (83 %)
<i>totiž***</i>	460	24 (5 %)	436 (95 %)
<i>proto***</i>	380	41 (10 %)	339 (90 %)
<i>tedy***</i>	307	33 (10 %)	274 (90 %)
<i>pak***</i>	295	78 (26 %)	217 (74 %)
<i>ovšem***</i>	285	64 (22 %)	221 (78 %)
<i>také***</i>	234	9 (3 %)	225 (97 %)
<i>navíc***</i>	182	26 (14 %)	156 (86 %)
<i>přítom***</i>	181	6 (3 %)	175 (97 %)
<i>naopak***</i>	152	29 (19 %)	123 (81 %)
<i>dále***</i>	117	6 (5 %)	111 (95 %)
<i>tak***</i>	112	18 (16 %)	94 (84 %)
<i>rovněž***</i>	106	6 (5 %)	100 (95 %)
<i>přesto***</i>	99	14 (14 %)	85 (86 %)
<i>například***</i>	97	9 (9 %)	88 (91 %)
<i>zároveň***</i>	94	12 (12 %)	82 (88 %)

Tabulka 6.17: Nejfrekventovanější primární konektory vyskytující se v souřadných konstrukcích PDiT preferující mezivětné užití; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p -hodnota $\leq 0,001$)

by; v případech, kdy. Není tedy překvapivé, že je tento typ vztahu vyjadřován spíše vnitrovětně, ačkoli v absolutních číslech převažují u sekundárních konektorů vztahy mezivětné. Z výsledků našeho výzkumu tedy vyplývá, že podobně jako u primárních konektorů, i u sekundárních konektorů neovlivňuje mezivětné či vnitrovětné užití ani tak sémantický typ diskurzního vztahu jako jejich konkrétní forma.

Jak jsme uvedli výše, sekundární konektory mají často příslovečný charakter (srov. *kromě toho, z tohoto důvodu, kvůli tomu* aj.). U těchto struktur převažuje mezivětné užití, srov. příklad 80 z PDiT. V tomto ohledu jsou podobné primárním konektorům s příslovečným původem (typu *proto*), u kterých rovněž převažuje užití mezivětné.

- (80) *Mövenpick provozuje několik desítek hotelů nejen v Evropě, ale i v Asii a Africe. Kromě toho je známý i jako obchodní a potravinářská firma.*

PDiT obsahuje celkem 40 výskytů konektoru *kromě toho*, 36 z nich je mezivětných, 4 vnitrovětné (vždy ve spojení se spojkou *a*). Pokud se tedy tyto struktury zkombinují s primárními konektory typu *a, ale*, převládá u nich užití vnitrovětné, srov. příklad 81 z PDiT.

Sekundární konektory					
Typ vztahu	Mezivětné	%	Vnitrovětné	%	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
Připustka	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 6.18: Podíl mezivětných a vnitrovětných vztahů vyjádřených sekundárními konektory v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p -hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$

- (81) *Rychlejšímu rozvoji telefonizace brání zejména nedostatek ústředěn, a kromě toho se provozovaná zařízení vyznačují nízkými technickými parametry i značnou opotřebeností.*

Zároveň pokud sekundární konektory obsahují podřadící spojku, resp. pokud tvoří podřadící strukturu (srov. kataforické výrazy typu *kromě toho, že; z toho důvodu, aby; díky tomu, že*), rovněž u nich převládne vnitrovětnost. Zajímavé je, že se tyto sekundární konektory chovají jako podřadící spojky – mohou měnit pořadí argumentů, resp. umožňují předsunutí vedlejší věty před větu hlavní: *Dostal jsem špatnou známku, přes-*

tože jsem se hodně učil. vs. *Přestože jsem se hodně učil, dostal jsem špatnou známku.* (více k chování podřadicích spojek viz Quirk a kol., 1985).

Obdobně se sekundární konektory typu *díky tomu, že* mohou vyskytovat v pořadí argumentů Arg1_Arg2 (příklad 82i) i Arg2_Arg1 (příklad 82ii). Zajímavé je, že mohou být použity i v konstrukcích, ve kterých je Arg2 vložen do Arg1 (příklad 82iii), a podílejí se tak na výstavbě textu také z hlediska aktuálního členění větného.

- (82) i) *Do východní části jezdí jenom díky tomu, že tam rodiče koupili v loděnici nějaký prám.*
- ii) *Díky tomu, že jsme se přizpůsobili potřebám trhu, náš firemní obrat plynule roste.*
- iii) *Uvedení spisovatelé jsou díky tomu, že publikovali převážně v angličtině, často omylem považováni za Angličany.*

6.4 Faktory ovlivňující výběr diskurzních konektorů

Pro vyjádření jednotlivých sémantických vztahů v textu má autor na výběr poměrně široký repertoár konektorů, například vztah důsledku může být vyjádřen primárními konektory jako *proto, a tak, tedy* a sekundárními konektory typu *z tohoto důvodu; díky tomu; to je důvod, proč; kvůli tomu* aj. (viz seznamy primárních a sekundárních konektorů v přílohách). Je zajímavé sledovat, za jakých podmínek, podle jakých principů či v jakých kontextech volí autor jeden konkrétní konektor.

V následující části práce se nejprve zaměříme na to, jak můžeme sledovat vhodnost užití diskurzních konektorů v různých kontextech, tj. testovat jejich kontextovou (ne)závislost (kapitola 6.4.1), a posléze se zaměříme na obecné principy a tendence v užívání diskurzních konektorů (kapitola 6.4.2).

6.4.1 Testování kontextové (ne)závislosti diskurzních konektorů

V kapitole 6.1.1 jsme vymezili diskurzní konektory jako jazykové výrazy se spojovací funkcí v textu, které mezi textovými jednotkami vyjadřují určitý významový vztah a které jsou zároveň (do jisté míry) kontextově nezávislé, tj. jsou vhodné pro vyjádření daného sémantického vztahu ve velkém množství kontextů, srov. diskurzní konektory *přesto* a *navzdory tomu* v příkladech 83–85.

- (83) *Nikam se mu nechtělo. Přesto / Navzdory tomu na výlet odjel.*
- (84) *Moc jsem si s novým kolegou nerozuměl. Přesto / Navzdory tomu jsem s ním musel trávit hodně času.*
- (85) *V poslední době jsme moc netrénovali. Přesto / Navzdory tomu jsme zápas vyhráli.*

Dále jsme v práci popsali jazykové prostředky, kterými je v některých případech také možné vyjádřit v textu určitý významový vztah, ale které jsou (na rozdíl od diskurzních konektorů) silně závislé na kontextu, srov. příklad 86.

(86) *Foukal silný vítr a od rána bylo zataženo. Navzdory tomuto počasí se akce konala.*

Prostředky typu *navzdory tomuto počasí* nazýváme volné konektivní fráze – od konektorů se liší tím, že mohou být použity pro vyjádření daného vztahu pouze v několika specifických kontextech.

Kontextovou nezávislost konektivních prostředků (resp. její míru) můžeme ověřovat následujícím způsobem. Vybrané výrazy – *přesto*, *navzdory tomu*, *navzdory tomuto počasí* – zasazujeme do kontextů s příslušným významovým vztahem, v našem případě vztahem přípustky, a sledujeme, zda jsou vybrané výrazy pro dané kontexty vhodné, srov. příklad 87.

- (87) i) *Foukal silný vítr a od rána bylo zataženo. Přesto / Navzdory tomu / Navzdory tomuto počasí se akce konala.*
 ii) *Moc jsem si s novým kolegou nerozuměl. Přesto / Navzdory tomu / *Navzdory tomuto počasí jsem s ním musel trávit hodně času.*

V některých případech je výsledek testování jednoznačný, resp. vyhodnocování kontextové nezávislosti některých prostředků je do jisté míry intuitivní – spojení *navzdory tomuto počasí*, na rozdíl od výrazů *přesto* či *navzdory tomu*, by pravděpodobně většina čtenářů intuitivně označila jako silně kontextově závislou frázi.

Zajímavější je podívat se ale na konektivní prostředky typu *navzdory této skutečnosti*, na základě těchto faktorů nebo vzhledem k této zkušenosti, tj. zaměřit se na hranici mezi sekundárními konektory a volnými konektivními frázemi.

6.4.1.1 Hranice mezi sekundárními konektory a volnými konektivními frázemi

V následující části práce jsme se rozhodli provést test několika konektivních struktur v PDiT, u kterých není vyhodnocení jejich kontextové (ne)závislosti (podle našeho názoru) zcela jednoznačné.

Nejprve jsme museli zvolit významovou kategorii, v rámci které budeme analýzu provádět – jako vhodnou kategorii jsme určili významový vztah důsledku, který je v PDiT jedním z nejfrekvencovanějších vztahů jak pro primární, tak sekundární konektory, a který tedy nabízí široké spektrum kontextů i široké spektrum konektivních struktur.

Pro naši analýzu jsme se rozhodli využít konektivní fráze s podobnou strukturou – zvolili jsme kombinaci sekundární předložky a anaforického výrazu typu *na základě těchto faktorů, díky všem těmto opatřením, na základě zmíněných aktivit, vzhledem k této zkušenosti, vzhledem k těmto skutečnostem, na základě těchto poznatků*.

Dále jsme postupovali ve dvou krocích. Nejprve jsme testovali vybrané sekundární předložky (které podle dat PDiT mohou vyjadřovat vztah důsledku) ve spojení se zájmenem *ten*. Podle výsledků testu jsme posléze vybrali jednu konkrétní předložku ve spojení s obecnějšími jmény typu *faktor*, *opatření*, *aktivita* atd. a sledovali jsme míru jejich kontextové (ne)závislosti na mnoha kontextech.

V prvním případě jsme tedy analyzovali konektivní struktury *díky tomu*, *kvůli tomu*, *na základě toho*, *vzhledem k tomu* a *vinou toho*. Pro testování jejich kontextové (ne)závislosti jsme přitom zvolili sto různých kontextů z PDiT, ve kterých se vyskytuje diskurzivní vztah důsledek. Z těchto kontextů jsme vyjmuli původní konektory a dosazovali do nich testované konektivní struktury. Zjišťovali jsme, pro jaké procento kontextů jsou vhodnými ukazateli tohoto typu vztahu. Zvolili jsme přitom padesát kontextů z PDiT, ve kterých se v originálním, autentickém textu objevil sekundární konektor (či volná konektivní fráze), a padesát kontextů s primárními konektory *proto* a *a tak*.

Do konečných výsledků testování kontextové (ne)závislosti jsme přitom pro jednotlivé výrazy nezapočítávali jejich originální kontexty – netestovali jsme, zda je tedy příslušný výraz vhodný do kontextu, ve kterém se reálně objevil, ale zda je vhodným konektivním prostředkem pro ostatní kontexty.³⁵

Souhrnné výsledky první části výzkumu zachycuje tabulka 6.19 a graf na obrázku 6.2 (výzkumu se zúčastnili tři lingvisté s praktickými zkušenostmi se zpracováním a analýzou diskurzivních vztahů; následující tabulky a grafy zobrazují konečné výsledky, které jsou průměrem hodnot získaných od tří anotátorů).

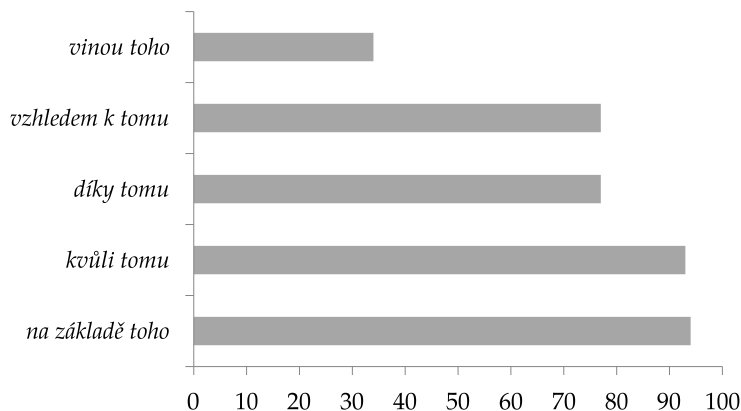
Testované výrazy	Výskyt ve vybraných kontextech	Přijatelnost ve zbylých kontextech	Míra kontextové (ne)závislosti %
<i>na základě toho</i>	0	94	94
<i>kvůli tomu</i>	3	90	93
<i>díky tomu</i>	7	72	77
<i>vzhledem k tomu</i>	1	76	77
<i>vinou toho</i>	1	34	34

Tabulka 6.19: Konektivní prostředky ve formě předložkových frází se zájmenem *ten*

Škálu kontextové (ne)závislosti konektivních prostředků ve formě předložkových frází obsahujících zájmeno *ten* zobrazuje graf na obrázku 6.2.

Z grafu je patrné, že spojeními s nejvyšší mírou kontextové nezávislosti jsou podle naší analýzy *na základě toho* (94 %) a *kvůli tomu* (93 %). Sekundární předložku *na základě* jsme proto použili v další části našeho výzkumu, ve které jsme se zaměřili na testování konektivních prostředků obsahujících obecnější podstatná jména (typu *opatření*, *skutečnost*). S předložkou *na základě* jsme testovali taková obecnější jména, která se v PDiT alespoň jednou vyskytla ve spojení se sekundární předložkou (daná podstatná jména v těchto případech plní anaforickou funkci). V další fázi výzkumu jsme

³⁵ Míru kontextové (ne)závislosti konektivních prostředků jsme vypočítávali následujícím způsobem: např. spojení *díky tomu* se ukázalo jako vhodné pro 72 (neoriginálních) kontextů z celkových 93 (nazapočítáváme 7 originálních kontextů z celkových 100 vybraných, ve kterých se *díky tomu* původně objevilo), tj. kontextová nezávislost spojení *díky tomu* v našem testu je 0,77 (72/93), tedy 77 %.



Obrázek 6.2: Škála kontextové (ne)závislosti konektivních prostředků ve formě předložkových frází se zájmenem *ten*

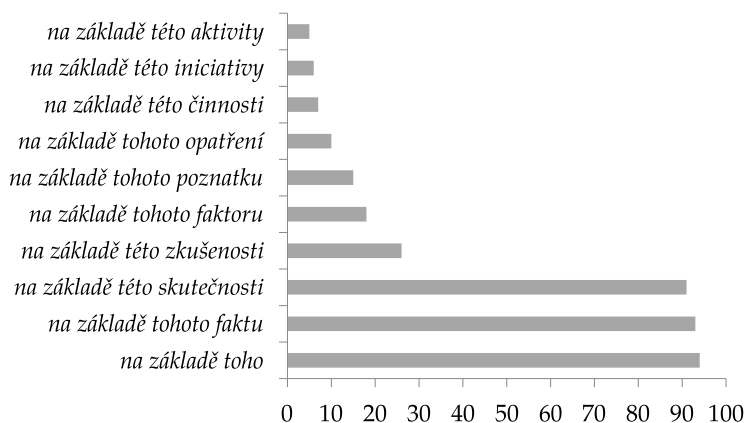
tedy zkoumali následující výrazy: *na základě této skutečnosti, na základě tohoto opatření, na základě této aktivity, na základě těchto faktorů, na základě těchto poznatků, na základě této zkušenosti, na základě této činnosti, na základě tohoto faktu, na základě této iniciativy*. Do analýzy jsme zároveň pro srovnání zahrnuli také spojení *na základě toho*. V těchto případech bude tedy testování kontextové nezávislosti v podstatě vyhodnocovat míru lexikální naplněnosti či vyprázdňenosti daných anaforických výrazů, resp. jejich schopnost významově pokrýt co největší počet různých kontextů.

Dané výrazy jsme testovali znovu ve stu kontextech (stejných jako v první fázi výzkumu). U užitých obecných jmen jsme přitom 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 6.20 a graf na obrázku 6.3.

Z výsledků vyplývá, že tři konektivní prostředky (*na základě toho, na základě tohoto faktu, na základě této skutečnosti*) dosahují vysoké kontextové nezávislosti (přesahující 90 %), a tím výrazně převyšují zbylé hodnocené výrazy. Naše výsledky přitom nejsou překvapivé – slova *skutečnost* a *fakt* se (především v odborných textech) vyskytují v podobných funkcích jako zájmeno *ten*, srov. vyjádření typu *předmětem našeho zájmu je to / skutečnost, zda..., to / tento fakt můžeme doložit na následujícím příkladu*, tj. především při anaforickém a kataforickém odkazování v textu. Možnost vzájemné zastupitelnosti zmíněných slov ve velkém množství kontextů se tedy ukázala i v našem výzkumu konektivních prostředků.

Testované výrazy	Výskyt ve vybraných kontextech	Přijatelnost ve zbylých kontextech	Míra kontextové (ne)závislosti %
<i>na základě toho</i>	0	94	94
<i>na základě tohoto faktu</i>	0	93	93
<i>na základě této skutečnosti</i>	0	91	91
<i>na základě této zkušenosti</i>	0	26	26
<i>na základě tohoto faktoru</i>	1	18	18
<i>na základě tohoto poznatku</i>	1	15	15
<i>na základě tohoto opatření</i>	0	10	10
<i>na základě této činnosti</i>	0	7	7
<i>na základě této iniciativy</i>	0	6	6
<i>na základě této aktivity</i>	0	5	5

Tabulka 6.20: Konektivní prostředky ve formě předložkových frází s předložkou *na základě*



Obrázek 6.3: Škála kontextové (ne)závislosti konektivních prostředků ve formě předložkových frází s předložkou *na základě*

6.4.1.2 Kontextová (ne)závislost vybraných primárních a sekundárních konektorů

V další části výzkumu jsme se zaměřili na testování kontextové (ne)závislosti vybraných primárních a sekundárních konektorů. Opět jsme zvolili konektory vyjadřující vztah důsledku, do analýzy jsme přitom zahrnuli strukturně odlišné konektory – srov. primární konektory *proto* a *a tak* a sekundární konektory *z tohoto důvodu*; *to je*

důvod, proč; kvůli tomu; díky tomu; z toho důvodu; na základě čehož; z uvedeného důvodu; vinou toho; vzhledem k tomu.

Test jsme prováděli opět na 100 původních kontextech z PDiT (stejných jako v předchozím výzkumu). U flektivních struktur jsme při testování nerozlišovali jednotné a množné číslo, 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 tabulce 6.21 je nicméně pro ilustraci užita pouze základní forma (tj. jednotné číslo). V tabulce uvádíme také četnost zkoumaných konektorů v Českém národním korpusu, konkrétně v korpusu SYN2013PUB³⁶ (obsahujícím 935 milionů slov z publicistických textů).

Testované výrazy	Výskyt ve vybraných kontextech	Přijatelnost ve zbylých kontextech	Míra kontextové (ne)závislosti %	Výskyty v ČNK
<i>z uvedeného důvodu</i>	1	98	99	189
<i>z tohoto důvodu</i>	8	90	98	8 801
<i>kvůli tomu</i>	3	90	93	53 755
<i>z toho důvodu</i>	3	83	86	8 623
<i>proto</i>	25	62	83	733 198
<i>to je důvod, proč</i>	6	77	82	1 884
<i>a tak</i>	25	61	81	228 532
<i>díky tomu</i>	7	72	77	40 376
<i>vzhledem k tomu</i>	1	76	77	39 497
<i>vinou toho</i>	1	34	34	188
<i>na základě čehož</i>	1	21	21	39

Tabulka 6.21: Kontextová (ne)závislost vybraných primárních a sekundárních konektorů

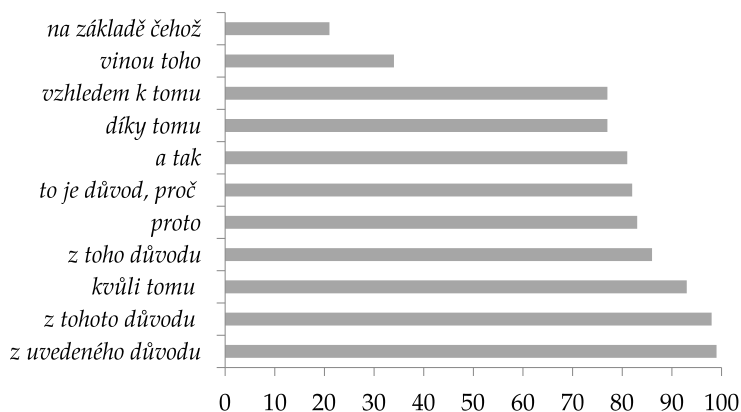
Při vyhodnocování výsledků jsme se nejprve zaměřili na vztah kontextové nezávislosti konektorů a jejich frekvence. Zkoumali jsme především to, zda při určování určitého jazykového prostředku jako diskurzního konektoru můžeme zohledňovat vedle míry kontextové nezávislosti také jeho frekvenci v textech.

Z výsledků je patrné, že míra kontextové nezávislosti daného konektoru koreluje s jeho frekvencí pouze v některých případech. Např. spojení *na základě čehož* má mezi 11 zkoumanými konektory jak nejnižší kontextovou nezávislost, tak nejnižší počet výskytů v korpusu SYN2013PUB. Obdobně vyrovnané výsledky vykazuje dále např. *vinou toho*. Na druhou stranu např. spojení *z uvedeného důvodu* dosáhlo nejvyššího stupně kontextové nezávislosti, ale zároveň má třetí nejnižší počet výskytů. Vzhle-

³⁶ U jednotlivých výskytů jsme blíže nezkoumali jejich funkci v textu, tj. tabulka zobrazuje absolutní čísla výskytů u jednotlivých výrazů nalezených v korpusu SYN2013PUB.

dem k tomu nebereme frekvenci jazykových prostředků v textech jako hlavní kritérium pro vymezení kategorie diskurzních konektorů.

Dále jsme se zaměřili na výslednou míru kontextové (ne)závislosti zkoumaných konektorů, srov. graf na obrázku 6.4.



Obrázek 6.4: Škála kontextové (ne)závislosti primárních a sekundárních konektorů

Z výsledků výzkumu je zřejmé, že většina vybraných konektorů dosahuje vysokého stupně kontextové nezávislosti, pohybujícího se mezi 77 % (*vzhledem k tomu*) a 99 % (*z tohoto důvodu*). Znatelně nižší kontextové nezávislosti dosáhly v našem testu výrazy *na základě čehož* (21 %) a *vinou toho* (34 %).

Spojení *na základě čehož* (21 %) je výrazně limitováno svým syntaktickým chováním – je silně vázáno na vnitrovětné vztahy, resp. jeho mezivětné užití by mohlo být vnímáno jako syntaktický nedostatek. Na míře kontextové nezávislosti tohoto spojení se tedy v našem zkoumání projevil výběr kontextů z PDiT, resp. to, že většina z nich obsahovala textové vztahy mezivětné. Prozkoumali jsme tedy znovu 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 kontextová nezávislost 81 %.

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*, 2003). 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. předložku *díky*, která se v textech PDiT objevuje i v negativních kontextech typu *díky jeho chorobě*), jiné si

tuto vázanost ponechávají (jako *vinou*). Zatímco tedy spojení *díky tomu* může být použito i v některých kontextech, ve kterých se původně vyskytovalo *vinou toho* (alespoň podle dat PDiT), obdobná nahraditelnost v opačném směru nefunguje, srov. příklad z PDiT, ve kterém jsme spojení *díky tomu* nahradili spojením *vinou toho*:

- (88) *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* ve funkci diskurzních konektorů³⁷ 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, budeme analogicky i spojení *vinou toho* řadit mezi diskurzní konektory.

Náš experiment poukázal na skutečnost, že existují určitá omezení, která zabraňují užití diskurzního konektoru v určitém kontextu. Tato omezení mohou být přitom různého charakteru (např. syntaktického či sémanticko-pragmatického, jak jsme ukázali na konektorech *na základě čehož* a *vinou toho*). V následující části práce (kapitola 6.4.2) se na podobná omezení blíže zaměříme.

6.4.2 Principy a tendence v užívání diskurzních konektorů

Při tvorbě koherentního textu 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. Jak jsme naznačili výše, na výběr příslušného konektivního prostředku přitom mají vliv různé faktory.³⁸

Gramatika: souřadnost vs. podřadnost, mezivětnost vs. vnitrovětnost

Jak jsme naznačili výše, užití některých diskurzních konektorů je výrazně limitováno jejich syntaktickým chováním. Některé konektory jsou vázány na vnitrovětné užití, jejich mezivětné užití je rodilými mluvčími vnímáno jako syntaktická neobratnost. Na vnitrovětné užití jsou vázány především podřadící struktury (srov. primární konektory ve formě podřadících spojek typu *protože, když, pokud, aby, kdyby* či sekundární konektory typu *kromě toho, že; za účelem, aby*).³⁹ Podřadící konektory (resp. jejich výsky-

³⁷ Sémantickými a pragmatickými aspekty diskurzních konektorů se podrobně zabývá J. Hoffmannová (srov. např. 1983, 1984). Ve svých pracích přináší podrobnou klasifikaci diskurzních konektorů právě na základě podílu sémantických a pragmatických rysů.

³⁸ K faktorům ovlivňujícím výběr diskurzních konektorů viz také Rysová a Rysová (2018).

³⁹ V některých případech (např. v uměleckých textech) může být mezivětné užití podřadících konektorů funkční, jedná se o tzv. parcelaci výpovědi.

ty) dohromady tvoří více než čtvrtinu všech výskytů konektorů v PDiT. Vázanost na vnitrovětné užití tak postihuje relativně vysoký počet výskytů konektorů obecně.

Vnitrovětné užití je dále preferováno některými souřadícími konektory (typu *a, ale, či*), jiné (především konektory adverbiálního původu jako *proto, však* aj.) preferují mezivětnost, více viz kapitola 6.3.3. V těchto případech se ovšem jedná o tendence v užívání konektorů, nikoli o syntaktické omezení (většina souřadících konektorů dovoluje užití vnitrovětné i mezivětné, i když jedno z nich vždy poměrně výrazně převažuje).

Gramatika: slovosled

Diskurzní konektory se mezi sebou liší i z hlediska slovosledného. Zatímco konektory ve formě souřadících spojek neadverbiálního původu stojí většinou na hranici mezi dvěma diskurzními argumenty (příklad 89), adverbiální konektory mívají více možností, kde být ve větě umístěny (příklady 90 a 91). Konektory ve formě podřadících spojek zase dovolují měnit pořadí hlavní a vedlejší věty, srov. Quirk a kol., 1985 (příklad 92).⁴⁰ V některých kontextech tak nejsou konektory s podobným významem (např. *proto, takže*), ale různými slovoslednými vlastnostmi vzájemně zastupitelné.

- (89) *Ještě jednou se ohlédla přes rameno a sledovala racky kroužící nad vodou.*
- (90) *Sklo lze recyklovat stále dokola. **Proto** je jeho sběr velmi důležitý. / Jeho sběr je **proto** velmi důležitý.*
- (91) *Taborou každý den projíždějí karavany vezoucí velké bohatství. **Kvůli tomu** vzniklo už několik šarvátek. / Vzniklo **kvůli tomu** už několik šarvátek.*
- (92) *Toho hosta budu muset obsloužit já, **protože** moje kolegyně odběhla do kuchyně. / **Protože** moje kolegyně odběhla do kuchyně, budu toho hosta muset obsloužit já.*

Sémantika

Dalším faktorem ovlivňujícím užívání diskurzních konektorů je sémantika. Některé z konektorů vyjadřujících tentýž typ vztahu jsou synonymní (zaměnitelné) v určitém množství kontextů (příklad 93), ale najdeme kontexty, kde zaměnitelné nejsou (příklady 94 a 95). Tyto konektory tak mají význam podobný, ale ne totožný, jde o částečná synonyma.

- (93) *Proních asi 10–12 dní jsou větší otoky a modřiny a pacient má sádra fixující nos, **kvůli tomu** / **a tak** je nezbytné zůstat v domácí péči.*
- (94) *Odpoledne měl volno, **a tak** / **?kvůli tomu** se toulal po městě.*
- (95) *Fanoušci skandovali mé jméno, byl to krásný pocit. **Kvůli tomu** / **?A tak** vlastně hrají fotbal.*

⁴⁰ Gramatickými vlastnostmi konektorů v češtině se podrobně zabývá studie Š. Zikánové a kol. (2015b).

Jednotlivé konektory se obecně liší v tom, jak široké či úzké možnosti použití mají. Například konektor *a* má širší význam než konektor *proto*. Zatímco *a* může stát mezi propozicemi s velmi různorodými typy vztahů, od konjunkce po opozici, konektor *proto* se vyskytuje v podstatě výhradně mezi propozicemi se vztahem důsledku.

Nositeli užšího významu bývají často sekundární konektory, mají totiž schopnost vyjadřovat specifitější odstíny daného typu vztahu (např. struktury *důvodem je a příčinou je* vyjadřují jemnější podtypy obecně kauzálního vztahu). Sekundární konektory se v textu také často objevují v modifikované podobě – tím svůj význam (a zároveň možnost využití) ještě zužují – srov. *hlavním / možným / podstatným důvodem bylo* aj.

Pragmatika

Na výběru diskurzních konektorů se dále podílejí jejich pragmatické aspekty. Jak jsme ukázali výše, konektor *vinou toho* je silně vázán na negativní kontexty, možnosti jeho použití ve větším spektru kontextů jsou tak značně limitovány. Většina konektorů v češtině je v tomto smyslu neutrální (tj. nejsou omezeny na pozitivní ani negativní kontexty, příp. tyto konotace postupně ztrácejí, viz *díky tomu*). Pragmatický prvek mohou konektorům přidávat také modifikace – srov. *cynicky řečeno, s nepřesvědčivým odůvodněním* apod.

Stylistika

Konektory (či jejich varianty) se liší také z hlediska stylistické platnosti. V jazyce se vyskytují konektory formální (*tedy, z tohoto důvodu, z uvedeného důvodu*), neutrální (*proto*) i neformální (*teda*). Nejširší možnosti užití mají konektory neutrální. Konektory formální a neformální se vyskytují v závislosti na typu textu.

Užití konektoru ve stylově opačně zabarveném kontextu je čtenáři vnímáno jako neobratné (srov. užití neformálního konektoru *teda* v odborném textu, příklad 96).

(96) ? Z výsledků našeho výzkumu *teda* vyplývá následující.

Vzdálenost a salience

Autorův výběr vhodného konektivního prostředku může být ovlivněn také tím, jak vzdálené argumenty je třeba propojit. Pokud se mezi diskurzními argumenty nachází ještě další, rozsáhlejší text, může být nutné místo konektoru užít volnou konektivní frázi, která má schopnost čtenáři připomenout předchozí kontext, srov. příklad 97. Užití konektoru by v tomto případě mohlo způsobit nedorozumění.

(97) V letošním ročníku jsme se díky vašim hlasům stali absolutními vítězi ankety a získali jsme nemalou finanční podporu. Stále se z toho radujeme. Celkově jsme totiž obdrželi téměř 800 tisíc hlasů. Nejvíce nás ale těší, že nerozhodovala komise, ale zákazníci – tedy veřejnost. *Díky této výhře* uskutečníme 130 dětských klauniád a rozsvítíme tak oči stovek malých pacientů v nemocnicích po celé České republice.

Pokud bychom v příkladu 97 užili místo volné konektivní fráze *díky této výhře* konektor *proto*, mohl by se čtenář domnívat, že vztah jím vyjádřený se váže k bezprostředně sousední propozici. Naprostá většina diskurzních vztahů v PDiT se totiž vyskytuje právě mezi sousedními propozicemi – čtenář tedy tento typ vztahu očekává spíše než vztah zasahující daleko do textu. Autor v příkladu 97 ovšem neříká, že může uskutečnit 130 dětských klauniád *proto*, že o anketě rozhodovali zákazníci, ale *proto*, že díky vítězství v anketě získal nemalou finanční podporu.

Volné konektivní fráze tedy hrají důležitou roli v připomínání předchozího kontextu čtenáři a při vyvozování diskurzních vztahů mezi vzdálenějšími argumenty. Jsou tak významné z pohledu tzv. salience, která je spojená s pronikáním podnětů do vědomí, resp. s jejich aktivovaností v mysli, srov. k tomu E. Hajičová (1993, 2013).

Princip minimálního úsilí: jazyková ekonomie

V předchozí části práce jsme popsali různé jazykové faktory ovlivňující výběr diskurzních konektorů v konkrétních kontextech, resp. specifikovali jsme, jaké objektivní jazykové aspekty omezují konektory v jejich užití. Nyní se zaměříme na preference v užívání konektorů na základě jejich frekvence v textech PDiT, tj. na to, jaké konektory autor textu skutečně používá. Podíváme se především na to, zda v tomto ohledu můžeme vysledovat v jazyce určité obecné tendence.

Výsledky naší anotace ukázaly, že PDiT obsahuje 21 416 diskurzních vztahů vyjádřených konektory – 20 255 z nich (94,6 %) tvoří primární konektory a 1 161 (5,4 %) sekundární konektory. Je tedy zřejmé, že autoři textů PDiT jasne preferují konektory primární před sekundárními. Pro autora je patrně jednodušší používat primární konektory, protože se jedná o ustálené, kratší, většinou jednoslovné, neohebné výrazy. Sekundární konektory naopak vykazují velkou míru variability (lexikální, morfologické i syntaktické) a bývají výrazně delší. Z toho vyplývá, že se při užívání konektorů uplatňuje princip tzv. jazykové ekonomie. Autor se snaží s využitím minimálního úsilí dosáhnout maximálního výsledku, srov. Zipfův zákon minimálního úsilí (1949), a využívá vždy co nejjednodušší cestu, která vede ke koherentnímu textu. Podobnou tendenci můžeme pozorovat i na používání konektorů v angličtině. Jak uvádí Prasadová a kol. (2010, 2014), v Penn Discourse Treebank se objevilo 18 459 (96,7 %) konektorů a 624 (3,3 %) alternativních lexikálních vyjádření konektorů (altlexů). I v angličtině tedy jasne převažuje užívání primárních konektorů nad dalšími možnostmi vyjadřování textových vztahů. Můžeme se tedy domnívat, že princip minimálního úsilí v užívání diskurzních konektorů platí napříč jazyky.

Předpoklad, že autor volí ve vyjadřování textových vztahů nejjednodušší cestu, se potvrzuje i při pohledu na frekvence výskytů jednotlivých konektorů. Zcela nejfrekventovanějším konektorem je spojka *a*. Její výskyty tvoří 27 % výskytů všech konektorů (v PDiT 2.0 se konektor *a* objevil v 5 766 případech, PDiT přitom celkově obsahuje 21 416 výskytů konektorů). První tři nejfrekventovanější konektory (*a*, *však*, *ale*) pokrývají 40 % výskytů všech konektorů (připadá na ně 8 554 výskytů), prvních pět (*a*, *však*,

ale, když, protože) tvoří dohromady 45 % (9 653 případů), prvních deset (*a, však, ale, když, protože, totiž, pokud, proto, tedy, aby*) pokrývá 54 % (11 508 případů).

PDiT přitom obsahuje přibližně 570 různých forem primárních konektorů a přibližně 400 různých forem, resp. realizací sekundárních konektorů (různé typy složených konektorů a modifikovaných konektorů jsou zde započítány jako samostatné jednotky, viz příloha). Celkem tedy PDiT obsahuje přibližně 1 000 rozrůzněných forem a realizací konektorů. Ve více než polovině případů, kdy autor potřeboval v textu použít diskurzní konektor, však vybíral pouze z deseti z nich. Jinými slovy deset nejfrekventovanějších konektorů stačilo na pokrytí více než poloviny komunikačních situací, v nichž bylo potřeba explicitně vyjádřit sémanticko-pragmatický textový vztah. Ačkoli má tedy autor na výběr z širokého repertoáru konektorů a jejich forem, používá přednostně jen několik z nich. Pět nejfrekventovanějších konektorů v PDiT zároveň významově pokrývá čtyři základní typy anotovaných vztahů: vztahy navazovací (*a*), kontrastivní (*však, ale*), časové (*když*) a kauzální (*protože*).

Princip maximální srozumitelnosti

Na druhé straně je nutné se ptát, proč v jazyce existuje tak široká množina ostatních, méně používaných konektorů a jejich forem. Zdá se, že z hlediska jazykové ekonomie by mnoho konektorů nemuselo v jazyce vůbec existovat. Je otázka, proč mít pro jeden typ vztahu k dispozici více synonymních konektorů (např. *proto, tedy, tak, takže, tudíž, kvůli tomu, v důsledku toho, díky tomu, z tohoto důvodu*), když pro většinu situací stačí jeden základní.

Například vztah konjunkce je v datech PDiT vyjadřován celkem 195 konektivními formami, nejčastěji konektorem *a* (tvoří 74 % všech výskytů tohoto typu vztahu v PDiT). Další konektory vyjadřující konjunkci jsou např. výrazy *také, což, dále, rovněž* (jejichž frekvence postupně výrazně klesá). Obdobně vztah opozice je nejčastěji signalizován konektorem *však* (ve 39 % všech vztahů opozice), ale současně může být vyjadřován i jinými výrazy jako *ale, ovšem, sice_ale, jenže* aj. (celkem existuje 109 konektivních forem vyjadřujících v PDiT vztah opozice). Podobně bychom mohli pokračovat i u ostatních sémantických vztahů. Jak už jsme naznačili výše, většina konektorů (resp. jejich konkrétních forem) se v PDiT objevuje velice zřídka, často v jednom či ve dvou výskytech – pouze jednou se objevuje 283 forem primárních konektorů (50 % z 570) a 264 forem sekundárních konektorů (66 % ze 400). Přítomnost těchto málo frekventovaných konektorů v korpusu znamená, že v jazyce existují kontexty, ve kterých autor upřednostnil méně obvyklé konektory před těmi „typickými“. Jinými slovy pro vyjádření stejného sémantického vztahu je v jazyce zapotřebí mít k dispozici více synonymních konektorů, a to i zmíněné málo frekventované konektory (např. *v důsledku toho*), protože existují kontexty, do kterých se (podle autorova výběru) hodí více než konektory frekventované (např. *proto*).

Ze skutečnosti, že se takto široký repertoár synonymních konektorů v jazyce dlouhodobě udržel, je zároveň patrné, že dané konektory zcela synonymní nejsou, tj. v jazyce najdeme kontexty, které vyžadují užití specifičtějšího, obecně méně frekventova-

ného konektoru. Plná synonyma totiž mají tendenci z jazyka mizet jako neekonomické varianty (srov. Cruse, 2000), a to nejen v případě konektorů.

Jak jsme popsali výše, konektory (patřící k jednomu významovému typu vztahu jako *proto*, *tedy*, *a tak*, *z tohoto důvodu* aj.) se vzájemně odlišují v různých jazykových aspektech, jedná se tedy pouze o částečná synonyma. Jednotlivé konektory mohou vyjadřovat různé sémantické nuance daného vztahu, obdobně modifikace specifikuje jejich význam apod. Významově specifitější bývají většinou právě méně frekventované konektory (viz rozdíl v četnosti obecnějších primárních a konkrétnějších sekundárních konektorů). V některých kontextech autoři potřebují právě tyto konektory k přesnějšímu vyjádření svého komunikačního záměru (chtějí například sdělit, že se něco děje za *výjimečných* podmínek, *z jasného* důvodu nebo *hlavně* *proto*).

Široký repertoár diskurzních konektorů také umožňuje autorovi, aby se vyvaroval nefunkčního opakování slov. V některých případech je třeba vyjádřit dva diskurzní vztahy stejného typu na malém prostoru. Proto je užitečné mít k dispozici jak základní, vysoce frekventované konektory pro daný vztah (například *proto* pro vztah důsledku), tak jejich méně frekventovaná synonyma (například *tudíž*), viz příklad 98.

- (98) *Snižováním počtu průduchů a někdy i přivíráním průduchů rostlina spotřebuje méně vody na svůj růst i při optimální fotosyntéze. **Proto** je zvyšování koncentrace oxidu uhličitého pro rostliny výhodné. Jsou **tudíž** zbytečné obavy, že zvyšování koncentrace oxidu uhličitého v atmosféře přivodí katastrofu našich kulturních rostlin.*

Skutečnost, že autor využívá široký repertoár konektorů, je tedy způsobena snahou vytvořit čtenáři co nejsrozumitelnější, nejkoherentnější text (viz také Griceův kooperační princip, 1975).

7

Závěr

V naší práci jsme se zaměřili na popis a rozbor diskurzních konektorů v češtině. Konektory jsme vymezili na základě jejich funkce v textu, a to jako jazykové výrazy, které vyjadřují určitý významový vztah mezi dvěma úseky textu (např. vztah důsledku, opozice, specifikace aj.) a které zároveň mohou tento typ vztahu vyjadřovat ve velkém počtu různých kontextů. Tím odlišujeme konektory od dalších konektivních prostředků (tzv. volných konektivních frází), které jsou naopak silně závislé na kontextu.

Jak jsme v našem výzkumu ukázali (srov. především kapitolu 6.1), diskurzní konektory přitom tvoří značně heterogenní skupinu výrazů, a to po stránce lexikální, syntaktické i sémantické. Konektory se v textu navíc různě kombinují, některé se objevují s modifikacemi. V textech se tak můžeme setkat s takto různorodými příklady konektorů a jejich kombinací: *a, však, a právě proto, buď_ nebo, kvůli tomu, nebo, ale, proto, důvodem je, s jasným cílem, i když, a kvůli tomu, ale naopak, když_tak, hlavně proto, v případě, že_pak, jedině kvůli tomu, především vinou toho* apod. V naší práci jsme se proto snažili vyčlenit v rámci kategorie diskurzních konektorů několik obecných podskupin, díky kterým je možné konektory blíže popsat a charakterizovat a lépe pochopit jejich funkci ve strukturaci textu.

V naší práci jsme vymezili následující opozice **diskurzních konektorů**:

a) **konektory primární** (*však, a, nebo, ale, proto...*) vs. **sekundární** (*důvodem je, kvůli tomu, s jasným cílem...*) lišící se ve stupni gramatikalizace;

b) **konektory inherentně referenční** (anaforické: *proto, přesto, potom, kvůli tomu...*; kataforické: *protože; z toho důvodu, že; kvůli tomu, aby...*) vs. inherentně **nereferenční** (*ale, a, nebo, i...*);

c) **konektory složené** (vyskytující se ve stejném diskurzním argumentu: *i když, a kvůli tomu, ale naopak, nebo* tvořící korelativní páry: *buď_ nebo, když_tak, v případě, že_pak*) vs. **nesložené** (tj. skládající se ze synchronního hlediska z jednoho konektoru, viz konektory jednoslovné, gramatikalizované: *nebo, ale*, či víceslovné, negramatikalizované: *z tohoto důvodu*);

d) **konektory modifikované** (obsahují hodnotící, modální či intenzifikační výraz, který upřesňuje, případně zesiluje význam příslušného diskurzního vztahu: *hlavně proto, s jasným cílem, jediné kvůli tomu, především vinou toho*) vs. **nemodifikované** (*proto, s cílem, kvůli tomu, vinou toho*);

e) **konektory vnitrovětné** (podřadící konektory typu *protože; když; kromě toho, že; souřadící konektory neadverbiálního původu jako a, či; dále konektory s nimi zkombinované typu a proto, ale; konektory obsahující negaci jako nebo, nýbrž a korelativní páry typu buď_nebo*) vs. **mezivětné** (konektory adverbiálního původu nebo s adverbiální funkcí: *proto, tedy, díky tomu, za tímto účelem*).

Uvedené aspekty se mohou vzájemně kombinovat – 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*).

V naší práci jsme se dále zaměřili především na opozici konektorů primárních a sekundárních a na výsledky jejich zpracování v Pražském diskurzním korpusu. Z výsledků je patrné, že se sekundární konektory vyskytují v textech s podstatně nižší frekvencí než konektory primární (sekundární konektory tvoří pouze 5,4 % všech vztahů v PDiT vyjádřených konektory).

V závěru práce jsme se zaměřili na principy a tendence v užívání diskurzních konektorů. Zjistili jsme, že na výběr vhodného konektivního prostředku působí několik faktorů zároveň – na jedné straně volí autor takové prostředky, jejichž použití je pro něj při vytváření textu v dané situaci nejjednodušší (tj. v případě konektivních výrazů se jedná častěji o ustálenější, krátké, mnohdy jednoslovné konektory). Zároveň musí brát autor 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 autora a zároveň co nejvyšší srozumitelnost textu pro adresáta, do určité míry proti sobě. Vzájemná interakce obou těchto faktorů je ovšem v každodenní komunikaci běžná a přirozená.

V naší práci jsme postupně ukázali, 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 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.

Summary

The monograph focuses on delimitation and description of discourse connectives as essential contributors to overall text coherence. Discourse connectives have a function of signalizers of semantico-pragmatic relations within a text and as such, they contribute to the comprehensibility and intended interpretation of a text as a whole.

In the monograph, we examine discourse connectives in Czech. The analysis is carried out on the data of the Prague Discourse Treebank 2.0 (PDiT). The results of the research are presented in four parts. The first one (Section 6.1) focuses on delimitation and characteristics of discourse connectives in Czech, the second one (Section 6.2) describes annotation of connectives in the PDiT and demonstrates its evaluation, the third one (Section 6.3) presents results of this annotation, and the fourth one (Section 6.4) discusses linguistic principles and tendencies in using discourse connectives.

In the discussion on the delimitation of discourse connectives, we present both Czech and foreign established approaches to discourse connectives and we try to contribute to the description of connectives via a corpus analysis of Czech material.

Discourse connectives

We specify discourse connectives as language expressions that signal semantico-pragmatic discourse relations within a text (such as opposition, specification, reason...) and at the same time, that have an ability to express these types of relations in various contexts. For example, the connective *proto* (*so*) expresses discourse relation of result in both of the following contexts: *Je mi horko, sundám si proto svetr.* (*I'm hot, so I take off my sweater.*) / *Do hotelu je to ještě daleko, vezmeme si proto taxi.* (*The hotel is still far away, so we take a taxi.*). As such, discourse connectives are nearly context independent (i.e. they may be used for expressing the relevant type of relation in several different contexts).

This delimitation applies to a wide range of language expressions, e.g. to both single-word conjunctions such as *a* (*and*), *ale* (*but*), *protože* (*because*) and multiword phrases such as *za této podmínky* (*under this condition*), *z tohoto důvodu* (*for this reason*), *na závěr* (*in conclusion*) etc. Therefore, we distinguish two types of connectives: primary and secondary.

Primary and secondary connectives

We describe primary connectives as language expressions with a connective function that are grammaticalized, mostly single-word (cf. *a* – *and*, *ale* – *but*, *proto* – *therefore*, *když* – *when*, *protože* – *because*, *nebo* – *or*) or less often multiword (forming correlative pairs such as *bud'nebo* – *either_or* or complex forms such as *i když* – *even though*). Primary connectives are lexically fixed, they do not have a function of sentence elements, they are uninflected and mostly do not allow for internal modification.

Secondary connectives are not (yet) fully grammaticalized language structures with a connective function in a text. In contrast to primary connectives, they are less fixed – their variability concerns lexical realization (cf. *za této / té / dané podmínky* – *under this / that / the given condition*), morphology (*důvod je* vs. *důvodem je* – both meaning *the reason is*, the noun *reason* once being in nominative or in instrumental case) or word order (cf. *to je příčinou* – *that is the cause* vs. *je to příčinou* – lit. *is that the cause*). Secondary connectives are integrated into the sentence structure as sentence elements (*kvůli tomu* – *due to this*) or they function as sentence modifiers (*stručně řečeno* – *simply speaking*). Secondary connectives are often inflected (*za této podmínky* – *under this condition* / *za těchto podmínek* – *under these conditions*) and they allow for internal modification (*hlavní / jedinou / další podmínkou je* – *the main / the only / another condition is*).

Secondary connectives contain so-called core words, i.e. semantically “strong” words forming the basis of the structure – typically, the core words refer to one of the semantico-pragmatic types of discourse relations. The core words may be e.g. nouns (*podmínka* – *condition*, *důvod* – *reason*, *příčina* – *cause*...), prepositions (*kvůli* – *due to*, *díky* – *thanks to*, *navzdory* – *despite*, *na základě* – *on the basis of*...) or verbs (*shrnout* – *to sum up*, *vysvětlit* – *to explain*, *předcházet* – *to precede*...).

We may observe that many connectives with the same core word are formed according to the same general scheme – e.g. the structures *za této podmínky* (*under this condition*), *za té podmínky* (*under that condition*), *za dané podmínky* (*under the given condition*) belong to the scheme “*za(under)_Pron.anaph/Adj.anaph_podmínka(condition).2*” (the preposition *za* – *under* combined with an anaphoric expression and the noun *podmínka* – *condition* in genitive). We call these structures lexical realizations and we delimit them as forms of secondary connectives with the same core word and the same scheme but with slightly different lexical items (cf. *těto / té / dané* – *this / that / the given*). The particular lexical realizations of secondary connectives may appear in various grammatical variants in the text. These variants differ in morphology (*za této podmínky* – *under this condition* vs. *za těchto podmínek* – *under these conditions*) or in word order (*to je příčinou, proč* – *that is the cause why* vs. *je to příčinou, proč* – lit. *is that the cause why*).

There are many primary and secondary connectives in Czech that fully correspond to the above-mentioned characteristics. At the same time, there are some connectives that differ in some points – e.g. Czech primary connectives *aby* (*so that*), *kdyby* (*if*) may be inflected or *proto, že* (*because*) may be modified (cf. *jednoduše proto, že* – *sim-*

ply because). Therefore, primary and secondary connectives do not form two strictly separated classes but rather a scale of expressions with a different degree of grammaticalization.

As said above, primary connectives are more fixed expressions than secondary ones and they appear also more frequently in the texts (as was found out during the annotation of connectives in the PDiT). Therefore, we consider primary connectives as the centre of discourse connectives and secondary once as their periphery.

Discourse connectives (both primary and secondary) are also implicitly referential – they assume the presence of two text units they connect. At the same time, some connectives contain an internal reference (cf. connectives *proto* – *therefore*, *z tohoto důvodu* – *for this reason*, *kvůli tomu* – *due to this* etc.). These referential connectives may be then divided into anaphoric and cataphoric.

Free connecting phrases

The semantico-pragmatic discourse relations may be expressed also by other language means than discourse connectives – cf. examples such as *Ceny benzínu se rapidně zvýšily. Kvůli tomuto zvýšení mnoho rodin nebude moct jezdit autem tak často.* – *Gasoline prices have risen rapidly. Because of this increase, many families will not be able to drive so often.* We call the phrases like *kvůli tomuto zvýšení* (*because of this increase*) free connecting phrases. They may also express discourse relations in texts (here the relation of result); however, they may be used only in a few specific contexts. The main difference between these phrases and discourse connectives is that free connecting phrases are highly context dependent.

Secondary connectives in the Prague Discourse Treebank

In the monograph, we describe the way in which secondary connectives are annotated in the Prague Discourse Treebank 2.0 (which is one of the cores of our results) and we present its evaluation. The PDiT contains almost 50,000 of sentences from the newspaper texts that are annotated in dependency trees. The annotators of secondary connectives captured these expressions in two steps. Firstly, they marked the core of the connective (e.g. the noun *podmínka* – *condition*) and then the whole structure (*za této podmínky* – *under this condition*). In this way, it was possible to capture all the core words and the specific lexical realizations of discourse connectives in the PDiT.

The annotators also marked the semantico-pragmatic types of relations and the extent of discourse arguments (i.e. the two units of a text that are interlinked with a discourse connective). The discourse relations could be annotated between two adjacent sentences as well as between long-distant arguments or even between the whole paragraphs. The annotation captures both intra- and inter-sentential relations.

The evaluation of annotation was focused especially on two aspects. The inter-annotator agreement on the existence of a discourse relation expressed by a second-

dary connective was 0.70 F1 and the agreement on the semantico-pragmatic types of relations was 82% (0.78 of Cohen's κ).

In the present monograph, the main results of the annotation are described in several larger parts covering frequency of connectives (Section 6.3.1), semantico-pragmatic relations they express (Section 6.3.2) and the opposition of intra- vs. inter-sentential relations (Section 6.3.3).

Within the analysis of connective frequency, we present the proportions of primary and secondary connectives in the PDiT. We have found out that secondary connectives covered 5.4% within all connective occurrences, i.e. the appearance of primary connectives in texts was much more common (PDiT contained 20,255 of primary connectives and 1,161 of secondary connectives). The existence of the free connecting phrases in the PDiT was even lower (151 of occurrences).

We present also twenty most frequent discourse connectives in the PDiT. These are: *a* (and), *ale* (but), *když* (when), *protože* (because), *totiž* (that is), *pokud* (if), *;*, *proto* (therefore), *tedy* (so), *aby* (so that), *pak* (then), *ovšem* (however), *-li* (if), *také* (too), *neboť* (for), *–*, *zatímco* (while), *nebo* (or), *navíc* (moreover). All of them are primary connectives (among primary connectives, we count also two punctuation marks: colon and dash). It turned out that also the individual primary connectives were more frequent than the secondary ones.

We have compared the frequency of secondary connectives in Czech with similar expressions in English. The comparison has been carried out on the data of the Penn Discourse Treebank (PDTB) capturing so-called alternative lexicalizations of discourse connectives. These expressions are comparable to secondary connectives and free connecting phrases together. The results demonstrate that PDTB contained 3.3% of alternative lexicalizations and PDiT 6.1% (of secondary connectives plus free connecting phrases). Primary connectives thus prevailed both in Czech and English.

We have compared also the existence of selected secondary connectives in written and spoken language (on the data of the PDiT and corpus DIALOG). We have found out that selected secondary connectives were more frequent in written Czech. This may be due to the fact that many secondary connectives (like *příčinou bylo* – *the cause was*, *z tohoto důvodu* – *for this reason*) are rather formal and appear thus more typically in written texts. Written Czech also exhibited greater lexical diversity than spoken Czech (the written texts contained more diverse lexical realizations of secondary connectives). Secondary connectives in written and spoken texts also differed in semantico-pragmatic relations. Whereas secondary connectives were used for semantically “stronger” relations in written texts, the use of secondary connectives in spoken texts was sometimes rather vague (cf. especially the structure *to souvisí s* – *this is related to* and its variants that were often used for a weak connection of two unrelated topics). Such use of secondary connectives is rather pragmatic than semantic, and in this sense, the spoken texts exhibited higher intentional vagueness.

In the next step, we have evaluated annotation of discourse connectives in terms of semantico-pragmatic relations they expressed. We found out that primary and secondary connectives expressed the individual relations with similar frequency – the

most frequent relations for both were conjunction, reason and result, condition, concession, asynchronous relations (precedence and succession) and confrontation. On the other hand, they differed especially in the relation of opposition – the second most frequent relation for primary connectives but rather infrequent for secondary ones.

The results have also demonstrated that primary and secondary connectives differ in their preference of intra- vs. inter-sentential relations. Primary connectives occur mostly (in 70%) in intra-sentential relations while secondary connectives rather in inter-sentential relations (in 63%). This fact is related to the part-of-speech and syntactic characteristics of the individual connectives. For example, whereas subordinating and basic coordinating conjunctions prefer intra-sentential usage, adverbs (as well as adverbial phrases and conjunctions with an adverbial origin) are used typically inter-sententially.

Contextual (in)dependence of discourse connectives and factors influencing the use of connectives in texts

The final part of results deals with possibilities of measuring contextual independence of selected discourse connectives. We have carried out an experiment with selected connective expressions and we demonstrate how it is possible to delimit discourse connectives as a class based on their suitability for various contexts. Then we have taken selected primary and secondary connectives of result (e.g. *proto – therefore, tak – so, z tohoto důvodu – for this reason, to je důvod, proč – that is the reason why...*) and we have tested them in a set of 100 contexts from the PDiT (containing the relation of result). We have looked at contexts where the insertion of a selected connective was not possible, i.e. on potential language constraints on discourse connectives (both primary and secondary). These constraints (or factors influencing the use of connectives) cover grammar (coordination vs. subordination, intra- vs. inter-sentential usage and word order), semantics, pragmatics, stylistics, and salience (connected with the use of connectives over a long-distant argument).

In the monograph, we demonstrate that discourse connectives are not a closed class of expressions – as a whole, they are constantly under development. They form a wide range of expressions – from fully grammaticalized to lexically variable structures. In this sense, we may imagine them as a set of expressions that has its centre (primary connectives) and its periphery (secondary connectives). The research conducted on a large language material may help improve understanding of relations within a text and thus enhance general awareness of the overall text coherence.

Seznam obrázků

3.1	Hierarchie abstraktních objektů podle Ashera (1993: 57)	26
3.2	Ukázka anotace textů podle teorie RST (Mann, Thompsonová, 1988: 252) . .	46
4.1	Rozvržení počtu slovních jednotek na jednotlivých anotačních rovinách v Pražském závislostním korpusu (viz Hajič a kol., 2006)	52
4.2	Ukázka anotace konektorů v Pražském diskurzním korpusu 1.0, viz příklad 39: <i>Elektrina v bytech funguje pouze několik hodin denně. Poprvé od počátku války minulý týden srbská strana navíc zablokovala přívod plynu.</i>	54
4.3	Složení korpusu SYN2013PUB podle titulů	56
6.1	Ukázka dvouúrovňové anotace sekundárních konektorů v Pražském diskurzním korpusu na souvětí: <i>Počet zaměstnaných osob klesl loni o dvě procenta, což znamená, že v Maďarsku loni pracovaly asi čtyři milióny osob.</i>	83
6.2	Škála kontextové (ne)závislosti konektivních prostředků ve formě předložkových frází se zájmenem <i>ten</i>	112
6.3	Škála kontextové (ne)závislosti konektivních prostředků ve formě předložkových frází s předložkou <i>na základě</i>	113
6.4	Škála kontextové (ne)závislosti primárních a sekundárních konektorů	115

Seznam tabulek

4.1	Množství dat v Pražském závislostním korpusu (viz Hajič a kol., 2006)	52
6.1	Charakteristika primárních a sekundárních konektorů	75
6.2	Mezianotátorská shoda pro anotaci sekundárních konektorů v PDiT 2.0 . . .	85
6.3	Srovnání mezianotátorské shody pro anotaci primárních a sekundárních konektorů v Pražském diskurzním korpusu	85
6.4	Srovnání výsledků anotace primárních a sekundárních konektorů v PDiT . .	86
6.5	Diskurzní konektory a volné konektivní fráze v PDiT	87
6.6	Diskurzní konektory s nejvyšším počtem výskytů v PDiT	88
6.7	Srovnání anotace diskurzních vztahů v češtině a angličtině	89
6.8	Struktury obsahující vybraná jádrová slova v PDiT a korpusu DIALOG	90
6.9	Shrnující výsledky pro vybrané konektivní struktury v korpusech PDiT a DIALOG	91
6.10	Typy diskurzních vztahů vyjádřené primárními konektory v PDiT	94
6.11	Typy diskurzních vztahů vyjádřené sekundárními konektory v PDiT; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$. .	95
6.12	Typy diskurzních vztahů vyjádřené primárními a sekundárními konektory v Pražském diskurzním korpusu – procentuální vyjádření	97
6.13	Sekundární konektory typu <i>výjimkou je</i> v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$	99
6.14	Podíl mezivětných a vnitrovětných vztahů vyjádřených primárními konektory v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$	104

6.15 Mezivětnost vs. vnitrovětnost primárních konektorů v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$)	105
6.16 Nejfrekventovanější primární konektory vyskytující se v souřadných konstrukcích v PDiT preferující vnitrovětné užití; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$)	106
6.17 Nejfrekventovanější primární konektory vyskytující se v souřadných konstrukcích PDiT preferující mezivětné užití; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$)	107
6.18 Podíl mezivětných a vnitrovětných vztahů vyjádřených sekundárními konektory v Pražském diskurzním korpusu; statisticky významný rozdíl (měřeno pomocí chí-kvadrát testu) je označen *** pro hladinu významnosti 0,001 (p-hodnota $\leq 0,001$), ** pro $p \leq 0,01$, * pro $p \leq 0,05$	108
6.19 Konektivní prostředky ve formě předložkových frází se zájmenem <i>ten</i>	111
6.20 Konektivní prostředky ve formě předložkových frází s předložkou <i>na základě</i>	113
6.21 Kontextová (ne)závislost vybraných primárních a sekundárních konektorů	114

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
PDiT – Prague Discourse Treebank, Pražský diskurzní 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

A

Přílohy

A.1 Seznamy konektorů

V následujících přehledech jsou uvedeny frekvenční a abecední seznamy primárních a sekundárních konektorů anotovaných v Pražském diskurzním korpusu. V seznamech jsou zachycené jednotlivé formy konektorů tak, jak se objevily v textech PDiT, tj. neuvádíme zde lemmata či základní tvary konektorů, ale všechny konkrétní varianty, které se v textech vyskytly (např. *na straně jedné_na straně druhé, na jedné straně_na druhé straně*). 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 diskurzních 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 PDiT 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 diskurzního vztahu se v PDiT 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 dovršování gramatikalizace, nicméně způsob jejich zápisu ještě kolísá.

Konektory jsou v seznamech rozděleny na primární a sekundární, zároveň jsou zde (už bez explicitního rozlišení) zahrnuty i konektory složené⁴¹ 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*⁴²..., konektory složené ze tří složek: *a i proto, a poté i, a tedy i, i tak ale, i tak však, – dokonce i...*), jsou řazeny

⁴¹ V seznamu jsou zahrnuty i korelativní páry, jsou zapisovány s podtržítkem (*bud_nebo*).

⁴² 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*.

mezi primární konektory. Konektory složené z primárních a sekundárních konektorů (např. *a pokračuje, a kromě toho, 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. Ze seznamů zároveň vyplývá, že modifikované primární konektory (na rozdíl od modifikovaných sekundárních konektorů) se v 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 typu *k tomu, na to, s čímž, za což* analogické k jednoslovným konektorům typu *proto, přitom, zato, předtím* (u kterých navíc v některých případech v textech PDiT 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 obtížné zařaditelnou skupinu tvoří příslovce typu *vzápětí, nakonec, současně, následně* nebo *posléze* (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 příslovčí 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.

A.1.1 Abecední seznam forem primárních konektorů z Pražského diskurzního korpusu

–	#neg_jen_ale především	#neg_ba naopak
– a	#neg_jen_ale také	#neg_i
– a nakonec	#neg_jen_ale též	#neg_jako
– a přitom	#neg_jen_nýbrž	#neg_jen
– a to	#neg_jen_především	#neg_jen i
– a zároveň	#neg_jenom_ale i	#neg_jenom
– aby	#neg_nic_jen	#neg_mnohem spíše
– což	#neg_pouze_ale	#neg_naopak
– dokonce	#neg_pouze_ale i	#neg_nýbrž
– dokonce i	#neg_pouze_i	#neg_pouze
– leda	#neg_pouze_nýbrž	#neg_přitom
– např.	#neg_sice_však	#neg_spíš
– například	#neg_a naopak	#neg_spíše
– pokud	#neg_a pouze	#neg_však
– proto	#neg_a spíš	:
– přesto	#neg_a spíše	: aby
– přitom	#neg_a spíše jenom	: např.
– tedy	#neg_ale	: například
– totiž	#neg_ale daleko spíše	: ovšem
– tzn.	#neg_ale i	: proto
– vždyť	#neg_ale jen	: přece
#neg ⁴³	#neg_ale pouze	: tedy
#neg a	#neg_ale především	: totiž
#neg ani_jako spíš ⁴⁴	#neg_ale spíš	a
#neg jen_a	#neg_ale spíše	a ani
#neg jen_ale	#neg_ale také	a bezprostředně poté
#neg jen_ale i	#neg_ale zato	a dále

⁴³ #neg je značka pro větnou negaci.

⁴⁴ Korelativní páry jsou zapisovány s podtržítkem (*bud_nebo*).

a dokonce	a taky	ale stejně
a dokonce i	a tedy	ale také
a hlavně	a tedy i	ale třeba i
a hned pak	a teprve potom	ale však
a i	a též	ale zároveň
a i proto	a tím	ale zato
a ještě	a tím i	alespoň
a konečně	a to	aneb
a nadále	a to i	anebo
a nakonec	a to jen	ani
a naopak	a to ještě	ani ne_ale spíše
a naopak vzápětí	a to proto	ani ne_jako spíše
a například	a tudíž	ani_ani
a navíc	a vzápětí	aniž
a ne	a zároveň	aniž_tak
a nebo	a zase	aspoň
a nikoli	a zejména	ať
a nikoliv	a zrovna tak	ať již
a pak	a_b	ať už
a pokud	aby	ať už_-li
a posléze	aby nakonec	avšak
a posléze i	aby posléze	avšak jen
a potažmo	aby tak	avšak pouze
a poté	ač	až
a poté i	ačkoli	až když
a potom	ačkoliv	až na
a později	ale	až nakonec
a proto	ale ani	až poté
a přece	ale dokonce	až poté, kdy
a především	ale i	až tehdy, až
a předtím	ale i tak	ba
a přesto	ale jen	ba co víc
a při tom	ale nakonec	ba i
a přímo	ale naopak	bud_anebo
a přitom	ale například i	bud_či
a rovněž	ale ne	bud_nebo
a současně	ale přece	budto
a stejně	ale přece jen	budto_nebo
a stejně tak	ale především	byť
a tak	ale přesto	což
a také	ale přitom	což ale
a také proto	ale spíše	což ovšem

což však	jednak_a dále	když_když
čemuž	jednak_a hlavně	když_tak
či	jednak_a jednak	koneckonců
či spíše	jednak_jednak	leč
čili	jednak_především	leč přece jen
čímž	však	ledaže
dál	jednak_také	-li
dále	jelikož	-li_jestliže
dále pak	jen	-li_pak
dále_a nakonec	jen když	-li_potom
dokonce	jen tehdy_-li	-li_tak
dokonce ani	jen však	mezi tím
dokonce aniž	jenom	mezitím
dokonce i	jenomže	mimo jiné
dokonce i když	jenže	mimo to
dokud	jenže i potom	mimoto
dokud_tak	jenže také	mj.
dříve, než	jestli	nadále
hned vzápětí	jestli pak	nadto
i	jestli_tak	nakonec
i ale	jestliže	nakonec totiž
i kdyby	jestliže potom	naopak
i kdyby_tak	jestliže_pak	naopak však
i když	jestliže_tak	naprosto ne
i když pouze	ještě	např.
i když přesto	ještě předtím	napřed_a teprve potom
i nebo spíše právě	ještě spíše	napřed_pak
i přesto	ježto	například
i přesto, že	jinak	například, když
i tak	jinak však	následně
i tak ale	již dříve	natož
i tak však	kdežto	natož pak
i tehdy, kdy	kdežto naopak	navíc
i však	kdyby	ne
jak	kdyby_pak	ne však
jak_tak	kdyby_potom	ne_ale
jakkoli	kdyby_tak	ne_nýbrž
jakkoliv tak	kdýkoli	ne_pouze
jakmile	kdýkoliv	neb
jako	když	nebo
jako vždy, když	když předtím	nebo alespoň
jednak_a	když už	nebo dokonce

nebo i	nejprve	pokud_tak
nebo ještě	nejprve_a pak	poněvadž
nebo naopak	nejprve_a poté	posléze
nebo pouze	nejprve_dále	poté
nebo spíš	nejprve_pak	poté, co
nebo spíše	nejprve_poté	poté, kdy
neboli	nejprve_potom	potom
neboť	nejprve_později	potom, co
nejdřív_pak	nejprve_vzápětí	potom, kdy
nejdřív_potom	nejprve_záhy	pouze
nejdříve_pak	než	pouze tehdy, pokud
nejdříve_potom	nicméně	pouze tehdy_-li
nejen	nikoli	pouze_nikoli
nejen navíc i	nikoli pouze	právě proto
nejen, že_ale zejména	nikoli tedy	pro to, aby
nejen:	nikoli však	proto
nejen_ale	nikoli_ale	proto také
nejen_ale ani	nikoliv	proto, aby
nejen_ale dokonce	nikoliv jenom	proto, že
nejen_ale i	nikoliv ovšem	proto_-li
nejen_ale naopak	nikoliv však	protože
nejen_ale především	nikoliv_ale	protože pokud
nejen_ale přímo	nikoliv_nýbrž	protože_tak
nejen_ale rovněž	nýbrž	přece
nejen_ale rovnou	obdobně pak	přece jen
nejen_ale také	ostatně i	přece jenom
nejen_ale též	ovšem	před tím
nejen_ale zároveň	ovšem pouze	před tím, než
nejen_nýbrž i	ovšem přeci jen	především pak
nejenom i	ovšem také	především pokud
nejenom_ale i	ovšemže	především však
nejenom_ale také	pak	předtím
nejenom_ale zejména	pak ovšem	předtím, než
nejenomže_ale	pak totiž	přesto
nejenže	pakliže	přesto ale
nejenže navíc	po tom, co	přesto však
nejenže_ale	po tom, kdy	přesto, že
nejenže_ale ani	pokud	přestože
nejenže_ale i	pokud ovšem	při tom
nejenže_ale ještě	pokud potom	příčemž
nejenže_ale také	pokud_-li	příčemž i
nejenže_rovněž	pokud_pak	příčemž ještě

příčemž navíc	stejně tak ale	vlastně totiž
příčemž zároveň	stejně tak i	však
přitom	tak	však i
přitom jen	tak ale	však mezitím
přitom navíc	tak např.	však naopak
přitom také	tak například	však navíc
přitom však	tak tedy	však nejen
přitom zase	tak, aby	však pouze
resp.	také	však přece i
respektive	také proto	však přeci jen
rovněž	taky	však přesto
sice	takže	však současně
sice_a	takže nakonec	však také
sice_ale	takže například	však zároveň
sice_ale aspoň	takže vlastně zase	vzápětí
sice_ale i tak	teda	vždyť
sice_ale pouze	tedy	vždyť také
sice_ale přesto	tehdy	za to
sice_ale taky	tehdy, až	záhy poté
sice_ale zároveň	tehdy, jestliže	zároveň
sice_ale zato	tehdy, když	zároveň ale
sice_avšak	tehdy, pokud	zároveň naopak
sice_jenže	tehdy_-li	zároveň však
sice_jenže také	tehdy_pokud	zas
sice_nicméně	teprve	zase
sice_nicméně však	teprve poté	zase naopak
sice_ovšem	teprve potom	zatím
sice_přesto	těž	zatím ale
sice_přesto však	tím	zatímco
sice_však	tím ovšem	zatímco naopak
sice_zato	tím, že	zato
snad jen	tj.	zato ale
sotva	totiž	zato však
souběžně	totiž navíc	zčásti_zčásti
současně ale	třeba	zda
současně však	třebaže	zdali
spíš	tudíž	zejména
spíše	už	zejména pokud
spíše však	už jen	zprvu_později
stejně jako	už ovšem	zvláště když
stejně tak	už předtím	

A.1.2 Frekvenční seznam forem primárních konektorů z Pražského diskurzního korpusu

a	příčemž	jinak
však	a proto	nejen_ale i
ale	či	a tedy
když	poté, co	a tudíž
protože	i	jelikož
totiž	jestliže	nakonec
pokud	dokonce	pokud_tak
:	#neg	a ne
proto	nicméně	a přitom
tedy	jenže	a dokonce
aby	avšak	a ani
pak	a pak	a naopak
ovšem	aniž	jen
-li	potom	i kdyby
také	a také	#neg_spíše
neboť	ačkoli	např.
–	vždyť	-li_pak
zatímco	zase	tím, že
nebo	sice_však	a navíc
navíc	až	a potom
přitom	ačkoliv	předtím
i když	než	nejen_ale
což	a i	pro to, aby
sice_ale	ani	#neg_pouze
naopak	dokud	třeba
takže	jakmile	tudíž
a tak	pokud_pak	zároveň však
a to	poté	třebaže
dále	#neg_nýbrž	vzápětí
kdyby	zato	přesto však
tak	ač	a poté
rovněž	#neg_jen	stejně tak
proto, že	byť	ne_ale
přesto	ať	nýbrž
#neg_ale	anebo	sice_ovšem
přestože	přesto, že	mezitím
například	bud_nebo	ať už
zároveň	čímž	: totiž

jestli	přece	#neg_spíš
pouze	ještě	a nikoliv
a zároveň	bud_anebo	a později
ne	jednak_jednak	a současně
posléze	a stejně tak	a stejně
čili	tehdy_-li	a taky
ale zároveň	a posléze	ale i tak
a ještě	a přece	ale přece
kdežto	#neg a ⁴⁵	i tak
tehdy, když	#neg_jen_ale i	jestli pak
– totiž	#neg_jen i	jestliže_tak
ale také	#neg_jenom	kdyby_pak
těž	a hlavně	když_když
tj.	a nakonec	leč
a tím	a vzápětí	mimoto
#neg_však	a zase	mj.
kdykoliv	ale ani	nadto
sice_avšak	ani_ani	ne však
jestliže_pak	čemuž	nebo alespoň
jak	i přesto	nejen_ale především
předtím, než	jak_tak	nejenže_ale i
kdyby_tak	jenom	nejprve_potom
nikoli	jestli_tak	potom, co
jenomže	následně	právě proto
jakkoli	nebo dokonce	proto také
při tom	nejdříve_pak	proto, aby
ale přesto	nejen_ale dokonce	přesto ale
a přesto	nejprve_pak	sice
sotva	ovšem také	sice_ale přesto
poněvadž	před tím, než	sice_nicméně
-li_tak	stejně jako	sice_zato
však také	tím	souběžně
mimo jiné	však i	současně však
nejen_ale také	však přesto	spíše
a rovněž	však zároveň	zas
poté, kdy	#neg_jen_ale ⁴⁶	– a ⁴⁷
taky	#neg_ale jen	– dokonce
ale i	#neg_naopak	– tedy

⁴⁵ Od této značky k další mají všechny formy konektorů v PDiT shodně 4 výskyty a jsou řazeny abecedně.

⁴⁶ Od této značky k další mají všechny formy konektorů v PDiT shodně 3 výskyty a jsou řazeny abecedně.

⁴⁷ Od této značky k další mají všechny formy konektorů v PDiT shodně 2 výskyty a jsou řazeny abecedně.

#neg jen_ale především	nejenom_ale i	– přitom
#neg jen_ale také	nejenže	– tzn.
#neg_a naopak	nejenže_ale	– vždyť
#neg_ale spíše	nejenže_ale také	#neg ani_jako spíš
: např.	nejprve_a pak	#neg jen_a
: tedy	nejprve_a poté	#neg jen_ale též
a dokonce i	nikoli však	#neg jen_nýbrž
a konečně	nikoli_ale	#neg jen_především
a například	nikoliv_nýbrž	#neg jenom_ale i
a nebo	pakliže	#neg nic_jen
a potažmo	po tom, co	#neg pouze_ale
a především	pokud potom	#neg pouze_ale i
ale ne	pouze tehdy_-li	#neg pouze_i
ale především	přičemž ještě	#neg pouze_nýbrž
ale přitom	přitom však	#neg sice_však
ale zato	resp.	#neg_a pouze
aneb	sice_ale zato	#neg_a spíš
až když	sice_přesto	#neg_a spíše
až nakonec	spíš	#neg_a spíše jenom
ba	tak, aby	#neg_ale daleko spíše
budťo	tehdy, pokud	#neg_ale i
dále pak	už	#neg_ale pouze
dokonce i	už předtím	#neg_ale především
dokud_tak	však pouze	#neg_ale spíš
dříve, než	zároveň ale	#neg_ale také
i přesto, že	zejména pokud	#neg_ale zato
i tehdy, kdy	zprvu_později	#neg_ba naopak
jednak_a	– a nakonec ⁴⁸	#neg_i
jen když	– a přitom	#neg_jako
kdyby_potom	– a to	#neg_mnohem spíše
když_tak	– a zároveň	#neg_přitom
koneckonců	– aby	: aby
ledaže	– což	: například
mimo to	– dokonce i	: ovšem
nadále	– leda	: proto
natož	– např.	: přece
natož pak	– například	a bezprostředně poté
ne_nýbrž	– pokud	a dále
nejen	– proto	a hned pak
nejen_ale rovněž	– přesto	a i proto

⁴⁸ Od této značky k další mají všechny formy konektorů v PDiT shodně 1 výskyt a jsou řazeny abecedně.

a nadále	avšak pouze	ještě spíše
a naopak vzápětí	až na	ježto
a nikoli	až poté	jinak však
a pokud	až poté, kdy	již dříve
a posléze i	až tehdy, až	kdežto naopak
a poté i	ba co víc	kdykoli
a předtím	ba i	když předtím
a při tom	bud' či	když už
a přímo	budto_nebo	leč přece jen
a také proto	což ale	-li_jestliže
a tedy i	což ovšem	-li_potom
a teprve potom	což však	mezi tím
a též	či spíše	nakonec totiž
a tím i	dál	naopak však
a to i	dále_a nakonec	naprosto ne
a to jen	dokonce ani	napřed_a teprve potom
a to ještě	dokonce aniž	napřed_pak
a to proto	dokonce i když	například, když
a zejména	hned vzápětí	ne_pouze
a zrovna tak	i ale	neb
a_b	i kdyby_tak	nebo i
aby nakonec	i když pouze	nebo ještě
aby posléze	i když přesto	nebo naopak
aby tak	i nebo spíše právě	nebo pouze
ale dokonce	i tak ale	nebo spíš
ale jen	i tak však	nebo spíše
ale nakonec	i však	neboli
ale naopak	jakkoliv tak	nejdřív_pak
ale například i	jako	nejdřív_potom
ale přece jen	jako vždy, když	nejdříve_potom
ale spíše	jednak_a dále	nejen navíc i
ale stejně	jednak_a hlavně	nejen, že_ale zejména
ale třeba i	jednak_a jednak	nejen:
ale však	jednak_především	nejen_ale ani
alespoň	však	nejen_ale naopak
ani ne_ale spíše	jednak_také	nejen_ale přímo
ani ne_jako spíše	jen tehdy_-li	nejen_ale rovnou
aniž_tak	jen však	nejen_ale též
aspoň	jenže i potom	nejen_ale zároveň
ať již	jenže také	nejen_nýbrž i
ať už_-li	jestliže potom	nejenom i
avšak jen	ještě předtím	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
 přitom jen
 přitom navíc
 přitom také
 přitom 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

A.1.3 Abecední seznam forem sekundárních konektorů z Pražského diskurzního korpusu

– ostatně	ač ve skutečnosti
– prostě	ale místo toho
– vlastně	ale na druhé straně
#neg_místo toho	ale o to víc
#neg_raději	ale stejně
#neg_rozhodně	ale to neznamená, že
#neg_to jenom	ale ve skutečnosti
a bude pokračovat	anebo ještě přesněji
a do třetice	anebo přesněji
a dodal jsem	ani v tomto případě
a dodejme	argumentoval
a doplňuje	argumentoval tím, že
a chtělo by se dodat	argumentovali tím, že
a jako důvod uvedla, že	argumentovalo tím, že
a ještě k tomu	argumentuje, že
a k tomu	argumentují tím, že
a k tomu ještě	argumentují, že
a kromě toho	avšak na druhé straně naopak
a kvůli tomu	bude pokračovat
a měly k tomu pochopitelné důvody	byla podmínkou_v opačném případě
a na závěr	což neznamená, že
a naproti tomu	což znamená, že
a pokračoval	což znamenalo, že
a pokračovali	dále dodal
a pokračovat bude	dám příklad
a pokračuje	díky tomu
a stejně	díky tomu, že
a tím pádem	dlužno dodat, že
a to je důvod, proč	do třetice
a to znamená, že	dodal
a v důsledku toho	dodala
a v tomto případě	dodává
a ve skutečnosti	dodávám
a vlastně	dodejme
a z toho důvodu	dodejme rovněž
a z tohoto důvodu	doplnil
a za druhé	doplňuje
aby pokračovala	dospěl k závěru, že

došli zkrátka k závěru, že
 důsledkem je
 důvod je jednoduchý
 důvod je složitější
 důvod je zřejmý
 důvod spočívá v tom, že
 důvodem bylo to, že
 důvodem bylo, že
 důvodem je
 důvodem je skutečnost, 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 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 neposlední 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 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
 zaprvé_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
 znamenalo to, že

A.1.4 Frekvenční seznam forem sekundárních konektorů z Pražského diskurzního 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 neznámá, 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 pokračuje⁴⁹
což znamená, že
do třetice
jak dodal
k tomu
kvůli tomu
od toho, aby
oproti tomu
pokračoval
tím spíš, že
to však neznámá, že
v tom případě
a kromě toho⁵⁰
ale ve skutečnosti
argumentují tím, že
což neznámá, že
doplňuje
důsledkem je
jak je vidět
jinými slovy
k čemuž
kromě toho, že
navzdory tomu, že
podmínkou je, aby

⁴⁹ Od této značky k další mají všechny formy konektorů v PDiT shodně 4 výskyty a jsou řazeny abecedně.

⁵⁰ Od této značky k další mají všechny formy konektorů v PDiT shodně 3 výskyty a jsou řazeny abecedně.

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_raději⁵¹
 a pokračoval
 a z tohoto důvodu
 ale na druhé straně
 ale to neznamená, 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 neznamená, 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
 – ostatně⁵²
 – 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

⁵¹ Od této značky k další mají všechny formy konektorů v PDiT shodně 2 výskyty a jsou řazeny abecedně.

⁵² Od této značky k další mají všechny formy konektorů v PDiT shodně 1 výskyt a jsou řazeny abecedně.

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
 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
 nedosti na tom
 nejen kromě toho
 nemluvě o tom, kdy
 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, 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 verzi 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 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
 zaprvé_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

A.1.5 Frekvenční seznam forem primárních a sekundárních konektorů z Pražského diskurzního korpusu s ohledem na vyjadřované významové vztahy

Disjunktivní alternativa: primární konektory

nebo	a nebo
či	a ⁵³
bud'_nebo	bud'_či
anebo	bud'to_nebo
jinak	i
bud'_anebo	nebo ještě
nebo alespoň	nebo naopak

Disjunktivní alternativa: sekundární konektory

případně	popřípadě
----------	-----------

Ekvivalence: primární konektory

tedy	:
čili	a tak
tak	aneb
tj.	– tzn.
a	a tedy i
–	a tím
takže	neboli
a tedy	potom
tudíž	proto
což	

Ekvivalence: sekundární konektory

prostě	srozumitelně řečeno
jinak řečeno	srozumitelněji řečeno
jinými slovy	upřesnil
	z toho vyplývá, že

⁵³ Od šedé značky dále mají v této příloze všechny formy konektorů v daném významovém vztahu v PDiT shodně 1 výskyt a jsou řazeny abecedně.

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

Explicace: primární konektory

totiž
 :
 –
 vždyť
 protože
 když
 neboť
 : totiž
 poté, co

přece
 proto
 proto, že
 – totiž
 přece
 aby
 nakonec totiž
 takže

Explicace: 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
dokonce
nejen_ale i
a dokonce
a navíc
nejen ale
a ještě
však
#neg jen
přitom
a to
nejen_ale také
a hlavně
nejen_ale dokonce

i
#neg jen_ale i
nebo dokonce
a i
nejenže_ale i
nadto
ale
#neg jen_ale
ani
nejen_ale především
ale i
nejenže
a především
nejenom ale i

#neg jen_ale především
 ovšem
 #neg jen_i
 – dokonce
 ale také
 ale ani
 nejenže_ale také
 dokonce i
 nejen
 natož
 nejenže_ale
 nejen_ale rovněž
 natož pak
 ale především
 #neg jen_ale také
 ba
 a
 – a to
 – dokonce i
 #neg jen_ale též
 #neg jen_nýbrž
 #neg jenom
 #neg jenom_ale i
 #neg pouze_ale
 #neg pouze_ale i
 #neg pouze_i
 #neg pouze_nýbrž
 #neg_ale i
 #neg_ale také
 #neg_i
 :
 a dokonce i
 a to i
 a to jen
 a to ještě
 ale dokonce
 ale například i
 ale spíše
 ale třeba i
 avšak

ba co víc
 ba i
 či
 či spíše
 dokonce ani
 dokonce aniž
 jednak_a hlavně
 jednak_především však
 ještě
 koneckonců
 nebo pouze
 nejen, že_ale zejména
 nejen_ale ani
 nejen_ale naopak
 nejen_ale přímo
 nejen_ale rovnou
 nejen_ale též
 nejen_ale zároveň
 nejen_navíc i
 nejen_nýbrž i
 nejenom i
 nejenom_ale také
 nejenom_ale zejména
 nejenomže_ale
 nejenže_ale ani
 nejenže_ale ještě
 nejenže_navíc
 nejenže_rovněž
 pokud
 přičemž
 přičemž ještě
 přičemž navíc
 přitom navíc
 však i
 však navíc
 však nejen
 však současně
 však také
 vždyť
 zato

Gradace: sekundární konektory

kromě toho
nedosti na tom
nehledě na to, že
a ještě k tomu
a k tomu
a k tomu ještě
k tomu

kromě toho ovšem i
kromě toho, že
ne dosti na tom
nejen kromě toho
nemluvě o tom, kdy
nemluvě o tom, že
o to víc
tím spíše, že

Konfrontace: primární konektory

zatímco
naopak
však
ale
zase
ovšem
a naopak
zato
a
kdežto
přitom
jenže
nicméně
sice_ale
avšak
když
jestliže
-li
–
jinak
a ne
a přitom
a zrovna tak
ale naopak
aniž

i když
jakkoliv_tak
jestliže_tak
kdežto naopak
nakonec
ne
nikoli
ovšem také
přestože
přičemž
sice však
sice zato
spíše však
tak
už
už jen
však mezitím
však naopak
zároveň naopak
zároveň však
zas
zase naopak
zatímco naopak
zato ale
zato však

Konfrontace: sekundární konektory

naproti tomu
 na druhé straně
 oproti tomu
 na druhé straně však
 na druhou stranu
 rozdíl je v tom, že
 ve skutečnosti
 a naproti tomu
 ale na druhé straně
 avšak na druhé straně naopak
 místo toho
 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 první pohled však
 na rozdíl od toho
 na straně druhé
 na straně jedné_a na straně druhé
 s rozdílem, že
 s tím kontrastuje skutečnost, že
 sice_ale na druhé straně
 v opačném případě
 v prvním případě_ve druhém případě
 v tom, že zatímco
 však na druhé straně

Konjunkce: primární konektory

a
 také
 což
 dále
 rovněž
 přitom
 přičemž
 pak
 i
 navíc
 zároveň
 a také
 a to
 –
 a i
 ani
 však
 zase
 aniž
 čímž
 současně
 :

a ani
 ale
 stejně tak
 když
 též
 ovšem
 či
 a stejně tak
 a rovněž
 a pak
 dokonce
 taky
 jednak_jednak
 ani_aní
 čemuž
 stejně jako
 ještě
 aby
 ale také
 a tak
 jak tak
 a přitom

a taky
 mimoto
 jinak
 ovšem také
 jednak_a
 a navíc
 však také
 dále pak
 tedy
 – a
 mimo to
 potom
 a zase
 – a nakonec
 a dále
 a dokonce i
 a hned pak
 a ještě
 a nakonec
 a potom
 a přímo
 a případně i
 a současně i
 a též
 a zároveň
 a_b
 ale ani
 ale přitom
 dál
 i ale

jako
 jednak_a dále
 jednak_a jednak
 jednak_také
 jestliže
 mimo jiné
 nebo
 nejen ale i
 obdobně pak
 ostatně i
 podobně i
 především však
 při tom
 přičemž i
 přičemž ještě
 přičemž zároveň
 přitom také
 resp.
 sice
 sice_a
 stejně tak ale
 stejně tak i
 tak
 tak tedy
 totiž
 však i
 vždyť také
 zas
 zatímco
 zčásti_zčásti

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 neposlední 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

budťo
–
nebo i

Konjunktivní alternativa: sekundární konektory

případně	popřípadě
----------	-----------

Nepravá podmínka: primární konektory

pokud	:
jestliže	-li_tak
-li	jestli_pak

Nepravá podmínka: sekundární konektory

v případě, že

Nepravá příčina a důsledek: primární konektory

totiž	:
tedy	a tedy
proto	aby
takže	neboť
tak	proto, že
protože	přece

Nepravá příčina a důsledek: sekundární konektory

kvůli tomu, že	to znamená, že
o důvod víc	

Nepravý kontrast: primární konektory

však	ať už
ale	avšak
ovšem	i když přesto
sice_ale	jenže
i když	přitom
nicméně	sice_ale zároveň

Nepravý kontrast: sekundární konektory

Neobsahuje.

Nesoučasnost (předčasnost a následnost): primární konektory

pak	zprvu_později
když	a bezprostředně poté
poté, co	a konečně
a pak	a naopak vzápětí
potom	a posléze i
než	a poté i
až	a předtím
poté	a teprve potom
jakmile	aby
nakonec	aby nakonec
později	aby posléze
předtím	až když
a potom	až poté, kdy
a poté	hned vzápětí
vzápětí	jak
dokud	jak tak
posléze	ještě předtím
předtím, než	již dříve
sotva	kdykoli
a posléze	když předtím
a vzápětí	mezitím
nejprve_pak	napřed_a teprve potom
poté, kdy	napřed_pak
před tím, než	nejdřív_pak
kdykoliv	nejdřív_potom
nejdříve_pak	nejdříve_potom
následně	nejprve
a později	nejprve záhy
a	nejprve_dále
a nakonec	nejprve_poté
potom, co	nejprve_později
nejprve_potom	nejprve_vzápětí
až nakonec	po tom, kdy
už předtím	potom, kdy
dále	před tím
nejprve_a pak	tak
nejprve_a poté	tehdy, když
po tom, co	teprve
nadále	teprve poté
dříve, než	záhy poté

Nesoučasnost (předčasnost a následnost): sekundární konektory

pokračoval
 pokračuje
 krátce nato
 a bude pokračovat
 a na závěr
 a pokračovat bude
 aby pokračovala
 bude pokračovat
 jako poslední
 má pokračovat

mají pokračovat
 následoval poté, co
 následovat bude
 od okamžiku, kdy
 pokračovala
 pokračovat by mělo
 pokračují
 s tím, že pak
 v tom pokračují
 v závěru

Opozice: primární konektory

však
 ale
 ovšem
 sice_ale
 jenže
 avšak
 nicméně
 sice_však
 přitom
 aniž
 a
 zato
 a přitom
 zároveň však
 naopak
 ale zároveň
 sice_ovšem
 sice_avšak
 i když
 jenomže
 nýbrž
 přičemž
 však zároveň
 přesto
 leč
 a ne

a naopak
 ale také
 přesto však
 přitom však
 sice_ale zato
 ale i
 ale přesto
 sice_zato
 však také
 ale přece
 a přesto
 –
 sice_přesto
 a zase
 což
 však i
 zároveň ale
 ale zato
 současně však
 a současně
 – přitom
 #neg jen_a
 #neg pouze
 #neg_ale
 #neg_jenom
 : ovšem

a to
ale ani
ale i tak
ale ne
ale přece jen
což ale
což ovšem
čímž
i však
jenže také
jinak však
leč přece jen
naopak však
nikoli však
nikoliv
nikoliv ovšem
ovšem také
pak
přece jenom
přestože

rovněž
sice
sice_ale aspoň
sice_ale i tak
sice_ale přesto
sice_ale taky
sice_jenže
sice_jenže také
sice_nicméně
sice_nicméně však
sice_přesto však
tak ale
tedy
už
už ovšem
však přece i
však přesto
zas
zatímco

Opozice: sekundární konektory

na druhé straně
na druhou stranu
na druhé straně však
ale místo toho
ale na druhé straně
ale o to víc
argumentovali tím, že
dlužno dodat, že
na druhé straně ovšem
na druhé straně zase
na jednu stranu
naproti tomu

navzdory tomu však
nutno ovšem dodat
o to víc
sice_ale ve skutečnosti
současně ale
v čemž
v této souvislosti
ve skutečnosti
však nutno dodat
však ve skutečnosti
z jedné strany_ze strany druhé
a stejně

Podmínka: primární konektory

pokud
-li
když
kdyby

jestliže
pokud_pak
pokud_tak
dokud

-li_pak
 jestli
 jakmile
 kdyby_tak
 jestliže_pak
 pak
 tehdy_když
 jak
 potom
 tehdy_-li
 i kdyby
 jestli_tak
 -li_tak
 až
 kdyby_pak
 ale
 a pak
 i když
 ať
 zejména pokud
 a
 pakliže
 tehdy_pokud
 jestli_pak
 pokud_potom
 jestliže_tak
 jen když
 když_tak
 pouze tehdy_-li
 –
 – leda
 – pokud

a pokud
 a také
 aniž
 ať už
 ať už_-li
 až poté
 až tehdy, až
 dokud tak
 i tehdy, kdy
 jako vždy, když
 jen tehdy_-li
 jestliže_potom
 jinak
 kdyby_potom
 kdyby_pak
 když už
 -li_potom
 ovšem
 pokud ovšem
 pouze tehdy_pokud
 protože
 protože pokud
 především pokud
 při tom
 tak
 tehdy
 tehdy, až
 tehdy, jestliže
 teprve potom
 zdali
 zvláště když

Podmínka: sekundární konektory

v případě, že
 za podmínky, že
 v takovém případě
 v tomto případě
 v tom případě
 podmínkou je, aby
 s tím, že

v případě, kdyby
 v případech, kdy
 s podmínkou, že
 byla podmínkou_v opačném případě
 jedním z těchto důvodů je
 skutečnost, že

jen v případě, že
jinými podmínkami bylo, že
podmínka je to nevyhnutelná
podmínkou je
podmínkou je, že
pouze v případě, když
pouze v případě, že
pouze v těch případech, kdy
s jednou podmínkou

stejně
v opačném případě
v případě
v případě toho, že
v případě, kdy
v tom případě, že
ve spojení s tím, že
za podmínky, kdyby

Příčina a důsledek: primární konektory

protože
proto
totiž
neboť
tedy
takže
a tak
proto, že
a proto
tak
:
–
když
vždyť
a tudíž
jelikož
pak
a
tím, že
a tedy
tudíž
aby
– totiž
a tím
poněvadž
: totiž
tím
proto také
právě proto
přece

a potažmo
čili
-li
– proto
– vždyť
a i proto
a také proto
a tím i
a to
a to proto
ale
což
čímž
dále
i
jestliže
ježto
jinak
koneckonců
například
neb
pak ovšem
pak totiž
poté, co
poté, kdy
právě proto, že
proto, aby
protože_tak
přitom
také proto

takže nakonec
totiž navíc

za to

Příčina a důsledek: sekundární konektory

vzhledem k tomu, že
za to, že
z tohoto důvodu
díky tomu
díky tomu, že
v této souvislosti
z toho důvodu, že
tím pádem
s tím, že
to znamená, že
tím spíše, že
kvůli tomu
na to, aby
k tomu, aby
ze dvou důvodů
tím spíše
s odůvodněním, že
upřesnil
z toho důvodu
důsledkem je
a z tohoto důvodu
důvodem je skutečnost, že
ze stejného důvodu
z jednoduchého důvodu
z těchto důvodů
je to z toho důvodu, že
výsledkem je, že
zásluhou toho, že
stejně

a jako důvod uvedla, že

a kvůli tomu
a měly k tomu pochopitelné důvody
a tím pádem
a to je důvod, proč
a to znamená, že
a v důsledku toho

a v tomto případě
a vlastně
a z toho důvodu
argumentoval
argumentovali tím, že
což znamená, že
což znamenalo, 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
 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
 přesto
 přestože
 ale
 však
 ačkoli
 přitom
 ačkoliv
 ač
 byť
 přesto, že
 ať
 nicméně
 i kdyby

třebaže
 ať už
 přesto však
 sice_ale
 aniž
 ovšem
 když
 jakkoli
 a přece
 i přesto
 a přesto
 ale přesto
 i tak
 však přesto

a přitom
 sice_ovšem
 avšak
 přesto ale
 sice však
 a
 ale i tak
 sice_ale přesto
 jenže
 i přesto, že
 i přes to, že
 a to
 – a přitom
 – přesto
 a současně
 ale nakonec
 ale přece
 ale přitom
 ale zároveň
 aniž tak

ať již
 což však
 dokonce i když
 i
 i kdyby_tak
 i přesto, že
 i tak ale
 i tak však
 jestliže
 jinak
 kdyby
 poté, co
 přece jen
 přitom zase
 sice
 sice_nicméně
 současně však
 tedy
 tím ovšem
 zároveň však
 zatím ale

Přípustka: sekundární konektory

to neznamená, že
 to však neznamená, že
 to ovšem neznamená, že
 což neznamená, že
 navzdory tomu, že
 to ale neznamená, že
 ale to neznamená, že
 a stejně
 a ve skutečnosti
 ač ve skutečnosti
 ale stejně
 ale ve skutečnosti
 ani v tomto případě

argumentovalo tím, že
 fakt neznamená, že
 i přes tato fakta
 když ve skutečnosti
 na rozdíl od toho, co
 navzdory tomu
 neznamená to, že
 o to více
 přes tyto skutečnosti
 to ale znamenalo, že
 v této souvislosti nezbyvá dodat, že
 však navzdory těmto faktům
 vzhledem k tomu, že
 z faktu nevyplývá, že

Rektifikace: primární konektory

#neg_ale
 #neg
 #neg_nýbrž
 #neg_jen
 #neg_spíše
 ale
 a ne
 #neg_pouze
 ne_ale
 ne
 #neg_však
 naopak
 nýbrž
 nikoli
 #neg_a
 spíše
 a nikoliv
 #neg_spíš
 #neg_naopak
 tedy
 ne však
 #neg_ale jen
 spíš
 však
 nikoli_ale
 #neg_ale spíše
 ne_nýbrž
 #neg_jenom
 #neg_jen i
 #neg_a naopak
 nikoliv_nýbrž
 – tedy
 #neg –
 #neg_a spíš
 #neg_a spíše
 #neg_a spíše jenom
 #neg_ale daleko spíše

#neg_ale pouze
 #neg_ale především
 #neg_ale spíš
 #neg_ale zato
 #neg_anijako spíš
 #neg_ba naopak
 #neg_jako
 #neg_mnohem spíše
 #neg_přitom
 #neg_sice_však
 a naopak
 a nikoli
 ale ne
 ani ne_ale spíše
 ani ne_jako spíše
 či
 i nebo spíše právě
 jen
 ještě spíše
 kdyby
 naprosto ne
 ne pouze
 nebo spíš
 nebo spíše
 nikoli pouze
 nikoli tedy
 nikoli však
 nikoliv ale
 nikoliv jenom
 nikoliv však
 pouze
 pouze_nikoli
 resp.
 respektive
 sice_ale
 teda
 totiž
 zato

Rektifikace: sekundární konektory

místo, aby
 místo toho
 ale ve skutečnosti
 místo toho, aby
 #neg_raději
 – vlastně
 #neg_rozhodně
 #neg_místo toho
 #neg_to jenom

anebo ještě přesněji
 anebo přesněji
 lépe řečeno
 místo toho však
 pravděpodobnější verzí je, že
 přesněji
 přesněji řečeno
 rozdílem je skutečnost, že
 s rozdílem, že
 ve skutečnosti

Restriktivní opozice: primární konektory

však
 ale
 ovšem
 jen
 jenže
 pouze
 nicméně
 i když
 jinak
 sice_ale
 avšak
 přitom
 jenom
 však pouze
 ledaže
 než
 –
 a
 a to
 ale jen
 ale však

alespoň
 aniž
 aspoň
 avšak jen
 avšak pouze
 i když pouze
 jen však
 jenže i potom
 když
 ne
 ovšem pouze
 ovšem přeci jen
 přesto však
 přitom jen
 sice_ale pouze
 sice_nicméně
 sice_ovšem
 sice_však
 snad jen
 však přeci jen

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

přitom

tj.

tedy

a nadále

a přitom

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

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 ANNO-DIS 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. *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer Academic Publishers, 1993.

ASHER, Nicholas; LASCARIDES, Alex. *Logics of Conversation*. Cambridge University Press, 2003. ISBN 9780521659512.

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.

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*, 1962a, 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*, 1962b, s. 29–37.
- BAUER, Jaroslav. Spojky. *Český jazyk 8*, 1958, s. 220–226.
- BAUER, Jaroslav; GREPL, Miroslav. *Skladba spisovné češtiny*. Praha: Státní pedagogické nakladatelství, 1980.
- 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 re-formulation. *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.

- CARLSON, Lynn; MARCU, Daniel; OKUROWSKI, Mary Ellen. Building a discourse-tagged corpus in the framework of rhetorical structure theory. *Current and new directions in discourse and dialogue*. Dordrecht: Springer, 2003, s. 85–112.
- 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.
- CRUSE, Alan D. *Meaning in Language: an Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press, 2000.
- 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.
- 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 posloupností 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 Texebene. 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.
- DANLOS, Laurence; RYSOVÁ, Kateřina; RYSOVÁ, Magdaléna; STEDE, Manfred. Primary and secondary discourse connectives: definitions and lexicons. *Dialogue & Discourse* 9 (1), 2018, s. 50–78.
- 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. 2018-08-18]. 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 aktualního členění větného*. Brno: Masarykova univerzita, 1978.
- FISCHER, Kerstin. Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume. In *Approaches to discourse particles*. Amsterdam: Elsevier, 2006, s. 1–20.
- 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.

HAJIČ, Jan a kol. *Prague Dependency Treebank 2.0*. Philadelphia: Linguistic Data Consortium, 2006.

HAJIČ, Jan a kol. *Prague Dependency Treebank 3.5*. Data/software, Univerzita Karlova, MFF, ÚFAL, Praha, 2018.

HAJIČOVÁ, Eva. *Issues of Sentence Structure and Discourse Patterns*. Praha: Charles University, 1993.

HAJIČOVÁ, 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.

HAJIČOVÁ, Eva. Contextual boundness and discourse patterns revisited. *Discourse Studies* 15 (5), 2013, s. 535–550.

HAKULINEN, Auli. 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. 2018-08-18]. Dostupné z WWW: <<http://www.etymonline.com>>.

HARRIS, Zellig Sabbettai. Discourse Analysis. *Language* 28 (1), 1952, s. 1–30.

HARRIS, Zellig Sabbettai. *Methods in Structural Linguistics*. Chicago: University of Chicago Press, 1951.

HAUSENBLAS, Karel. Gramatická syntax textu. In *Referati od X. zasedanie na slovenskí literaturní 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. *Tra-vaux linguistiques de Prague* 1, 1964, s. 67–83.
- 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. *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 Organisation des Textes. In KOŘENSKÝ, J.; HOFFMANNOVÁ, J. (eds.). *Text and pragmatic aspects of language*, 1984, s. 101–140.
- HOLUB, Josef; KOPEČNÝ, František. *Etymologický slovník jazyka českého*. Praha: Státní nakladatelství učebnic v Praze, 1952.
- 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. K problematice syntaktických vztahů parataktických. *Naše řeč* 74 (4–5), 1991, s. 181–188.
- HRBÁČEK, Josef. *Nárys textové syntaxe spisovné češtiny*. Praha: Trizonia, 1994. ISBN 80-85573-51-2.
- HRBÁČEK, Josef. Text, promluva a nadpromluvvé 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. *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. Nejčastější konektivní prostředky kauzálního vztahu v Pražském závislostním korpusu. *Studie z aplikované lingvistiky* 1–2, 2012, s. 35–52.

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.

JOHANNSSEN, 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.

KARLÍK, Petr. Juxtapozice. In KARLÍK, P. a kol. (eds.). *CzechEncy – Nový encyklopedický slovník češtiny*. 2017 [cit. 2018-08-18]. Dostupné z WWW: <<https://www.czechency.org/slovník/JUXTAPOZICE>>.

KARLÍK, Petr; ŠKRABALOVÁ, Hana Gruet. Slučovací vztah. In KARLÍK, P. a kol. (eds.). *CzechEncy – Nový encyklopedický slovník češtiny*. 2017 [cit. 2018-08-18]. Dostupné z WWW: <https://www.czechency.org/slovník/SLUČOVACÍ_VZTAH>.

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.
- MATHESIUŠ, Vilém. Jazykozpytné poznámky k řečnické výstavbě souvislého výkladu. *Slovo a slovesnost* 7, 1943, s. 114–129.
- MATHESIUŠ, Vilém. O tak zvaném aktuálním členění větěném. *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.

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.

PETÖFLI, 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.

PETÖFLI, János Sándor. *Per una teoria semiotica integrata del testo. Dalla TeSWeST alla Testologia Semiotica*. Carabba, 2011. ISBN-10 8863442037. ISBN-13 978-8863442038.

PETÖFLI, János Sándor. *Scrittura e interpretazione. Introduzione alla Testologia Semiotica dei testi verbali*. Roma: Carocci, 2004. ISBN 8843031333.

PETÖFLI, János Sándor. Towards an Empirically Motivated Grammatical Theory of Verbal Texts. *Studies in Text Grammar* 19, 1973, s. 205–275.

PETÖFLI, 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; MÍROVSKÝ, Jiří; NEDOLUZHKO, Anna; JÍNOVÁ, Pavlína; ZIKÁNOVÁ, Šárka; HAJIČOVÁ, Eva. 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; JÍNOVÁ, Pavlína; ZIKÁNOVÁ, Šárka; BEDŘICHOVÁ, Zuzanna; MÍROVSKÝ, Jiří; RYSOVÁ, Magdaléna; ZDENKOVÁ, Jana; PAVLÍKOVÁ, Veronika; HAJIČOVÁ, Eva. *Manual for Annotation of Discourse Relations in Prague Dependency Treebank*. Technická zpráva č. 2012/47, Praha: ÚFAL MFF UK, 2012a.

POLÁKOVÁ, Lucie; JÍNOVÁ, Pavlína; ZIKÁNOVÁ, Šárka; HAJIČOVÁ, Eva; MÍROVSKÝ, Jiří; NEDOLUZHKO, Anna; RYSOVÁ, Magdaléna; PAVLÍKOVÁ, Veronika; ZDENKOVÁ, Jana; PERGLER, Jiří; OCELÁK, Radek. *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. 2018-08-18]. 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*. Marrackech, 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.

REJZEK, Jiří. *Český etymologický slovník*. Praha: LEDA, 2001.

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. Discourse Connectives: From Historical Origin to Present-Day Development. In *New perspectives on cohesion and coherence: Implications for translation*. Berlín: Language Science Press, 2017, s. 11–34. ISBN 978-3-946234-72-2.

RYSOVÁ, Magdaléna. Jazykové prostředky vyjadřující textové vztahy v češtině a jejich zpracování v Pražském závislostním korpusu. *Bohemistika* 13 (1), 2013, s. 57–71. ISSN 1642–9893.

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; 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.

RYSOVÁ, Magdaléna; RYSOVÁ, Kateřina. Secondary Connectives in the Prague Dependency Treebank. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, Uppsala, Sweden: Uppsala University, 2015, s. 291–299. ISBN 978-91-637-8965-6.

RYSOVÁ, Magdaléna; RYSOVÁ, Kateřina. Primary and secondary discourse connectives: Constraints and preferences. *Journal of Pragmatics* 130, 2018, s. 16–32. ISSN 0378-2166.

RYSOVÁ, Kateřina; RYSOVÁ, Magdaléna. Discourse Connectives and Reference. In *TextLink – Structuring Discourse in Multilingual Europe – Final Action Conference*. Toulouse, France: University of Toulouse, 2018a, s. 122–128.

RYSOVÁ, Kateřina; RYSOVÁ, Magdaléna. The Correlation between Discourse-Anaphoric Devices and an Overall Communicative Competence in Learners' Essays. In *EDULEARN18 Proceedings*. Valencia, Spain: IATED Academy, 2018b, s. 2144–2154. ISBN 978-84-09-02709-5, ISSN 2340-1117.

- 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* 29 (1), 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, s. 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*. 3., opravené vyd. Praha: Academia, 2003.
- 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.

SVOBODA, Karel. K třídění vedlejších vět podle spojovacích výrazů. *Slovo a slovesnost* 29, 1968, s. 142–151.

SVOBODA, Karel. Věty se spojkou *aniž*. *Naše řeč* 55, 1972, s. 97–103.

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. *Perspectives* 8, 1998a, 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*, Praha, 1998b.

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. *Journal of Pragmatics* 3, 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*. Tübingen: Niemeyer, 1980b.

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), 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 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; HAJIČOVÁ, Eva; HLADKÁ, Barbora; JÍNOVÁ, Pavlína; MÍROVSKÝ, Jiří; NEDOLUZHKO, Anna; POLÁKOVÁ, Lucie; RYSOVÁ, Kateřina; RYSOVÁ, Magdaléna; VÁCL, Jan. *Discourse and Coherence. From the Sentence Structure to Relations in Text*. Praha: ÚFAL, 2015a. ISBN 978-80-904571-8-8.
- ZIKÁNOVÁ, Šárka; POLÁKOVÁ, Lucie; JÍNOVÁ, Pavlína; NEDOLUZHKO, Anna; RYSOVÁ, Magdaléna; MÍROVSKÝ, Jiří; HAJIČOVÁ, Eva. Zachycení výstavby textu v Pražském závislostním korpusu. *Slovo a slovesnost* 76 (3), 2015b, s. 163–197.
- ZIPF, Georg Kingsley. *Human Behavior and the Principle of Least Effort*. Cambridge: Addison-Wesley Press, 1949.
- 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

Elektronický slovník staré češtiny. Ústav pro jazyk český AV ČR, v. v. i., oddělení vývoje jazyka, Praha, 2006–2018. Dostupné z WWW: <<http://vokabular.ujc.cas.cz>>.

Korpus DIALOG 1.1. Ústav pro jazyk český, AV ČR, Praha, 2012. Dostupné z WWW: <<http://ujc.dialogy.cz>>.

KŘEN, Michal; HNÁTKOVÁ, Milena; JELÍNEK, Tomáš; PETKEVIČ, Vladimír; PROCHÁZKA, Pavel; SKOUMALOVÁ, Hana. *SYN2013PUB: korpus psané publicistiky*. Ústav Českého národního korpusu FF UK, Praha, 2013. Dostupné z WWW: <<http://www.korpus.cz>>.

RYSOVÁ, Magdaléna; SYNKOVÁ, Pavlína; MÍROVSKÝ, Jiří; HAJIČOVÁ, Eva; NEDOLUZHKO, Anna; OCELÁK, Radek; PERGLER, Jiří; POLÁKOVÁ, Lucie; PAVLÍKOVÁ, Veronika; ZDEŇKOVÁ, Jana; ZIKÁNOVÁ, Šárka. *Prague Discourse Treebank 2.0*. Data/software, Univerzita Karlova v Praze, MFF, ÚFAL, Praha, 2016. Dostupné z WWW: <<http://hdl.handle.net/11234/1-1905>>.

Rejstřík

- abstraktní objekty, 54
- argumenty, *viz* diskurzní argumenty
- autosémantika, 66, 74
- diskurz, 8, 15, 17, 19–21
- diskurzní argumenty, 26, 67
 - explicitní argumenty, 102
 - implicitní argumenty, 102
 - nominální argumenty, 27, 68, 92, 97
 - situační argumenty, 102
 - slovesné argumenty, 27, 53, 68
 - vzdálené argumenty, 118
- diskurzní konektory, 7, 23, 37, 38, 41, 44, 60
 - anaforické konektory, 77, 123
 - kataforické konektory, 77, 123
 - mezivětné konektory, 106, 124
 - modifikované konektory, 79, 124
 - primární konektory, 61, 62, 67, 86, 93, 105, 113, 123
 - sekundární konektory, 61, 63–67, 80, 86, 93, 106, 110, 113, 123
 - složené konektory, 77–79, 123
 - vnitrovětné konektory, 105, 124
- diskurzní vztahy, 24
 - mezivětné vztahy, 27, 103
 - vnitrovětné vztahy, 27, 103
 - významové vztahy, 28, 32, 37, 93
- explicitní vztahy, 7, 9, 28
- frazémy, 63
- frekvence, 86, 87, 114
- gramatikalizace, 61, 66, 73
- historický vývoj, 72, 74
- hypersyntax, 16
- idiomy, 63
- implicitní vztahy, 7, 9, 28, 29
- intenční vágnost, 93
- jazyková ekonomie, 119
- juxtapozice, 101
- jádrová slova, 61, 66, 88
- koherence, 14, 95
- koheze, 14
- komunikace, 14, 15, 116
- komunikát, 17
- kondenzace, 92
- konektory, *viz* diskurzní konektory
- kontext, 60, 76, 109
 - kontextová nezávislost, 60, 110, 111, 113
- korelativní páry, 78
- lexikální rozmanitost, 91
- makrosyntax, 16
- mežianotátorská shoda, 83
- mluvená čeština, 90, 93
- modifikovatelnost, 65
- nadvětná syntax, 16
- nadvětné jevy, 15, 16
- podřadnost, 116
- pragmatika, 93, 96, 118
- Pražský diskurzní korpus, 10, 24, 32, 53, 55, 86
- Pražský závislostní korpus, 51, 53
- princip maximální srozumitelnosti, 120
- princip minimality, 26, 54
- princip minimálního úsilí, 119
- promluva, 17
- reference, 76
- salience, 118
- slovosled, 117

- souřadnost, 116
- spojky, 61
 - podřadící spojky, 61, 62
 - souřadící spojky, 61, 62
- stylistika, 118
- synonymie, 120
- synsémantika, 66, 74
- syntaktické fráze, 64

- tematické předěly, 103
- teorie koherence, 22
- teorie relevance, 23
- text, 13, 17
- textová gramatika, 16
- textová lingvistika, 13–16
- textová syntax, 16

- variabilita, 66
 - gramatické varianty, 66
 - lexikální realizace, 66
- volné konektivní fráze, 76, 82, 86, 87, 110
- význam, 23, 118
 - pojmový význam, 23, 38
 - procedurální význam, 23, 38