

Primární a sekundární diskurzní konektory

MAGDALÉNA RYSOVÁ – KATEŘINA RYSOVÁ

Primary and secondary discourse connectives

ABSTRACT: In the paper, we study discourse connectives and other discourse phenomena in Czech based on corpus data. We focus on the delimitation of connectives primarily from the functional point of view and we present their further classification into primary and secondary connectives based especially on their degree of grammaticalization. Special attention is paid to the variability of secondary connectives occurring in several lexical realizations and grammatical variants. We also present the frequency of the individual connectives in written texts and we analyze their syntactic behavior. We then discuss the relation of connectives to reference (a special group consists of the anaphoric and cataphoric connectives), connectives occurring next to each other in the text, and the possibilities of modification of connectives. Our analysis is based especially on the language data of the Prague Discourse Treebank 2.0. The aim of the paper is to present a detailed study of discourse connectives in written Czech and to present connectives as an open class covering both grammaticalized expressions and looser lexical phrases.

Key words: discourse connectives, primary connectives, secondary connectives, connective means, cohesion, coherence, text relations, Prague Discourse Treebank

Klíčová slova: diskurzní konektory, primární konektory, sekundární konektory, spojovací prostředky, koheze, koherence, vztahy v textu, Pražský diskurzní korpus

1. Úvod

Text – psaný nebo mluvený – je základním prostředkem lidské komunikace. Jak bylo v literatuře již mnohokrát popsáno (např. Halliday – Hasan, 1976), schopnost rozpoznat ucelený text od náhodného shluku vět na sebe nenavazujících je přirozenou součástí jazykové kompetence. Zmíněná ucelenost textu spočívá v jeho významové soudržnosti a srozumitelnosti – jednotky či úseky textu na sebe obsahově navazují a jsou vzájemně významově propojeny. Koherentní text tak obsahuje komplexní síť významových vztahů. Ty mohou být buď implicitní, tj. čtenář či posluchač je s ohledem na specifické a obecné vědění vyrozumívá z obsahu daných výpovědí, nebo mohou být vyjádřeny explicitně pomocí konkrétních jazykových prostředků, zejména tzv. diskurzních konektorů. Přítomnost konektorů usnadňuje čtenáři či posluchači orientaci ve významové výstavbě textu a napomáhá interpretaci vztahů mezi textovými jednotkami.

V naší studii se zaměřujeme na výzkum diskurzních konektorů v psané češtině. Na základě analýzy korpusových jazykových dat přinášíme podrobný popis konektorů a jejich základní klasifikaci. Snažíme se přitom postihnout charakteristické chování jednotlivých typů konektorů v konkrétních (kon)textech.

V článku se nejprve věnujeme vymezení základních pojmů, se kterými téma diskurzních konektorů úzce souvisí: jedná se o pojmy text a diskurz, jednotka textu a vztahy mezi jednotkami textu (oddíl 2). Nejvíce prostoru je věnováno popisu diskurzních konektorů a jejich podrobnější klasifikaci (oddíl 2.4). Postupně se tak v článku zabýváme vymezením konektorů především z funkčního hlediska (oddíl 2.4.1), jejich tříděním na

základě stupně gramatikalizace (oddíly 2.4.2 a 2.4.3), vztahem konektorů a reference (oddíl 2.4.4) a souvýskytem několika konektorů (oddíl 2.4.5). Následně v článku představujeme jazykové korpusy, které jsme využily pro náš výzkum (oddíl 3). Díky možnostem počítačového zpracování jazykových jevů v korpusech je zpřístupněno velké množství dat, i když značkování konektorů a jiných diskurzních jevů je v korpusech češtiny na rozdíl od některých jiných jazyků zatím sporadické. Značkování diskurzních konektorů a sémanticko-pragmatických textových vztahů obsahuje v současné době jen Pražský diskurzní korpus, jehož jazyková data byla pro náš výzkum využita primárně. Doplnkově byla použita i data Českého národního korpusu, který obsahuje automatické značkování slovních druhů, a tedy částečně i spojovacích prostředků. V analytické části článku se zaměřujeme na diskurzní konektory z hlediska jejich frekvence v textech (oddíl 4.1), dále na nejčastější typy složených konektorů (oddíl 4.2) a na víceslovné konektory s fakultativní referencí (oddíl 4.3). Závěrečné shrnutí je uvedeno v oddíle 5.

2. Vymezení základních pojmů

2.1. *Text a diskurz*

Při zkoumání významových vztahů v textu a konektorů, které je vyjadřují, je předmětem našeho zájmu „nadvětný útvar“. Pro něj se využívá různých termínů, jako je komunikát, text či diskurz (srov. Hausenblas, 1984). Představení různých pojetí není cílem tohoto článku – diskuzi o pojmech text a diskurz a jejich bližší definici nabízí například Tárnyiková (2002). V našem příspěvku používáme termíny text i diskurz pro ucelený nadvětný útvar, který je dále dělitelný, tj. skládá se z menších jednotek. Podstatou jeho zkoumání je rozbor významových vztahů mezi těmito jednotkami a jazykových prostředků, které mají schopnost dané vztahy vyjádřit (tj. především diskurzních konektorů). Pojem text (textový) zde tedy chápeme jako synonymum k pojmu diskurz (diskurzní).

2.2. *Jednotka textu*

V obecné rovině považujeme za jednotku textu tzv. diskurzní argument (Prasad et al., 2008). Diskurzní argumenty jsou definovány jako abstraktní objekty (Asher, 1993), které mohou vyjadřovat například děje, události, stavy, jevy. Dále jsou vymežovány na základě tzv. principu minimality (Prasad et al., 2007), podle kterého do argumentu patří pouze výpovědi (nebo jejich části), které nesou informace nezbytně nutné pro interpretaci příslušného diskurzního vztahu. V příkladu (1) – zachyceném na obrázku 1 v oddíle 3 – najdeme dva diskurzní argumenty: prvním z nich je výpověď *Oba Turci si přepadení vymysleli* (Arg1), druhým je výpověď *Byli v těchto dnech z ČR vypovězeni* (Arg2).¹

¹ Konektory v příkladech zvýrazňujeme tučně. Příklady pocházejí zejména z Pražského diskurzního korpusu a z Českého národního korpusu, okrajově z internetu (některé příklady jsou kráceny).

- (1) Jak vyplynulo z vyšetřování, oba Turci si přepadení vymysleli.

Z tohoto důvodu byli v těchto dnech z ČR vypovězeni.

Diskurzí argumenty mohou být nehierarchizované jednotky, srov. příklad (2). Mezi první a druhou větou je vztah opozice vyjádřený konektorem *ale* (příklad 2a), mezi druhou a třetí je vztah důsledku vyjádřený konektorem *čili* (příklad 2b).

- (2) Chtěla jsem dosáhnout na medaili, ale vypadla jsem už v prvním kole. Čili zítra poletím domů.
(2a) (Chtěla jsem dosáhnout na medaili,)*_Arg1* **ale** (vypadla jsem už v prvním kole.)*_Arg2*
(2b) (Vypadla jsem už v prvním kole.)*_Arg1* **Čili** (zítra poletím domů.)*_Arg2*

Diskurzí argumenty mohou v textu tvořit také hierarchickou strukturu, v níž je jeden argument součástí jiného argumentu, srov. příklad (3). Mezi prvními dvěma větami je vztah konjunkce vyjádřený konektorem *a* (příklad 3a). Mezi prvním větným celkem (který v sobě zahrnuje oba předchozí argumenty) a druhým větným celkem je vztah důsledku vyjádřený konektorem *takže* (příklad 3b).

- (3) Už na prvním stupni jsem hrál na kytaru a od šesté třídy jsem chodil na hodiny zpěvu. Takže mám dobrou hudební průpravu.
(3a) (Už na prvním stupni jsem hrál na kytaru)*_Arg1* **a** (od šesté třídy jsem chodil na hodiny zpěvu.)*_Arg2*
(3b) (Už na prvním stupni jsem hrál na kytaru a od šesté třídy jsem chodil na hodiny zpěvu.)*_Arg1* **Takže** (mám dobrou hudební průpravu.)*_Arg2*

Z hlediska formy mají diskurzí argumenty nejčastěji podobu jednoduchých vět, hlavních vět, vedlejších vět nebo souvětí. V některých případech ale může být diskurzí argumentem i posloupnost několika větných celků nebo celý odstavec – srov. příklad (4), ve kterém spojuje konektivní prostředek *a kvůli tomu všemu* rozsáhlý diskurzí argument (*Voda tu nebyla [...] vítr nedovolil klasům více povyrůst.*) s jednovětným argumentem (*tato krajina měla právo na zázrak a na útěchu*). Mezi nimi je vztah důsledku. V jiných případech mohou mít diskurzí argumenty také podobu jmenných frází – srov. příklad (5), v němž je druhým argumentem spojení *stárnutí obyvatelstva*.

- (4) Voda tu nebyla, měli jen tu, kterou přinášel déšť. V létě sázeli rajčata do hlíny hluboké pár palců, níž byl už jen vápenec, a sazenice pak rostly jen ze slunce. Když dech nížiny vystoupal vysoko do hor, změnil se nad Astrasem v jinovatku, vrátil se a pokryl rostliny, které jako by věděly, že nemají času nazbyt, a tak své plody rodily předčasně a ve spěchu. Rodily spoustu malých plodů a pak se, smířené se svým osudem, skláněly k zemi a usychaly. Bez váhání padaly k zemi, aby ji obohatily. A hlína se tak stávala výživnější. Obilí v Astrasu nerostlo výš než na dvě dlaně – to byla výška, kterou umožnil zimní déšť, a později zase vítr nedovolil klasům více povyrůst. **A kvůli tomu všemu** měla tato krajina právo na zázrak a na útěchu.
(5) V roce 1993 klesl počet porodů na 748 500. **Důsledkem je** stárnutí obyvatelstva.

Z hlediska pozice v textu mohou stát diskurzí argumenty bezprostředně vedle sebe (příklad 6) nebo od sebe mohou být odděleny jiným textem, který není součástí ani jednoho z nich (takové případy jsou však málo časté), srov. diskurzí vztah důsledku (příklad 7) mezi podtrženými argumenty signalizovaný konektorem *z tohoto důvodu*.

- (6) Úrokové sazby jsou vysoké. Proto splatit úvěr za čtyři roky je pro většinu nových vlastníků nemožné.
- (7) Výše sazeb havarijního pojištění se podle ředitele Jakuba Hradce řídí pouze rizikovostí příslušné značky. Reagoval tak na dohady ohledně jejich tvorby. Z tohoto důvodu je podle něho i nižší pojistné na škodovky než třeba na mercedesy.

V některých případech může dojít také k tomu, že mezi dvěma úseky textu najdeme dva různé diskurzní vztahy zároveň. V příkladu (8) je mezi dvěma diskurzními argumenty vztah opozice vyjádřený konektorem *ale* i vztah důsledku vyjádřený modifikovaným konektorem *právě proto*.

- (8) Všichni jsme unavení. **Ale právě proto** se musíme jít provětrat na čerstvý vzduch.

Diskurzní argumenty jsou ve většině případů v psaných textech explicitně vyjádřeny, v některých kontextech však mohou být argumenty (či jejich části) pouze implicitní, viz v komunikaci běžně užívané textové elipsy typu *Kdy přijedeš? Abych tě stihla včas vyzvednout na nádraží. (Kdy přijedeš? Ptám se, impl. Arg **abych** tě stihla včas vyzvednout na nádraží.)*. V některých případech může být diskurzní argument vyjádřen situací (srov. také Rouchota, 1996), nemusí být tedy nutně textový. Diskurzní významový vztah tak může být mezi výpovědí a situačním kontextem.

2.3. Vztahy mezi diskurzními argumenty

Po diskuzi o diskurzních argumentech se budeme věnovat vymezení a klasifikaci textových vztahů, které se mezi těmito argumenty mohou vyskytovat. Koherentní text tvoří síť rozmanitých vztahů – kromě sémanticko-pragmatických (kterým se více věnujeme v oddíle 2.3.1) se na koherenci textu podílejí například také vztahy izotopie (Greimas, 1966) či vztahy spojené s aktuálním členěním větným.

Vztahy izotopie specifikuje Daneš (1985) jako vztahy identity, částečné identity a sémantické blízkosti předmětů řeči, které lze vykládat nejen z pohledu sémantické ekvivalence, ale také z pohledu koreference daných výrazů. V pozdějších pojetích se pro izotopické řetězce používají spíše termíny anafora a koreference (Hrbáček, 1994).

Textové vztahy spojené s aktuálním členěním větným reflektují informační strukturu vět s ohledem na jejich zasazení do kontextu (Mathesius, 1939, 1943; Daneš, 1968; Hajičová, 1975; Hajičová – Partee – Sgall, 1998; Sgall – Hajičová – Buráňová, 1980 aj.). F. Daneš (1968, 1979, 1985) v návaznosti na aktuální členění podrobně rozpracovává téma organizace vyšších textových celků, tzv. tematické posloupnosti.

2.3.1. Významové vztahy v textu

V naší práci se dále zaměříme pouze na mezipropoziční vztahy sémanticko-pragmatické, které lze z formálního hlediska dělit na vztahy vnitrovětné (vztahy uvnitř větného celku, tedy v souvětí souřadném i podřadném) a mezivětné (vztahy nad úroveň větného celku). Při analýze textových vztahů patří mezi stěžejní témata jejich sémantická klasifikace. Existuje celá řada lingvistických přístupů, které se ve vymezení významových

vztahů v textu v různé míře vzájemně odlišují. Předmětem diskuzí je celkový počet vztahů, které v textu můžeme najít, ale i jejich přesná interpretace (ke srovnání různých pojetí viz Blakemore, 2002).

Pro výzkum konektorů v psané češtině jsme využily především data Pražského diskurzního korpusu (PDiT), který rozlišuje 22 významových typů, viz tabulku 1. Příklady k jednotlivým typům vztahů viz Zikánová et al. (2015b). Sémantické vztahy jsou v PDiT rozděleny do čtyř základních kategorií – rozlišují se vztahy kontrastivní, navazovací, kauzální a časové, které jsou dále děleny do několika podtypů.

Tabulka 1: Přehled sémanticko-pragmatických vztahů mezi textovými jednotkami v Pražském diskurzním korpusu.

KONTRASTIVNÍ VZTAHY	konfrontace	opozice	restriktivní opozice	pragmatický kontrast	připustka	rektifikace	gradace
NAVAZOVACÍ VZTAHY	konjunkce	konjunktivní alternativa	disjunktivní alternativa	exemplifikace	specifikace	ekvivalence	generalizace
KAUZÁLNÍ VZTAHY	příčina – důsledek	pragmatická příčina – důsledek	explikace	podmínka	pragmatická podmínka	účel	
ČASOVÉ VZTAHY	současnost	nesoučasnost					

Tato klasifikace významových vztahů pro češtinu byla inspirována značkováním vztahů v diskurzním korpusu Penn Discourse Treebank pro angličtinu (Prasad et al., 2008), který rozbor textu zakládá na vyhledávání diskurzních konektorů a jejich lexikálním popisu. Takto sestavená klasifikace textových vztahů vychází především ze sémantiky diskurzních konektorů zasazených do konkrétního kontextu.

2.3.2. *Implicitní významové vztahy a juxtapozice*

Významové vztahy v textu mohou být signalizovány explicitním jazykovým prostředkem (nejčastěji konektorem, příklad 9) nebo mohou být implicitní (příklad 10). Typ implicitního vztahu si čtenář či posluchač vyvodí z obsahu propozic, popř. z širšího kontextu. Takto lze vyvozovat například vztahy důsledkové či příčinné. Existují ale typy vztahů, které se běžně v implicitní podobě nevyskytují, například vztahy připustkové nebo podmínkové – resp. vyskytují se zřídka, například v kontextech typu: *Chceš nechtěš, vyzkoušet tě musím. Nemyslíš, zaplatíš. Čiň čertu dobře, peklem se ti odmění.*

(9) Přes naše území přejde tlaková výše. Můžeme **proto** očekávat hezké počasí.

(10) Přes naše území přejde tlaková výše. Můžeme očekávat hezké počasí.

Implicitní významové vztahy se vyznačují tím, že mezi příslušné diskurzní argumenty je možné vložit vhodný konektor, který čtenáři usnadní interpretaci daného vztahu: *Přes naše území přejde tlaková výše. Můžeme (proto/tedy/tak ad.) očekávat hezké počasí.* Pokud vložení vhodného konektoru není možné, jedná se o (do jisté míry) volné kladení tematicky navzájem souvisejících vět vedle sebe (juxtapozici). V konkrétních (kon)textech ovšem bývá obtížné od sebe implicitní významové vztahy a juxtapozici odlišit, proto jim nyní budeme věnovat pozornost.

Vztahem juxtapozice se v českém prostředí zabývá řada autorů. Například Hrbáček (1994) chápe juxtapozici (spolu s junkcí a konexí) jako jednu ze základních možností uspořádání výpovědi v textu. Junkcí jsou výpovědi propojovány v souřadném souvětí, konexí výpovědi oddělené koncovým signálem (tj. jde o řazení nad úroveň větňého celku). Výpovědi kladené v textu vedle sebe, které nejsou ve vztahu junkce ani konexe, jsou ve vztahu juxtapozice. Hrbáček navrhuje nerozlišovat juxtapozici a asyndeton, protože je obtížné v praxi rozlišit, kdy jde v textu o elipsu spojovacího prostředku a kdy jde o spojení vět, mezi nimiž je nulový významový vztah. Na druhou stranu například Bauer a Grepl (1980, s. 232) charakterizují juxtapozici jako zmíněný nulový vztah, resp. asyndetické spojení vět bez zjevného vnitřního vztahu (*Jděte mi k šípku, kdopak by do takové houpačky sedal.*).

V našem pojetí chápeme juxtapozici jako významově vyprázdňený vztah mezi dvěma diskurzními argumenty v rámci jednoho tematického celku, resp. jedná se o vztah mezi diskurzními argumenty rozvíjejícími stejné téma. Ačkoliv mezi těmito argumenty není těsnější významový vztah (např. vztah opozice, příčiny nebo specifikace), jejich kladení vedle sebe neporušuje koherenci textu. Rozdíl mezi implicitním sémanticko-pragmatickým textovým vztahem a vztahem juxtapozice spočívá zejména v tom, zda je možné mezi diskurzní argumenty vložit vhodný konektor.

Konektory chápeme jako prostředky, které mají vlastní sémantiku (například *ale* vyjadřuje nejčastěji opozici, *protože* příčinu, *proto* důsledek). Je-li však diskurzní vztah mezi textovými jednotkami významově vyprázdňený, není možné mezi tyto jednotky vložit konektor, který ve své podstatě významově vyprázdňený není. Základní funkce konektoru v textu je totiž dvojí: spojovat diskurzní argumenty a zároveň signalizovat, který typ sémanticko-pragmatického vztahu mezi argumenty je. Pokud dva diskurzní argumenty nejsou v užším významovém vztahu, pak na tomto místě textu nelze užít ani diskurzní konektor.²

V konkrétním (kon)textu však může být rozhodnutí, zda konektor lze, či nelze vložit, obtížné. Diskurzní konektory mají totiž různou míru sémantického naplnění. Zatímco například konektor *protože* je významově vyhraněný (vyjadřuje příčinu, v některých případech explikaci), konektor *a* je obecnější. Objevuje se mezi propozicemi, které jsou ve vztahu prosté konjunkce, ale i mezi propozicemi se vztahem příčiny a důsledku, specifikace, gradace, konfrontace, opozice, předčasnosti a následnosti nebo současnosti. Tyto vztahy jsou však spíše než sémantikou spojky *a* vyjádřeny významy příslušných propozic (a tím se blíží k implicitním textovým vztahům). Spojka *a* je sémanticky poměrně „volný“ prostředek, a má proto v textech obecně širší možnosti uplatnění

² Na druhou stranu pokud je v textu na určitém místě užít diskurzní konektivní prostředek, je zde také třeba hledat sémanticko-pragmatický textový vztah. V některých případech se může stát, že pro interpretaci nalezeného vztahu v textu nebude stačit předem zvolená množina typů vztahů (zachycená například v tabulce 1), pak je třeba množinu rozlišovaných vztahů rozšířit. V současné době v tabulce vztahů pro PDiT není například rozlišován vztah zřetele nebo způsobu, zpracováváním dalšího materiálu lze pravděpodobně najít i jiné (pod)typy významových vztahů, o které by bylo možné tabulku obohatit.

(*a* je zdaleka nejfrekventovanější konektor) než prostředek sémanticky silně vyhraněný. Proto je v některých případech také těžké rozhodnout, zda mezi asyndeticky navázané propozice ještě lze, nebo už nelze vložit spojku *a* – *a* zda se tedy jedná o juxtapozici. Konektivní prostředek typu *a* (či jeho případné ekvivalenty) tak může představovat můstek mezi juxtapozicemi a sémanticko-pragmatickými vztahy, zejména konjunkcí.

2.3.3. Tematické řetězce

Od diskuze o významových vztazích přecházíme k úvahám o vyšší strukturaci textu. Textové jednotky (argumenty) propojené významovými vztahy či vztahy juxtapozice dohromady tvoří jeden tematický celek. Text se však může skládat z několika takovýchto celků, resp. tematických řetězců, jejichž začátky a konce jsou ohraničeny tematickými předěly. Na text lze tedy pohlížet jako na hierarchickou strukturu skládající se z jednoho či více tematických celků, přičemž v textu může přirozeně docházet k funkčnímu střídání témat či podtémat.

Tematické předěly (podobně jako sémanticko-pragmatické vztahy) jsou v textu signalizovány explicitním jazykovým výrazem (typu *dále*, *mimochodem*, *další věcí je*, *odjinud*, *jinak*) nebo jsou pouze implicitní. V takových případech bývají alespoň naznačeny například novým odstavcem či v mluveném projevu pomlkou. Explicitní změnu tématu signalizovanou výrazem *jinak* ukazuje příklad (11).

- (11) Dneska na cvičení nepůjdu, jsem nastydlá. **Jinak** tu knížku jsem ti přinesla, nechala jsem ti ji v práci na stole.

Konektivní výrazy typu *jinak* nebo *dále* tedy nemusí v textu nutně spojovat dva diskurzivní argumenty v určitém významovém vztahu, ale mohou být použity jako signalizátory tematických předělů, a tím se podílet na vyšší organizaci textu.

2.4. Diskurzivní konektory

Konektivní prostředky pomáhají čtenáři či posluchači při interpretaci textu. V některých případech lze diskurzivní konektor v textu vynechat, aniž bychom narušili celkovou koherenci. Pokud ale čtenář nebo posluchač dobře nerozumí obsahu propozic (například v odborném textu), pak sám z textu neumí vyčíst, v jakém vztahu propozice jsou (zda stojí například v opozici, či zda jedna vyplývá z druhé). Pro základní orientaci ve vztazích mezi propozicemi je pak textový konektor nezbytný, srov. příklad (12), v němž autor kromě vyjádření toho, který vztah v textu je, explicitně upozorňuje i na to, který vztah v textu není.

- (12) Nekovalentní interakce jsou mnohem slabší než interakce kovalentní. **Nikoliv přesto, ale právě proto** hrají tak významnou úlohu v procesech ovládajících živou hmotu.

Diskurzivní konektory jsou tedy jazykové prostředky, které napomáhají porozumění významovým vztahům v textu, a tím i porozumění textu jako takovému. V tomto ohledu hrají významnou roli a je zapotřebí věnovat prostor jejich vymezení a bližší specifikaci. V následující části nejprve uvedeme přední zahraniční a české studie zabývající se

diskurzními konektory a posléze představíme naše vlastní vymezení těchto výrazů pro češtinu.

Jednotlivé zahraniční přístupy k diskurzním konektorům se mezi sebou výrazně liší, a to jak v nastavení definic spojovacích prostředků, tak i v užívané terminologii – například van Dijk (1979) používá označení pragmatické konektory, Schiffrinová (1987) diskurzní ukazatele, Bazanella (1990) fatické konektory, termín diskurzní operátory najdeme u Redekerové (1991), pragmatické operátory užívá Arielová (1994), diskurzní konektory Blakemoreová (2002) a diskurzní částice Fischerová (2006). Na konektivní výrazy lze zároveň pohlížet z mnoha hledisek, například z hlediska jejich slovnědruhov-é příslušnosti, pozice ve větě, lexikální ustálenosti, fonologického chování, významu, pravdivostních podmínek či pragmatiky. Jednotliví autoři pak často v definicích kladou důraz na různé aspekty.

Zwicky (1985) například charakterizuje diskurzní konektory jako prozodicky nezávislé výrazy oddělené od okolního textu pauzami. Zároveň podle něj bývají konektory monomorfematické, syntakticky oddělené od zbytku věty a obvykle ve větě obsazují první pozici. Důležitou vlastností konektorů je, že v textu vždy plní pragmatickou funkci.

Podle Schiffrinové (1987) patří mezi diskurzní ukazatele široké spektrum prostředků včetně parajazykových jevů a neverbálních gest. Lexikální ukazatele (v naší terminologii konektory) autorka charakterizuje jako výrazy, které jsou z věty syntakticky vydělitelné a které se obvykle ve větě vyskytují na první pozici.

Fischerová (2006) chápe konektory jako „malá“, sémanticky, syntakticky a často i prozodicky nezapojená neohebná slova. Zároveň však upozorňuje na to, že tato charakteristika se vztahuje především na typické konektory a že je potřeba se zaměřit i na ostatní výrazy se spojovací funkcí.

Prasadová (viz Prasad et al., 2008, 2010) vyděluje dvě kategorie výrazů fungujících v textu jako ukazatele významových vztahů. První kategorií tvoří diskurzní konektory – jde o spojovací výrazy patřící mezi podřadící a souřadící spojky, předložkové fráze nebo adverbia. Ostatní výrazy vyjadřující významové vztahy (typu *after that*) zařazuje do kategorie označené jako alternativní lexikální vyjádření diskurzních konektorů.

Urgellesová-Collová (2010) chápe diskurzní konektory jako fonologicky krátká a redukováná slova, která jsou syntakticky nezapojená do větné struktury (a mohou být tedy z věty vypuštěna). Zároveň podle ní konektory většinou nepřispívají k pravdivostním podmínkám daných výpovědí.

V českém prostředí se způsoby propojování textu zabývá řada autorů. Termín konektor v lingvistice nebyl zprvu ukotven, proto se kromě diskuze o konektorech zaměříme také na diskuzi k vymezení a charakteristice spojek.

Výzkumem spojovacích prostředků se podrobně zabývá například Žaža (1958), který kategorii spojek charakterizuje velmi široce. Pro naši práci je významný především Žažův rozbor víceslovných výrazů, které autor považuje za samostatné spojky. Za jednu spojku tak pokládá například výrazy typu *i když*, *i kdyby*, *jen když* nebo *jen co*. Označuje je jako spojky složené. Jedná se o spojky skládající se z více výrazů, z nichž alespoň jeden již samostatně neplní funkci, kterou má v rámci složené spojky. Význam jednotlivých komponentů složených spojek se tak stírá.

Naopak Havránek a Jedlička (1960) termínem spojka označují jen jednoslovné výrazy, i když spojovací funkci připisují i víceslovným spojkovým výrazům (*i kdyby, a proto*). V rámci kategorie spojek rozlišují dále spojky jednoduché (*ale, když, že*) a složené (*zatímco*). Zároveň upozorňují na skutečnost, že spojkové výrazy mohou být i dvojité (*bud' nebo, jednak jednak*).

Spojovací prostředky se podrobně zabývají také Bauer a Grepl (1968, 1980 aj.). Podle jejich pojetí se propojení vět může realizovat pomocí spojovacích prostředků nebo asyndeticky. Mezi spojovací prostředky autoři řadí spojky, relativna, tázací slova a také slova s vlastním věcným významem – jde o výrazy, které jsou větnými členy, ale svým lexikálním významem zároveň vyjadřují vztah mezi větami (například *naopak*).

Jeden z nejučenějších popisů výrazů s konektivní funkcí v češtině přináší Hoffmannová (např. 1983 nebo 1984). Diskurzivní konektory vymezuje ze sémanticko-pragmatického hlediska. Neomezuje tuto skupinu příslušností k určitým slovním druhům, ale upřednostňuje vymezení na základě funkce výrazu v textu. Zdůrazňuje přitom, že každý konektor obsahuje prvek sémantický (jde o vztahovou sémantiku v rámci textu) i pragmatický (postoj autora k textu nebo postoj autora k adresátovi).

Také Daneš (1985) uplatňuje při popisu konektorů zejména funkční hledisko. V charakteristice konektorů navazuje například na Mistríka (1968), když zdůrazňuje, že konektory „nejen připojují, ale zároveň i modifikují význam podobně jako příslovce“ (Daneš, 1985, s. 172), tj. obsahují „jisté momenty povahy pragmatické a sémantické“ (ibid.). Konektory Daneš chápe jako relátory (sémantické predikáty), jejichž argumenty tvoří větné propozice (nebo jejich seskupení v podobě větných či souvětých bloků). Daneš (ibid., s. 171) naznačuje, že konektory tvoří širokou, dále dělitelnou skupinu výrazů – „široce pojaté pole konektorů (jakožto jazykových výrazů) má zřetelně oblasti centrální a periferní, obsahuje výrazy slovnědruhově i funkčně dosti různé, v jejichž funkčním potenciálu nalézáme vedle funkce spojovací v nestejně míře a v různé distribuci a dominanci mnohé momenty další“.

Užší pojetí konektorů nabízí Hrbáček (1994). Za konektory označuje jazykové prostředky, které spojují větné celky a které zároveň nemají větněčlenskou platnost. Pro výrazy vyjadřující vztahy uvnitř souřadného souvětí užívá termín junktory. Za konektory lze označit například spojky, některá příslovce, částice nebo předložkové a zájmenné výrazy. Hrbáček dále popisuje i výrazy, které mají zároveň spojovací a větněčlenskou funkci (*důsledkem je*), ty však mezi konektory neřadí.

S modelem dvou základních souvětých typů vybudovaným na komunikačně-pragmatickém základu přichází Karlík (1995). Rozlišuje souvětí založená na zapojování jedné větné propozice do jiné propozice a souvětí založená na spojování větných propozic. Na základě tohoto modelu pak ukazuje, že zatímco v prvním případě existuje uzavřený a málo početný repertoár konektorů (*že, aby, zda, zdali*), v případě druhém existuje v komunikaci potřeba vytvářet nové spojovací prostředky pro explicitní postihnutí sémantických a pragmatických rysů (například *na jedné straně – na druhé straně, ale stejně*). Při popisu spojek pak Karlík (2016) uplatňuje gramatické testy. Ukazuje například, že spojky souřadící se mohou kombinovat s jednotkami různého typu (mohou spojovat věty, fráze, slova), zatímco spojky podřadící jsou v tomto směru omezené: *Petr a Pavel* vs.

**Petr že Pavel*. Souřadící spojky nejsou zároveň považovány za součást žádné ze spojovaných jednotek, zatímco spojky podřadící jsou součástí věty, na jejímž levém kraji se typicky objevují: *Spi, nebo bdí – *Nebo bdí, spi vs. Tuší, že omdlí – Že omdlí, tuší*.

Obsáhlou studii o tzv. argumentativních konektorech v češtině a ve francouzštině přináší Pešek (2011). Ve své práci představuje model fungování argumentativních konektorů, rozebírá mj. jejich rysy stylistické či vlastnosti textově-kompoziční a syntaktické. Vybrané konektory analyzuje v rámci teorie argumentace v jazyce a teorie optimální relevance. Pro vlastní rozbor si autor vybírá několik skupin (do jisté míry) významově ekvivalentních konektorů, pro češtinu např. *protože, poněvadž, jelikož, neboť* či *ale, však, avšak, ovšem, nicméně*, a zaměřuje se na jejich charakteristiku a vzájemné odlišnosti.

Spojovací výrazy v češtině podrobně zkoumá také Štěpán (2013). Popisuje je jako jazykové prostředky, které mají syntaktickou funkci spojovat dvě věty v souvětí a které zároveň vyjadřují významový vztah mezi těmito větami. Spojovací výrazy dále vymezuje na základě jejich formy – jsou jimi zpravidla prostředky jednoslovné napříč některými slovními druhy. Štěpán navazuje na koncepci Svobody (1972), jehož výklad prohlubuje a doplňuje o závěry získané prací s rozsáhlým korpusovým materiálem. Popisuje například kontaktní postavení tzv. různovětných podřadících výrazů spojujících více než dvě věty: *Připadá mi, že když jsem já chodila na střední školu, byly děti jiné. – Připadá mi, že děti byly jiné, když jsem já chodila na střední školu*. Podrobněji viz Štěpán (1977).³

Rada autorů ve svých pracích charakterizuje hlavně „typické konektory“. Ze zahraničních autorů se např. Fischerová (2006) věnuje „nejtypičtějším zástupcům konektorů“, Urgellesová-Collová (2010) analyzuje „prototypické“ vlastnosti konektorů, Prasadová (Prasad et al., 2008, 2010) vyděluje dvě skupiny – konektory a jejich alternativní vyjádření. Na heterogenost spojovacích prostředků v češtině a jejich funkční rozrůzněnost upozorňuje např. Daneš (1985). Ukazuje se tedy, že v rámci diskurzních konektorů existuje jejich jádro a periferie. Zatímco na charakteristice jádra se lingvisté často shodnou, při vymezování a popisu periferie nastávají v jednotlivých pojetích velké rozdíly. Proto je zajímavé zaměřit se dále zejména na analýzu méně prozkoumané části konektorů, tedy na jejich „periferii“.

2.4.1. *Definice diskurzního konektoru*

V této části se zabýváme dvěma hlavními otázkami: jak můžeme vymezit diskurzní konektory jako samostatnou množinu výrazů a jak je možné blíže charakterizovat její jádro a periferii. V našem článku využíváme vymezení diskurzního konektoru zavedené v práci M. Rysové a K. Rysové (2014). Na toto pojetí zde navazujeme a dále jej rozpracováváme.

³ Dále se spojovacími prostředky, resp. prostředky strukturace textu obecně, v českém prostředí zabývají například Hausenblas (1962), Kraus (1981), Adamec (1995), Nekula (1996), Tárnyiková (1998) či Štícha (Štícha a kol., 2013).

Signalizovat významový vztah mezi dvěma jednotkami textu mohou jazykové prostředky různého druhu, srov. příklad (13).

- (13) Zpěvačka bohužel onemocněla a přišla o hlas. **Proto** se koncert odkládá na jaro. / **A tak** se koncert odkládá na jaro. / **Z tohoto důvodu** se koncert odkládá na jaro. / **Kvůli tomu** se koncert odkládá na jaro. / **Kvůli její indispozici** se koncert odkládá na jaro.

Otázkou je, zda všechny výrazy zmíněné v příkladu (13) (lexikálně, slovnědruhově i syntakticky značně rozrůzněné) můžeme pokládat za diskurzni konektory. V naší práci definujeme konektory především na základě jejich funkce v textu, neomezujeme se na jejich podobu formální. Významným faktorem je zároveň míra vázanosti konektivního prostředku na konkrétní kontext.

Diskurzni konektor tedy vymezujeme na základě dvou hledisek, a to jako výraz, který signalizuje jeden ze sémanticko-pragmatických typů diskurzních vztahů (jako je příčina, podmínka, exemplifikace, specifikace aj.) mezi dvěma textovými jednotkami a zároveň je vhodný pro vyjádření tohoto typu významového vztahu ve velkém množství kontextů, tj. v tomto ohledu je do značné míry kontextově nezávislý⁴ – viz příklady (14a–d), v nichž lze v několika různých kontextech se vztahem důsledku užít tytéž diskurzni konektory.

- (14a) Zpěvačka bohužel onemocněla a přišla o hlas. **Proto** / **A tak** / **Z tohoto důvodu** / **Kvůli tomu** se koncert odkládá na jaro.
- (14b) Majitelka zámků miluje květiny. **Proto** / **A tak** / **Z tohoto důvodu** / **Kvůli tomu** založila zámeckou botanickou zahradu.
- (14c) Nechci na tu schůzku přijít první. **Proto** / **A tak** / **Z tohoto důvodu** / **Kvůli tomu** vyjdu radši o něco později.
- (14d) Má rodina je pro mě nade vše. **Proto** / **A tak** / **Z tohoto důvodu** / **Kvůli tomu** do práce pospíchat nebudu.

Z jazykových výrazů použitých v příkladu (13) uvedenému vymezení nevyhovuje fráze *kvůli její indispozici*, kterou za diskurzni konektor nepovažujeme – vzhledem k silné závislosti na konkrétním kontextu (k obdobným frázím viz oddíl 2.4.3).

2.4.2. Diskurzni konektory primární a sekundární

Výše uvedené definici konektoru vyhovuje velmi různorodá skupina výrazů. Jak jsme ukázaly, například vztah důsledku lze vyjádřit konektory jako *proto*, *a tak*, *z tohoto důvodu* nebo *kvůli tomu*. Proto diskurzni konektory dále dělíme na primární (například *tak*, *proto*, *tedy*, *tím*) a sekundární (například *díky tomu*; *kvůli tomu*; *to je důvod*, *proč*; *vzhledem k tomu*; *z tohoto důvodu*).

Za hlavní kritérium pro rozlišení primárních a sekundárních konektorů považujeme míru jejich gramatikalizace. Primární konektory jsou výrazy gramatikalizované, zatímco sekundární konektory jsou konstrukce, které ještě plnou gramatikalizací neprošly. Jako

⁴ Více k testování kontextové nezávislosti konektorů, jejich charakteristice a k faktorům ovlivňujícím výběr konektorů při tvorbě textu viz M. Rysová a K. Rysová (2018) a M. Rysová (2018).

vedlejší kritéria (která s gramatikalizací úzce souvisejí) pro odlišení těchto dvou typů konektorů můžeme dále uplatnit především (ne)plnovýznamovost (primární konektory bývají synsémantika, sekundární mívají autosémantické jádro),⁵ větněčlenskou platnost, modifikovatelnost a ohebnost, viz následující oddíl.

2.4.2.1. Primární konektory

Jako primární konektory označujeme spojovací výrazy, které jsou gramatikalizované, zpravidla jednoslovné (*ale, tedy, vždyť, a, protože, totiž* nebo *i*), v některých případech i víceslovné (srov. korelativní páry jako *bud' nebo, jednak jednak* či sousloví typu *i když*). Z lexikálního hlediska jsou primární konektory ustálené, ve větě neplní funkci větných členů, jsou neohebné a zpravidla nedovolují modifikaci.⁶

Funkci primárních konektorů plní především synsémantika, nejčastěji spojky (*Svítilo slunce a já jsem chtěla jít ven.*). Dále se jedná o konektivní užití částic (*Uklid' si ten nepořádek. – Ale to není moje!*), konektivní užití zájmen (*Účast byla velická, což nás moc potěšilo. / Zvolili ji premiérkou. Ve sněmovně tím vznikla napjatá atmosféra.*) a konektivní užití příslovcí, resp. spojovacích výrazů příslovečného původu (*Ubylo mechanické práce a naším hlavním úkolem je kontrola spalinových cest. Zároveň se práce výrazně zmodernizovala.*).

Primární konektory bývají v různých studiích (např. Rouchota, 1996; Fischer, 2006; Urgelles-Coll, 2010) popisovány jako typičtí zástupci diskurzních konektorů, jejich frekvence ve vyjadřování diskurzních vztahů je také mnohem vyšší než u konektorů sekundárních (viz oddíl 4.1). Z tohoto důvodu považujeme primární konektory za centrum množiny diskurzních konektorů a sekundární konektory za její periferii.

2.4.2.2. Sekundární konektory

Sekundární konektory vymezujeme jako (dosud) ne plně gramatikalizované jazykové struktury, které v textu plní spojovací funkci.⁷ Na rozdíl od konektorů primárních jsou méně ustálené – jejich variabilita se týká lexikálního obsazení (srov. *za této/té/dané podmínky*), morfologických tvarů (*příčina je vs. příčinou je*) či slovosledu (*to je důvod, proč vs. je to důvod, proč*). Ve větě plní sekundární konektory funkci větných členů (*díky tomu*) či větných modifikátorů (*jednoduše řečeno*), zpravidla také podléhají flexi (*z tohoto důvodu – z těchto důvodů*) a většinou dovolují modifikaci (*hlavní/další/jedinou příčinou je*).⁸

⁵ U sekundárních konektorů ovšem dochází k postupnému oslabování vlastního lexikálního významu jednotlivých složek a nového (tj. konektivního) významu nabývá celé spojení. Jinými slovy, u příslušných autosémantických složek je oslabován jejich původní význam ve prospěch fungování struktury jako celku.

⁶ Nelze například říci **obecně a*; v určitých případech některé primární konektory modifikované být mohou, např. *hlavně proto*, jejich modifikovatelnost je ovšem omezená – srov. **významně proto*.

⁷ Některé konektory mohou být homonymní s konstrukcemi bez konektivní funkce: *Mladé generace mají velkou šanci se prosadit i v silné světové konkurenci. Na druhé straně jsme málo sebevědomí, ještě stále se někdy bojíme vyjádřit svůj názor; nechceme vyčnívat z řady.* vs. *Hannah a Emmeline nakonec přece jen nebyly u kůlny, ale posedávaly na okraji jezírka na druhé straně.*

⁸ Modifikované konektory – srov. *s jasným cílem*; *z jednoduchých důvodů*; *možná díky tomu* nebo *jediným důvodem bylo* – obsahují hodnotící, modální či intenzifikační výraz, který sám o sobě nesignalizuje příslušný

Jednotlivá slova, která sekundární konektory tvoří, nemívají spojovací funkci ze své podstaty (jako ji mají spojky). Například slovo *důvod* není samo o sobě konektorem, platnost konektoru má až celá konstrukce, například *z tohoto důvodu* či *důvodem je*. I proto užíváme pro daná konektivní spojení označení konektory sekundární.

Sekundární konektory zpravidla obsahují tzv. jádrová slova, která tvoří sémantický základ dané konstrukce (většinou odkazují k určitému sémanticko-pragmatickému vztahu). Jádrová slova bývají často podstatná jména (jako *příčina*, *účel*, *podmínka*) nebo předložky (jako *díky*, *kvůli*, *navzdory*).

2.4.2.3. Hierarchizace sekundárních konektorů (lexikální realizace, gramatické varianty)

Konektivní konstrukce obsahující stejné jádrové slovo mohou být v jazyce velmi rozmanité – srov. například varianty konektorů s jádrovým slovem *důvod* (*z tohoto důvodu*; *to je důvod*, *proč*; *to byl hlavní důvod k tomu, že*; *z uvedených důvodů*), *podmínka* (*pod podmínkou, že*; *za této podmínky*; *podmínkou bylo*; *za daných podmínek*) či *příčina* (*z těchto příčin*; *příčinou je*; *příčina je*). Z uvedených příkladů je zřejmé, že konektivní fráze se stejným jádrovým slovem mohou mít různou syntaktickou strukturu, resp. syntaktické schéma (srov. *z těchto příčin* vs. *příčinou je*).

Ovšem i konstrukce kopírující stejné schéma se v textu mohou objevovat v různých podobách, např. s obměněným lexikálním obsazením (*za této/té/dané/uvedené podmínky*) či v různých gramatických formách. Konektivní konstrukce se tak mohou lišit morfolo- gicky (mohou se například vyskytovat v singularu i plurálu, viz *z tohoto důvodu* vs. *z těchto důvodů*, jádrové slovo lze použít v různých pádech – *důsledek je* vs. *důsledkem je*) či slovosledně (srov. konstrukce typu *to je výsledkem* vs. *je to výsledkem*).

Sekundární konektory tedy vykazují (ve srovnání s gramatikalizovanými konektory primárními) poměrně vysokou míru variability a je zapotřebí se zabývat otázkou, zda konektivní konstrukce se stejným jádrovým slovem můžeme považovat za samostatné konektory, nebo pouze za varianty jednoho a téhož konektoru.

V naší práci chápeme jako samostatné sekundární konektory takové konektivní konstrukce, které náleží k různým syntaktickým schématům. Konstrukce se stejným schématem (obsahující zároveň stejné jádrové slovo) označujeme jako varianty jednoho konektoru. Například tedy spojení *z tohoto důvodu* a *to je důvod, proč* chápeme jako dva různé sekundární konektory, zatímco předložkové fráze *z tohoto důvodu* a *z uvedeného důvodu* jako varianty jednoho a téhož konektoru.⁹

Sekundární konektory dále hierarchizujeme následujícím způsobem. Jednotlivé konstrukce s různým lexikálním obsazením (*za tím účelem* / *za tímto účelem* / *za daným účelem*) nazýváme *lexikální realizace* jednoho syntaktického schématu. Tyto realizace se mohou v textu dále objevovat v různých *gramatických variantách* (*za této podmínky* vs. *za těchto podmínek*; *to je hlavní důvod, proč* vs. *je to hlavní důvod, proč*).

významový vztah, ale pouze ho modifikuje, upřesňuje nebo vyjadřuje jeho intenzitu. Modifikované konektory se objevují hlavně mezi konektory sekundárními (vzhledem k jejich lexikálně-syntaktické povaze).

⁹ Patřící k obecnému schématu *z_Pron_{anaph}/Adj_{anaph}-důvod.2* skládajícímu se z předložky *z*, anaforického zájmena či přídavného jména a slova *důvod* v genitivu.

V hierarchii sekundárních konektorů tedy stojí nejvýše *obecné schéma* (např. *z_Pron_{anaph}/Adj_{anaph}_důvod.2*), dále jeho jednotlivé *lexikální realizace* (např. *z tohoto/toho/daného/uvedeného důvodu*) a posléze jejich *gramatické varianty* (např. *z těchto důvodů, z těch důvodů, z daných důvodů, z uvedených důvodů*). Za základní formu příslušného sekundárního konektoru přitom považujeme jeho nejfrekventovanější realizaci v základním gramatickém tvaru. Pro schéma *z_Pron_{anaph}/Adj_{anaph}_důvod.2* jde o konstrukci *z tohoto důvodu*.¹⁰

2.4.2.4. Diskurzni konektory jako otevřená množina výrazů

V jazyce existuje mnoho „prototypických“ primárních a sekundárních konektorů, které zcela odpovídají výše zmíněným charakteristikám, 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 rysech odlišují, například primární konektor *proto* může být modifikovaný (*hlavně/právě/jedině proto*). Z tohoto hlediska necháme primární a sekundární konektory jako striktně uzavřená a oddělená skupiny výrazů, ale jako škálu jazykových prostředků v různém stupni gramatikalizace. To znamená, že některé sekundární konektory se mohou stát primárními – díky přirozeným jazykovým změnám a případné budoucí gramatikalizaci.

Proces gramatikalizace můžeme pozorovat i na současných primárních konektorech (resp. na jejich původu) – některé totiž vznikaly ze struktur podobných současným sekundárním konektorům (srov. primární konektory *proto, přesto, přitom, potom, zato* původně vzniklé z předložkových frází a dosud negramatikalizované sekundární konektory typu *kvůli tomu, díky tomu*). Ustalování konektivní funkce u předložkových frází dokládají také cizojazyčné výrazy, které vznikly podobným způsobem jako české *proto* – srov. anglický konektor *therefore* (vzniklý spojením anaforického slova *there* a slova *fore*, staroanglické a středoanglické formy předložky *for*) či *daarfoor* v holandštině, *dafür* v němčině nebo *derfor* v dánštině.

Ke gramatikalizaci dochází i u dalších konektivních struktur, srov. současný primární konektor *because* v angličtině vzniklý z fráze *bi cause (by cause)* či konektor *for* vzniklý krácením – pravděpodobně ze staroanglických frází typu *for þon þy* (doslovně *for the (reason) that*), viz Harper (2001–2021). V tomto ohledu by tedy obdobný vývojový proces mohl nastat i u dalších, dosud ne plně gramatikalizovaných konektivních struktur.

2.4.3. Volné konektivní fráze

Významové vztahy nejsou v textu vyjadřovány pouze diskurzními konektory (ať už primárními, či sekundárními), ale mohou být signalizovány i jinými jazykovými prostředky, srov. příklad (15).

- (15) Vždy jsem se cítila méněcenná. **Kvůli tomuto vnitřnímu pocitu** jsem nikdy nedosáhla na vedoucí pozici.

¹⁰ Měřeno na datech PDiT 2.0.

V uvedeném příkladu signalizuje spojení *kvůli tomuto vnitřnímu pocitu* vztah důsledku. Obdobné konstrukce nazýváme termínem volná konektivní fráze (srov. M. Rysová – K. Rysová, 2014). Od diskurzních konektorů se liší tím, že jsou silně závislé na kontextu. Pro vyjádření příslušného diskurzního vztahu mohou být tedy použity pouze v určitých případech – srov. příklady (16a) a (16b). Volné konektivní fráze obsahují specifickou referenční složku (např. *tomuto vnitřnímu pocitu*, *tomuto počasí*, *tomuto zvýšení*), která je propojuje s širším kontextem – jedná se o způsob navazování pomocí integrovaného autosémantika, které shrnuje, explikuje nebo opakuje obsah či smysl předchozí textové jednotky.

(16a) Vždy jsem se cítila méněcenná. **Proto / Kvůli tomu / Kvůli tomuto vnitřnímu pocitu** jsem nikdy nedosáhla na vedoucí pozici.

(16b) Velvyslanec byl obviněn z podvodu a zneužití důvěry. **Proto / Kvůli tomu / *Kvůli tomuto vnitřnímu pocitu** musel složit svou funkci.

2.4.4. Diskurzní konektory a reference

Odkazování do kontextu má zásadní význam pro konektivní prostředky obecně. Podstatou diskurzních konektorů je, že propojují dva diskurzní argumenty a vyjadřují mezi nimi určitý sémanticko-pragmatický vztah. Jako takové mají schopnost odkazovat do okolního kontextu, resp. přítomnost konektoru v textu implikuje zároveň přítomnost dvou argumentů, které tento konektor spojuje (srov. např. Halliday – Hasan, 1976). Z širšího hlediska tak můžeme diskurzní konektory chápat jako výrazy implicitně referenční – např. konektory jako *a*, *nebo*, *ale* explicitní referenční prvek neobsahují, ale přesto zřetelně signalizují přítomnost dvou diskurzních argumentů.

V rámci široké skupiny diskurzních konektorů zároveň existují takové výrazy, které referenční komponent obsahují explicitně. Jedná se například o sekundární konektory jako *díky tomu*, *za této podmínky*, *z těchto důvodů* nebo *výsledkem toho je*. I mezi nimi ovšem můžeme pozorovat určité rozdíly – některé konektory (např. *díky tomu*) obsahují referenční (v tomto případě anaforický) prvek obligatorně, zatímco u jiných být vyjádřen může a nemusí (srov. *výsledkem toho je* vs. *výsledkem je*).

Na základě (ne)přítomnosti referenčního prvku tak můžeme diskurzní konektory rozčlenit na konektory:

- 1) bez inherentní reference (např. *a*, *ale*, *i*, *nebo*, *však*) a
- 2) s inherentní referencí, která může být:
 - a. fakultativní (srov. *příkladem je* vs. *příkladem toho je*) či
 - b. obligatorní – tyto konektory mohou být:
 - i. jednoslovné (*proto*, *zato*, *přítom*, *přesto*) a
 - ii. víceslovné (*díky tomu*, *za této podmínky*).

Inherentně referenční konektory můžeme dále dělit z hlediska jejich dosahu, resp. směru, kterým v textu působí, a to na konektory anaforické (*proto*, *potom*, *z tohoto důvodu*, *díky tomu*) a kataforické (*protože*; *kvůli tomu*, *aby*; *kromě toho*, *že*). V příkladu (17)

odkazuje sekundární konektor *z toho důvodu* anaforicky do předchozího kontextu. Příklad (18) ilustruje užití kataforického konektoru *z toho důvodu, že*, který naopak odkazuje do kontextu následujícího.

- (17) Chmel je popínavá rostlina a je třeba mu poskytnout podporu pro jeho růst. Z toho důvodu ^{anaf} jsou chmelnice vybaveny konstrukcemi zhotovenými obvykle z dřevěných sloupů o přibližné výšce 6 metrů.
- (18) Herec byl pro roli hlavní postavy slavného hokejisty osloven **z toho důvodu, že** ^{kataf} je legendárním sportovci velmi podobný.

2.4.5. Souvýskyt diskurzních konektorů

Dalším významným rysem diskurzních konektorů je, že se v textu skládají a kombinují. V jedné větě se tak často objevují dva konektory vedle sebe. Souvýskyt konektorů přitom může být různého charakteru. V některých případech vyjadřují konektory vyskytující se vedle sebe jeden diskurzní vztah a dohromady tak fungují jako jeden složený konektor. Kombinovat se tímto způsobem mohou konektory primární i primární a sekundární, srov. příklady (19–21) se složenými konektory *i když*, *a proto*, *a díky tomu* vyjadřujícími vztah přípustky (příklad 19) nebo důsledku (příklady 20 a 21).

Ve většině případů fungují složky složeného konektoru jako konektory i samostatně (*i když*, *a proto* apod.), ve výjimečných případech se konektivní výrazy samostatně nepoužívají, tj. fungují pouze jako část složeného konektoru (např. *bud' nebo* v rámci složeného konektoru *bud' nebo*).

- (19) Většina Němců chce zelenou energii, **i když** je dražší a méně spolehlivá.
- (20) V průběhu roku na něj budou kladeny značné finanční požadavky, **a proto** musí moudře hospodařit se svými výdaji.
- (21) Téměř každý vytěžený diamant má kvalitu drahokamu, **a díky tomu** se tak nákladná těžba z moře firmě vyplácí.

Mezi jednotlivými složenými konektory přitom můžeme sledovat rozdíly, a to především v pevnosti jejich spojení. Konektor *i když* je například specifický tím, že jeho jednotlivé složky nevyjadřují samy o sobě stejný typ vztahu, jako když jsou užitý vedle sebe (tj. samotné *i* ani samotné *když* nevyjadřuje vztah přípustky). Zároveň nemohou být složky tohoto konektoru od sebe odděleny jiným textem (**Většina Němců chce zelenou energii, i je když dražší a méně spolehlivá*), zatímco složky jiných, volnějších spojení ano (*V průběhu roku na něj budou kladeny značné finanční požadavky, a musí proto moudře hospodařit se svými výdaji*).

Častý souvýskyt některých spojovacích výrazů (zejména konektorů primárních se sémanticky obecnějším konektorem *a*) může vést až ke vzniku nového, plně gramatikalizovaného konektoru: *avšak, anebo*.

Z hlediska pozice v textu se mohou výrazy tvořící složený konektor vyskytovat ve stejném argumentu (příklady 19–21 výše), nebo každá složka v jiném argumentu (v tom případě se jedná o korelativní páry typu *bud' nebo*, srov. příklad 22).

- (22) **Bud'** kutil něco v dílně, **nebo** se díval na zprávy.

V textu se můžeme setkat také s případy konektorů zdvojených, srov. příklady (23) a (24). V nich se objevují dva sémanticky silně naplněné konektory, které jsou si významově velmi podobné. Pro zachycení daného typu vztahu by v textu stačilo použít jen jeden z nich. Přesto je užívání zdvojených konektorů v přirozeném jazyce poměrně běžný jev.

- (23) Dosavadní senátorku za Severočechy nelze zvolit, protože se lídři této podivné party rozhá-dali a vyškrtli ji. Je **však ale** možné zvolit hokejistu Jiřího Šlégra.
- (24) Firma nedokáže nastavit systém hodnocení a rozdělení působnosti obchodníků tak, aby na-vzájem nesoupeřili, **ale naopak** si pomáhali a učili se od sebe.

Tímto způsobem se mohou kombinovat dva souřadící konektory (příklady 23 a 24) i konektor souřadící a podřadící (příklad 25). V případech, kdy se ve vyjádření vztahu mezi týmiž úseky textu kombinuje spojka souřadící a podřadící, se stírá rozdíl mezi parataxi a hypotaxi.

- (25) Na sezonu 2016 rozhodně nikdy nezapomenu, **protože totiž** ono opravdu všechno zlé je pro něco dobré.

V některých případech dochází i k tomu, že se vedle sebe vyskytnou dva zcela samo-statné, sémanticky silně naplněné konektory, z nichž každý signalizuje jiný diskurzni vztah mezi týmiž diskurzními argumenty (příklad 26). Častý je samozřejmě také souvý-skyt dvou konektorů vyjadřujících vztah mezi jinými argumenty (srov. příklad 27, ve kte-rém konektor *ale* vyjadřuje vztah opozice mezi dvěma hlavními větami, zatímco konek-tor *protože* vztah příčiny mezi větou vedlejší a druhou větou hlavní).

- (26) Podle soudce byly důkazy dostatečně přesvědčivé, některé **proto dokonce** zamítl pro nadby-tečnost.
- (27) V posledních měsících se museli v mnoha směrech omezit, **ale protože** s penězi vždycky na-kládali opatrně, pořád jim zbývala malá rezerva.

V textu se tedy můžeme setkat s různými případy skládání a kombinace diskurzniích konektorů a vždy je zapotřebí zkoumat, k jakým úsekům textu se jednotlivé konektory vztahují, zda fungují dohromady jako jedna jednotka, či zda se jedná o dva samostatné konektory vyjadřující různé diskurzni vztahy.

3. Pražský diskurzni korpus

V další části práce se věnujeme rozboru a popisu chování diskurzniích konektorů v kor-pusových textech. Jako hlavní zdroj jazykových dat pro náš výzkum slouží Pražský dis-kurzni korpus 2.0 (PDiT: Prague Discourse Treebank, M. Rysová et al., 2016), který je vystavěný na textech Pražského závislostního korpusu (PDT, Hajič et al., 2006). Jedná se o korpus novinových textů skládající se z 3 165 dokumentů (49 431 vět, 833 195 pozic). PDiT 2.0 obsahuje anotaci sémanticko-pragmatických vztahů a jazykových prostředků, kterými jsou tyto vztahy vyjádřeny.

První verze diskurzniího korpusu (PDiT 1.0, Poláková et al., 2012a) obsahovala znač-kování sémanticko-pragmatických vztahů explicitně vyjádřených diskurzniími konek-tory. Diskurzni konektory byly v této verzi vymezeny jako spojky, vybraná příslovce

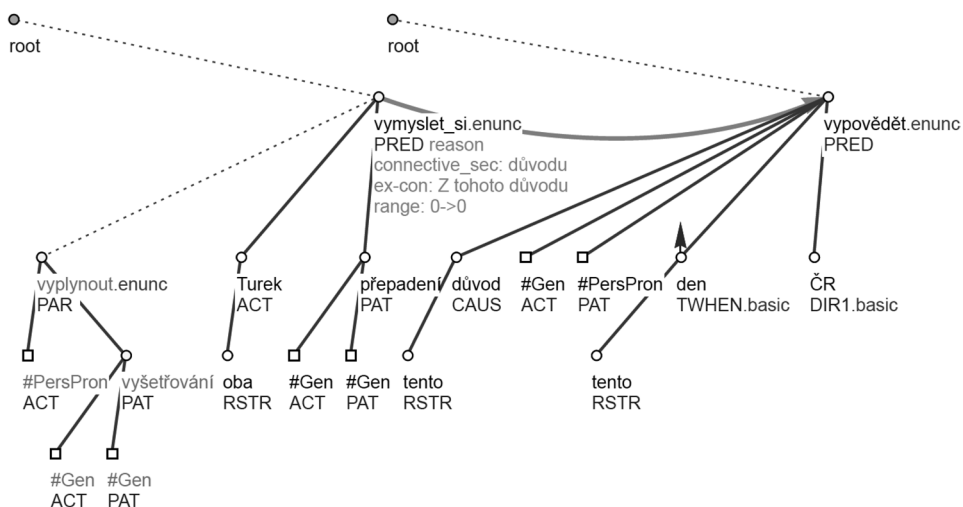
(např. *potom*) a vybrané částice (např. *jen*), některá interpunkční znaménka (dvojtečka, pomlčka), některá užití zájmen (např. *kromě toho*) a některé typy idiomatických víceslovných spojení (blíže k tomu Poláková et al., 2012b; Zikánová et al., 2015a).

Druhá verze PDiT 2.0 reflektuje dělení diskurzních konektorů na primární (např. *a*, *ale*, *totiž*, *když*) a sekundární (např. *z tohoto důvodu*, *za této podmínky*, *za tímto účelem*). PDiT 2.0 obsahuje jednak revizi předchozího značkování textových jevů (například některá idiomatická víceslovná spojení byla přeznačena jako sekundární konektory), jednak značkování nové sady diskurzních konektorů – sekundární konektory byly nově značeny na celých datech korpusu.¹¹

PDiT obsahuje značkování vnitrovětných (příklad 27 výše) i mezivětných (příklad 28) sémanticko-pragmatických vztahů.¹² Ukázka korpusového značení textových vztahů a konektorů, které tyto vztahy vyjadřují, je zachycena na obrázku 1 (příklad 28).

(28) Jak vyplynulo z vyšetřování, oba Turci si přepadení vymysleli.

Z tohoto důvodu byli v těchto dnech z ČR vypovězeni.



Obrázek 1: Ukázka z Pražského diskurzního korpusu 2.0. Diskurzni vztah důsledku mezi dvěma propozicemi vyjádřený sekundárním konektorem *z tohoto důvodu*.

Obrázek 1 zachycuje mezivětný důsledkový vztah vyjádřený sekundárním konektorem *z tohoto důvodu*. Textový vztah je vyznačen silnou šipkou, která vede mezi predikáty obou závislostních stromů. Z obrázku je patrné, že sekundární konektory jsou v korpusu značeny na dvou úrovních: nejprve je zachyceno jádrové slovo sekundárního konektoru

¹¹ Diskurzni anotace z PDiT se posléze stala součástí Pražského závislostního korpusu 3.5 (Hajič et al., 2018) a souborného korpusového vydání PDT-C 1.0 (Hajič et al., 2020).

¹² Tedy jednak vztahů uvnitř větného celku a jednak vztahů nad úrovní větného celku.

(zde *důvod*), poté je uveden konkrétní diskurzí konektor (*z tohoto důvodu*). Značkování jádrových slov sekundárních konektorů usnadňuje mj. vyhledávání konektorů s podobným významem, avšak jinou strukturou (srov. *z tohoto důvodu; to je důvod, proč*).

Diskurzí anotace také zachycuje, jak rozsáhlé jsou textové úseky, mezi nimiž vztah existuje. Ve velké většině případů jde o dvě sousední věty nebo o dva sousední větné celky (jako je tomu na obrázku 1) – tato informace je zachycena symbolem $0 \rightarrow 0$.

Na základě vyhodnocení dat PDiT 2.0 se pokusíme přispět k bližší charakteristice diskurzí konektorů v psané češtině. Doplňkově používáme také data Českého národního korpusu (SYN2015, Křen et al., 2015), který sice přímo neobsahuje značkování textových jevů, ale najdeme v něm značkování slovních druhů. Svým rozsahem převyšuje data PDiT, a některé jevy na něm lze proto sledovat podrobněji.

4. Diskurzí konektory v číslech

V následujících oddílech se postupně zabýváme frekvencí diskurzí konektorů v textech (4.1), nejčastěji se vyskytujícími složenými konektory (4.2) a užíváním konektorů s fakultativní referencí typu *příkladem (toho) je* (4.3).

4.1. Primární a sekundární konektory v Pražském diskurzí korpusu

Při analýze jazykových dat korpusu PDiT jsme se nejprve zaměřily na četnost výskytu diskurzí konektorů (primárních a sekundárních) a volných konektivních frází. Z výsledků je zřejmé, že autoři dávají přednost gramatikalizovaným primárním konektorům, které vyjadřují až 94 % všech sémanticko-pragmatických vztahů (20 255 výskytů). Přibližně 5 % (1 161 výskytů) diskurzí vztahů je vyjádřeno konektory sekundárními a jen asi 1 % (151 výskytů) kontextově závislými frázemi typu *kvůli tomuto malému zaváhání*, srov. příklady (29), (30) a (31).

- (29) Dužina z dýně obsahuje velké množství vody, která se při tepelném zpracování vylučuje. **Proto** primární konektor byste měli další tekutiny přilévat jen s mírou.
- (30) Při delším používání se přehřívala natolik, že se diskety teplem deformovaly, což se opravdu nedá považovat za bezpečné zacházení s daty. **Kvůli tomu** sekundární konektor dostal tento výrobek přezdívku „pec na pizzu“.
- (31) Podle plánu měl Sin utíkat jako první, dokud se nedostanou do bezpečné vzdálenosti za plot, když vtom nečekaně uklouzl a na zledovatělé pěšině, vyšlapané strážnými, se poroučel k zemi. **Kvůli tomuto malému zaváhání** volná konektivní fráze se stalo, že Pak dospěl k plotu jako první.

Ukazuje se, že autor textu dává přednost konektorům, které jsou využitelné ve velkém množství kontextů, před kontextově závislými frázemi. Důvodem je pravděpodobně to, že vhodný konektor stačí vybrat z množiny ustálených výrazů, ale kontextově závislé fráze musí autor textu nově formulovat.

V rámci konektorů pak autoři jednoznačně upřednostňují konektory primární před sekundárními. Důvod zřejmě spočívá v tom, že primární konektory jsou krátké, neohebné, již lexikalizované výrazy, zatímco sekundární konektory bývají víceslovné a vykazují velkou míru variability (lexikální, morfologické i syntaktické).

Z toho vyplývá, že se při užívání konektorů uplatňuje princip tzv. jazykové ekonomie. Autor se snaží s využitím minimálního úsilí dosáhnout maximálního výsledku, a najít tak co nejjednodušší cestu, která vede k vytvoření funkčního, koherentního textu, srov. Zipfův zákon minimálního úsilí (Zipf, 1949).

V dalším zkoumání jazykového materiálu PDiT jsme se zaměřily na frekvenci výskytu jednotlivých konektorů. Ukazuje se, že pro vyjadřování významových vztahů si autor opakovaně vybírá ze stále stejné množiny výrazů. Jednoznačně nejfrekventovanějším konektorem je spojka *a*, jejíž výskyty pokrývají 27 % výskytů všech konektorů dohromady (PDiT obsahuje celkem 21 416 výskytů konektorů, počet výskytů konektoru *a* je 5 766). První tři nejčastěji užívané konektory (*a*, *však*, *ale*) tvoří 40 % výskytů všech konektorů v PDiT (8 554 výskytů), prvních pět (*a*, *však*, *ale*, *když*, *protože*) pokrývá 45 % (9 653 výskytů) a prvních deset (*a*, *však*, *ale*, *když*, *protože*, *totiž*, *pokud*, *proto*, *tedy*, *aby*) dohromady tvoří 54 % (11 508 výskytů). I v tomto ohledu se tedy potvrzuje, že při vytváření textu volí autor co nejjednodušší cestu. V PDiT je zachyceno téměř 1 000 různých forem konektorů – přibližně 570 forem primárních a téměř 400 forem sekundárních konektorů (jako samostatné formy jsou zde chápány například různé typy složených a modifikovaných konektorů či lexikálních realizací). Pro vyjadřování sémanticko-pragmatických vztahů v textu však v polovině případů autor vybíral pouze z deseti z nich. Těchto deset konektorů tedy stačilo pokrýt více než polovinu komunikačních situací, ve kterých se objevil explicitní konektivní prostředek.

V rámci sledování frekvence výskytů konektorů v textech jsme se dále zaměřily na počet jejich užití uvnitř jednoho větného celku a nad jeho úrovní. V tomto ohledu můžeme vysledovat mezi jednotlivými typy konektorů značné rozdíly. Z výsledků našeho výzkumu vyplynulo, že primární konektory obecně preferují užití uvnitř jednoho větného celku (v 70 % výskytů v textech PDiT z celkových 20 255), zatímco u sekundárních konektorů převažuje užití mezi dvěma větnými celky (v 63 % výskytů z 1 161). Typické postavení primárních a sekundárních konektorů ilustrují příklady (32) a (33).

(32) Budu ji muset obsloužit, **protože** moje kolegyně odběhla do kuchyně.

(33) Stejně jako v medicíně, i při léčbě nežádoucího chování je třeba nejprve stanovit diagnózu. **Za tímto účelem** je třeba pozorně sledovat jak chování dané kočky, tak celé její okolí.

Pokud ale zkoumáme syntaktické chování konektorů v textu, nemůžeme se omezovat pouze na analýzu konektorů primárních a sekundárních jako ucelených skupin. Syntaktické chování je silně ovlivněno slovnědruhovou příslušností jednotlivých konektorů, a je tedy zapotřebí podívat se blíže na to, jaké konkrétní výrazy dané skupiny tvoří.¹³ Jak jsme ukázaly výše, u sekundárních konektorů obecně (v absolutních číslech) převládá mezivětné užití, pokud však struktura sekundárního konektoru obsahuje podřadící spojku, převažuje u ní užití vnitrovětné, viz příklad (34).

¹³ Podobnou problematikou se v nedávné době zabývala Jínová (2012) a Zikánová et al. (2015a). Zmíněné studie poukazují na to, že konektory ve formě podřadících spojek se používají většinou vnitrovětně (uvnitř souvětí), zatímco příslovce mezivětně (nad úrovní větného celku). Souřadící spojky jsou podle nich v tomto směru nevyhraněné.

- (34) Pohádali se **kvůli tomu**, že žena stále váhala, zda manžela opustit, nebo se k němu vrátit.

Převažující výskyt podřadicích spojek (ve funkci primárních konektorů, viz příklad 32, či jako součást sekundárních konektorů, viz příklad 34) uvnitř větného celku není překvapivý. Různorodější chování v tomto směru můžeme pozorovat u konektivních výrazů v souřadných konstrukcích. Mnohé z nich mají schopnost vyskytovat se vnitrovětně (příklad 35) i mezivětně (příklad 36). Zajímavé je, že ale velké množství z nich zároveň zřetelně preferuje jednu z těchto variant. Užití ve druhé, méně frekventované variantě nemusí být negramatické, ale v komunikaci není příliš časté. Například konektor *ale* se v 66 % vyskytuje uvnitř souvětí (srov. 1 267 výskytů v PDiT) a jen ve 34 % (420 výskytů v PDiT) je užíván mezi větnými celky. Typický výskyt konektoru *ale* tedy ilustruje příklad (35), zatímco příklad (36) dokládá gramaticky sice možné, ale menšinové užití tohoto spojovacího prostředku.

- (35) Mluví našťavaně, **ale** hlas má monotónní a nízký.

- (36) Odvykání často probíhá s pomocí náhražkových drog. **Ale** existují i možnosti detoxikace bez těchto náhražek.

Při bližším pohledu na četnost užití konektorů v souřadných konstrukcích v PDiT se tak ukazují obecnější tendence. Vnitrovětné postavení zřetelně preferují zejména konektory neadverbiálního původu – srov. např. *a* (vyskytující se uvnitř souvětí v 94 % případů, tj. v 5 428 výskytech v PDiT), *nebo* (87 %, 167 výskytů), *či* (100 %, 86 výskytů). Stejnou tendenci mají také výrazy, které s konektory neadverbiálního původu tvoří dohromady složený konektor – například *a tak* (uvnitř souvětí v 68 %, tj. v 97 výskytech v PDiT), *a to* (82 %, 97 výskytů), *a proto* (100 %, 86 výskytů), *a pak* (89 %, 50 výskytů). K užití uvnitř větného celku inklinují také korelativní páry – typu *sice_ale* (uvnitř souvětí v 98 %, tj. v 162 výskytech) – a výrazy s negací, například *neboť* (99 %, 220 výskytů), *nýbrž* (ve 100 %, tj. ve 12 výskytech).

Naopak převážně mezivětně se užívají zejména spojovací výrazy adverbiálního původu, srov. prvních pět nejfrekventovanějších konektorů tohoto typu: *však* (užité mezivětně v 83 % případů, tj. v 1 255 výskytech v PDiT), *totiž* (95 %, 436 výskytů), *proto* (90 %, 339 výskytů), *tedy* (90 %, 274 výskytů), *pak* (74 %, 217 výskytů). Primární konektory adverbiálního původu jsou v tomto ohledu velmi blízké sekundárním konektorům.¹⁴

Závěrem můžeme shrnout, že konektory ve formě podřadicích spojek nebo podřadicí spojující preferují užití vnitrovětné, nemůžeme ovšem říci, že konektory v souřadných konstrukcích obecně preferují užití mezi větnými celky. Rozhodující vliv má v těchto případech především (ne)adverbiální původ daného konektoru. Při popisu rozdílů v syntaktickém chování konektorů tak kromě tradiční opozice souřadnosti a podřadnosti hraje důležitou roli také opozice vnitrovětnosti a mezivětnosti (srov. k tomu Danlos et al., 2018).

¹⁴ Sekundární konektory ve větě často zastávají roli příslovočného určení, a zřejmě proto preferují mezivětné užití. Větněčlenská platnost většiny sekundárních konektorů je však již značně oslabená. Tyto struktury ji pomalu ztrácejí a posiluje se u nich funkce konektivní (podobný průběh můžeme pozorovat i u konektorů primárních typu *proto*, *přesto*).

4.2. Složené konektory v Pražském diskurzním korpusu a v Českém národním korpusu

V rámci rozboru jazykového materiálu PDiT jsme se dále zaměřily na nejčastější výskyty složených konektorů. Jak jsme popsaly v oddíle 2.4.5, jedná se o souvýskyty výrazů, které v daném kontextu fungují do jisté míry jako jedna jednotka – vyjadřují jeden textový vztah a váží se ke stejným argumentům. Složené konektory jsou v anotaci PDiT explicitně značeny. Zachyceny jsou přitom i případy složených konektorů, jejichž složky nestojí v textu nutně vedle sebe (srov. *a mám proto důvod k oslavám* vs. *a proto mám důvod k oslavám*), a výskyty korelativních párů typu *bud' _nebo*, v nichž každá část stojí v jiném argumentu téhož významového vztahu.¹⁵

Nejčastěji se v textech PDiT vyskytují tyto složené konektory: *i když* (178 výskytů), *sice _ale* (165), *a tak* (141), *a to* (118), *a proto* (86), *a pak* (56), *a také* (51), *sice _však* (41), *a i* (38), *a tudíž* (20), *a tedy* (20), *a dokonce* (19).

Nejfrekventovanějším složeným konektorem v PDiT je výraz *i když*. Mezi složky tohoto konektoru nelze vložit jiný text (na rozdíl od spojení v tomto ohledu volnějších *a tak*, *a proto*, *a pak*, *a také*, *a i* aj.), navíc konektory *i a když* vyjadřují samostatně jiné typy sémantických vztahů než dohromady. V tomto ohledu *i když* vykazuje vysokou míru gramatikalizace.

Mezi další frekventované složené konektory patří korelativní páry *sice _ale* a *sice _však*. Ve zbylých případech se jedná o spojení konektorů se sémanticky vyprázdněnějším konektorem *a* – zpravidla je u těchto složených konektorů oslabena sémantická funkce *a* ve prospěch dominantnější složky typu *proto*, *tudíž* nebo *dokonce*. Díky spojení s konektorem *a* mohou být tyto složky lépe užity ve vnitrovětných kontextech. Konektory adverbialního původu typu *proto*, *tudíž*, *dokonce* se totiž preferovaně vyskytují mezivětně, ve spojení s *a* u nich naopak převažuje užití uvnitř jednoho větného celku, srov. oddíl 4.1. Obdobně se s konektorem *a* mohou pojít i konektory sekundární, výskyt takových složených konektorů je ovšem v PDiT velmi nízký – srov. *a kromě toho* (4 výskyty), *a z toho(to) důvodu* (3) nebo *a kvůli tomu* (1).

Souvýskyt konektorů jsme dále zkoumaly na obsáhlejších datech Českého národního korpusu. Ten nezachycuje textové vztahy ani konektory, obsahuje ale pro nás relevantní značkování slovních druhů. Data PDiT ukazují, že se v rámci diskurzních konektorů kombinují především spojky, příslovce a částice, okrajově zájmena. V korpusu SYN2015 jsme proto hledaly souvýskyt těchto slovních druhů (resp. dvojice spojka – spojka, spojka – příslovce, spojka – částice a spojka – zájmeno). Vyhledané výskyty jsme dále prořídily – zahrnovaly jsme pouze souvýskyty dvou konektorů, nikoli například konektoru a obsahové spojky (typu *a že* nebo *že když*). Z analýzy jsme dále vyloučily konektory jako *a když*, *ale když*, *a pokud* (jednalo se především o spojení spojky souřadící a podřadící, z nichž každá spojuje jiný úseky textu a vyjadřuje jiný sémantický vztah).

¹⁵ Jiné typy souvýskytů (tj. konektory vyskytující se *sice* v textu vedle sebe, ale vyjadřující jiný typ vztahu mnohdy mezi jinými textovými argumenty) nejsou v PDiT explicitně značeny – srov. příklady typu *ale proto*, *jenže když*. Každý z těchto konektorů je zachycen samostatně, v rámci jiného textového vztahu.

Tímto způsobem jsme získaly nejčastější výskyty složených konektorů v korpusu SYN2015:¹⁶ *i když* (35 864), *a pak* (35 833), *a tak* (32 865), *a také* (15 831), *a ještě* (8 460), *a proto* (8 436), *a zároveň* (7 886), *a potom* (6 673), *a navíc* (6 089), *a přitom* (6 015), *a taky* (5 866), *a nakonec* (5 820), *a i* (5 779), *a znovu* (5 227), *i kdyby* (4 931), *i tak* (4 130), *a tedy* (2 767), *a tudíž* (1 742).

Výsledky ukazují, že nejfrekventovanější složené konektory jsou v textech PDiT i ČNK obdobné. Nejčastěji používaný složený konektor je výraz *i když* a i v dalších případech se jedná o spojení obsahující často sémanticky méně naplněné konektory *a* a *i*.

4.3. Konektory s fakultativní referencí v Pražském diskurzním korpusu a v Českém národním korpusu

Na jazykovém materiálu PDiT jsme dále zkoumaly výskyty referenčních konektorů v psané češtině. Jak jsme popsaly v oddíle 2.4.4, vztah k referenci, resp. k přímému či nepřímému odkazování do kontextu, je významnou oblastí výzkumu diskurzních konektorů obecně.

Referenční prvek přitom může být přímo součástí konektoru (srov. *proto*, *přitom*, *kvůli tomu*) nebo může být jeho užití fakultativní – např. *příkladem (toho) je*. Právě obdobným případům, tj. fakultativnímu užití reference v konektivních konstrukcích, se věnujeme v tomto oddíle. Zabýváme se především otázkou, jakou variantu daného konektoru užívají autoři častěji – zda konstrukci s referencí explicitně vyjádřenou (*příkladem toho je, důsledkem toho je*), či nevyjádřenou (*příkladem je, důsledkem je*), srov. příklady (37) a (38).

- (37) Přírodní léčiva a potravinové doplňky jsou obvykle vyráběny z přírodních zdrojů a dosud nejsou tak přísně kontrolovány jako běžné léky. **Důsledkem toho je**, že existují velké rozdíly v jejich kvalitě, účinnosti a bezpečnosti.
- (38) Na jedné polokouli je stále noc, na druhé den. Mezi oběma jsou velké teplotní rozdíly, které se planeta snaží vyrovnat. **Důsledkem je**, že v atmosférách takových planet pořádně fouká.

Korpus PDiT 2.0, na kterém jsme výzkum konektorů prováděly primárně, značkování víceslovných referenčních konstrukcí obsahuje přímo. Z tohoto korpusu jsme pro naši analýzu vybraly takové konektory s fakultativní referencí, které mají z formálního hlediska podobnou strukturu: *důvodem (toho) je*, *výsledkem (toho) je*, *výjimkou (z toho) je*, *příkladem (toho) je*, *podmínkou (toho) je*, *příčinou (toho) je*, *účelem (toho) je*, *důsledkem (toho) je*, *následkem (toho) je*.

PDiT 2.0 obsahuje celkem 82 výskytů uvedených konektivních konstrukcí, ve všech případech šlo o užití bez referenčního prvku. Počet výskytů daných konektorů je ovšem v PDiT poměrně nízký. Uvedené konstrukce (s referenčním prvkem i bez něj) jsme tedy

¹⁶ Vzhledem k tomu, že korpus SYN2015 neobsahuje přímo anotaci diskurzních konektorů, nejsou uvedená čísla zcela přesná; bylo například možné vyhledat pouze složené konektory, které se v textu vyskytují bezprostředně vedle sebe – uvedená čísla tedy nezachycují případy typu *a zítřka tedy na výlet nepojedu*, které zahrnuje korpus PDiT.

dále vyhledaly v rozsáhlejší korpuse bez diskurzí anotáce – v textech Českého národního korpusu. Výsledky výzkumu provedeného na textech SYN2015 jsou představeny v tabulce 2.

Tabulka 2: Výskyt vybraných sekundárních konektorů s fakultativním referenčním prvkem v datech korpusu SYN2015

SYN2015	Bez referenčního elementu			S referenčním elementem		
	(typ příkladem je)			(typ příkladem toho je)		
	Výskyt	i.p.m. ¹⁷	ARF ¹⁸	Výskyt	i.p.m.	ARF
výsledkem (toho) je	2 267	18,77	1 070,72	28	0,23	15,40
důvodem (toho) je	1 519	12,58	701,41	1	0,01	1,00
příkladem (toho) je	1 236	10,24	476,51	13	0,11	7,15
výjimkou (z toho) je	783	6,48	396,83	0	0,00	-
příčinou (toho) je	430	3,56	216,82	10	0,08	5,03
podmínkou (toho) je	392	3,25	185,80	1	0,01	1,00
účelem (toho) je	371	3,07	186,77	0	0,00	-
důsledkem (toho) je	362	3,00	156,67	46	0,38	24,15
následkem (toho) je	58	0,48	29,26	25	0,21	15,82
CELKEM	7 418	61,43	3 612,65	124	1,03	70,40

Z výsledků je patrné, že v textech korpusu SYN2015 výrazně (v 98 %) převažují výskyt bez referenčního prvku. Pokud tedy konektivní konstrukce dovoluje dvojí způsob užití (např. *důsledkem je* i *důsledkem toho je*), preferovaná je varianta s referencí implicitní.

I v tomto případě se uplatňuje zákon jazykové ekonomie – pokud má autor na výběr, upřednostňuje variantu kratší, a tím i ekonomičtější. Vynechání fakultativního referenčního elementu přitom nebrání porozumění textu. Čtenář dokáže i bez referenční části jednoznačně identifikovat, ke kterému argumentu se odkazuje, a pochopit význam diskurzího vztahu.

Zajímavé je, že vynechání referenčního prvku se projevuje i u některých dalších konektorů, u kterých bychom podle jejich struktury očekávali referenci obligatorní. Jedná se například o spojení *místo aby* (srov. variantu s kataforickou referencí *místo toho, aby*). U obdobných předložkových konstrukcí typu *kromě toho, že/aby; kvůli tomu, že/aby* nebo *díky tomu, že* je užití referenčního prvku obligatorní. U spojení *místo toho, aby* nicméně dochází k jeho vypouštění, srov. příklady (39) a (40).

- (39) Přešel jsem parkoviště, kelímky s kávou se přitom nakláněly na různé strany a v půli cesty domů se jeden převrátil z krabice ven. **Místo abych** ho nechal spadnout na zem, snažil jsem se ho chytit, a přitom se mi na chodník vyklopily i všechny ostatní.
- (40) Spousta restaurací řeší akcemi na slevových serverech nedostatek klientely. **Místo toho, aby** se majitelé zaměřili na podnik a snažili se přijít na to, proč je navštěvuje málo hostů, nabízejí hromadné slevy.

¹⁷ Zkratka i.p.m. (*instances per million*) vyjadřuje průměrný počet výskytů v hypotetickém korpusu o délce 1 milion slov.

¹⁸ Zkratka ARF (*average reduced frequency*) vyjadřuje průměrnou redukovanou frekvenci.

Kratší varianta *místo aby* je v textech dokonce mnohem frekventovanější – převažuje v textech PDiT (celkový počet výskytů této fráze je zde ale nízký, srov. 11 užití *místo aby* vs. 2 užití *místo toho, aby*) i v korpusu SYN2015 (a to v 78 % výskytů, srov. 1 807 užití *místo aby* vs. 524 *místo toho, aby*).¹⁹ Vyjadřování reference u diskurzních konektorů se může lišit v různých jazycích. Některé jazyky (například němčina) tíhnou ke gramatikalizaci referenčních struktur více, jiné méně. Mezijazykové rozdíly jsou dobře viditelné zvláště na sémantických ekvivalentech, srov. příklady (41a–c) s víceslovným konektorem *kromě toho* s obligatorní referencí v češtině, s gramatikalizovaným anaforickým konektorem *außerdem* v němčině a konektorem *besides* bez explicitní reference v angličtině.

(41a) Je moc pozdě a **kromě toho** mám moc práce.

(41b) Es ist zu spät und **außerdem** habe ich viel Arbeit vor mir.

(41c) It's too late, and **besides**,²⁰ I've got a lot of work to do.

5. Závěr

V naší studii jsme se věnovaly popisu a rozboru diskurzních konektorů v češtině. Nejprve jsme se zabývaly bližší charakteristikou základních pojmů spojených s rozbořením textu obecně, a to pojmů text/diskurz, jednotka textu a významové vztahy v textu. Největší prostor jsme věnovaly diskuzi o diskurzních konektorech jako hlavních signalizátorech textových vztahů. Konektory mají schopnost propojovat textové jednotky a zároveň svým významem dotvářet kontext, ve kterém se vyskytují. Jako takové mají v interpretaci sémanticko-pragmatických vztahů v textu zásadní roli.

Konektory jsme vymezily především z funkčního hlediska. Neomezovaly jsme je tedy například na výrazy určitých slovních druhů či na jejich roli ve větě stavbě, ale zaměřily jsme se na jejich funkci v textu. Z tohoto hlediska pokládáme za diskurzní konektor takový výraz, který má schopnost v textu vyjadřovat určitý významový vztah a který zároveň může být použit pro vyjádření tohoto vztahu v mnoha kontextech (srov. *Už jsem velmi unavený / Venku prší / Změnili mi program koncertu / [...] Proto dnes nikam nepůjdu*).

Diskurzní konektory tvoří značně heterogenní skupinu jazykových prostředků – v textech se můžeme setkat s konektory jednoslovnými i víceslovnými, s konektory modifikovanými nebo s různými kombinacemi dvou či více konektorů, srov. příklady typu *proto*, *a*, *a právě proto*, *bud' nebo*, *však*, *kvůli tomu*, *ale*, *důvodem je*, *s jasným cílem*, *i když*, *a kvůli tomu*, *ale naopak*, *když tak*, *jedině kvůli tomu*, *hlavně proto*, *nebo*, *v případě*, *že pak*, *především vinou toho*.

¹⁹ Vypuštění referenčního prvku je v tomto případě možné pouze u kataforické varianty *místo (toho,) aby*. Referenční element u anaforického spojení *místo toho* vypustit nelze, srov. příklad: *Když došel k autu, Sara nestáhla okénko. *Místo otevřela dveře a vrhla se mu do náruče.*

²⁰ Jednoslovný konektor *besides* pochází ze staroanglické fráze *be sidan* „by the side of“ (dativní koncovku *-an* nahrazuje genitivní *-s*, srov. Brinton, 2012), a v původní formě tedy vykazoval podobné chování jako současné negramatikalizované konektory sekundární.

Proto jsme se snažily diskurzí konektory dále specifikovat, a to na základě různých hledisek. V naší práci je charakterizujeme především podle míry jejich gramatikalizace – rozlišujeme konektory primární (*proto, ale, i, a* ad.) a sekundární (*z tohoto důvodu, díky tomu, s jasným cílem* ad.).

Dále můžeme rozlišovat konektory inherentně referenční (anaforické: *potom, přesto, proto, kromě toho* ad.; kataforické: *kvůli tomu, aby; z toho důvodu, že* ad.) a inherentně nereferenční (*ale, a, nebo, však, tak, i* ad.). Konektory mohou být také složené (vyskytující se ve stejném argumentu: *i když, a proto, a kvůli tomu*, nebo tvořící korelativní páry: *sice_ale, jestliže_pak, v případě, že_tak*) a nesložené (obsahující ze synchronního hlediska pouze jeden konektor; může jít o konektory jednoslovné, gramatikalizované: *a, však*, či víceslovné, negramatikalizované: *za této podmínky*). Některé konektory mohou být modifikované (obsahující hodnotící či modální výraz, který upřesňuje význam příslušného diskurzího vztahu: *hlavně proto, z jednoduchých důvodů, s jasným cílem, možná kvůli tomu*). Z hlediska způsobu užití můžeme rozlišovat konektory vnitrovětné (podřadící konektory, souřadící konektory neadverbiálního původu jako *a, či*, konektory s nimi zkombinované: *a tak* nebo *díky tomu, že*, korelativní páry: *sice_ale*, konektory obsahující negaci: *nebo, nýbrž*) a mezivětné (konektory adverbiálního původu nebo s adverbiální funkcí: *proto, díky tomu, za tím účelem*).

Dále jsme se věnovaly možnostem výskytu a popisu chování jednotlivých typů konektorů v psaných textech. Na datech korpusů PDiT 2.0 a SYN2015 jsme představily rozbor konektorů primárních a sekundárních z hlediska četnosti jejich výskytů v textech, rozbor složených konektorů a také konektorů s fakultativní referencí.

Rozbor frekvence výskytu diskurzí konektorů ukázal, že ačkoli je repertoár konektorů v češtině značně široký, v textech se s vysokou frekvencí objevují jen některé z nich – například výskytů prvních deseti nejfrekventovanějších konektorů v PDiT (*a, však, ale, když, protože, totiž, pokud, proto, tedy, aby*) pokrývají více než polovinu (54 %) výskytů všech konektorů v tomto korpusu. Je tedy zjevné, že z velké nabídky konektorů autoři obvykle volí jen některé z nich. K tomu je vede zřejmě snaha o vynaložení minimálního úsilí při dosažení maximálního výsledku. Obecně se princip jazykové ekonomie uplatňuje v různých oblastech textových vztahů, nejen v absolutních počtech výskytů konektorů. Výsledky našeho výzkumu ukazují, že autor zpravidla volí při vyjadřování významových vztahů úspornější cestu – například u konektorů s fakultativní referencí preferuje variantu s referenčním prvkem implicitním. Užití konektoru bez explicitní reference přitom nebrání porozumění danému vztahu.

Na výběr vhodného konektoru tedy působí současně více faktorů. Na jednu stranu si autor vybírá co nejekonomičtější způsob vyjádření textového vztahu (preferuje konektory gramatikalizované, mnohdy významově obecnější či vykazující určitou míru implicitnosti), na druhou stranu musí být jím vytvořený text dostatečně srozumitelný pro adresáta – nesprávná interpretace diskurzí konektorů a vztahů, které vyjadřují, může vést k nesprávné interpretaci celého textu. Úspora při vytváření textu tedy nesmí ohrozit jeho výslednou srozumitelnost.

Závěrem můžeme konstatovat, že v jazyce existuje široké spektrum konektorů (od ustálených výrazů po volnější, variabilní konektivní konstrukce), které jako celek

procházejí vývojem a postupným procesem ustalování. Díky své schopnosti propojovat (mnohdy rozsáhlé či různě se překrývající) textové jednotky a zároveň schopnosti svým významem dotvářet kontext poskytují konektory čtenáři mnohdy nezbytnou oporu při interpretaci sémanticko-pragmatických vztahů a usnadňují tím porozumění textu jako celku.

Zdroj financování

Tato práce byla podpořena Ministerstvem školství, mládeže a tělovýchovy České republiky, projekt č. LM2018101 LINDAT/CLARIAH-CZ.

ZDROJE

- HAJIČ, J., ET AL. (2018): *Prague Dependency Treebank 3.5*. Data/software. Praha: Ústav formální a aplikované lingvistiky MFF UK. Dostupné online na adrese: <<http://hdl.handle.net/11234/1-2621>>.
- HAJIČ, J., ET AL. (2020): *Prague Dependency Treebank – Consolidated 1.0 (PDT-C 1.0)*. Praha: Ústav formální a aplikované lingvistiky MFF UK. Dostupné online na adrese: <<http://hdl.handle.net/11234/1-3185>>.
- HAJIČ, J. – PANEVOVÁ, J. – HAJIČOVÁ, E. – SGALL, P. – PAJAS, P. – ŠTĚPÁNEK, J. – HAVELKA, J. – MIKULOVÁ, M. – ŽABOKRTSKÝ, Z. – ŠEVČÍKOVÁ-RAZÍMOVÁ, M. – UŘEŠOVÁ, Z. (2006): *Prague Dependency Treebank 2.0*. Data/software. Philadelphia: Linguistic Data Consortium. Dostupné online na adrese: <<http://hdl.handle.net/11858/00-097C-0000-0001-B098-5>>.
- KŘEN, M. – CVRČEK, V. – ČAPKA, T. – ČERMÁKOVÁ, A. – HNÁTKOVÁ, M. – CHLUMSKÁ, L. – JELÍNEK, T. – KOVÁŘÍKOVÁ, D. – PETKEVIČ, V. – PROCHÁZKA, P. – SKOUMALOVÁ, H. – ŠKRABAL, M. – TRUNEČEK, P. – VONDŘIČKA, P. – ZASINA, A. (2015): *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupné online na adrese: <<http://www.korpus.cz>>.
- POLÁKOVÁ, L. – JÍNOVÁ, P. – ZIKÁNOVÁ, Š. – HAJIČOVÁ, E. – MÍROVSKÝ, J. – NEDOLUZHKO, A. – RYSOVÁ, M. – PAVLÍKOVÁ, V. – ZDEŇKOVÁ, J. – PERGLER, J. – OCELÁK, R. (2012a): *Prague Discourse Treebank 1.0*. Data/software. Praha: Ústav formální a aplikované lingvistiky MFF UK. Dostupné online na adrese: <<http://ufal.mff.cuni.cz/pdit>>.
- RYSOVÁ, M. – SYNKOVÁ, P. – MÍROVSKÝ, J. – HAJIČOVÁ, E. – NEDOLUZHKO, A. – OCELÁK, R. – PERGLER, J. – POLÁKOVÁ, L. – PAVLÍKOVÁ, V. – ZDEŇKOVÁ, J. – ZIKÁNOVÁ, Š. (2016): *Prague Discourse Treebank 2.0*. Data/software. Praha: Ústav formální a aplikované lingvistiky MFF UK. Dostupné online na adrese: <<http://hdl.handle.net/11234/1-1905>>.

LITERATURA

- ADAMEC, P. (1995): Konektivní částice a jiné textově propojovací výrazy v současné češtině. In: J. Hasil (ed.), *Přednášky z 37. a 38. běhu Letní školy slovanských studií*. Praha: Univerzita Karlova, s. 59–64.
- ARIEL, M. (1994): Pragmatic operators. In: R. E. Asher – J. M. Y. Simpson (eds.), *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press, s. 3250–3253.
- ASHER, N. (1993): *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer.
- BAUER, J. – GREPL, M. (1968): *Skladba spisovné češtiny*. 2. vyd. Praha: Státní pedagogické nakladatelství.
- BAUER, J. – GREPL, M. (1980): *Skladba spisovné češtiny*. 3. vyd. Praha: Státní pedagogické nakladatelství.

- BAZANELLA, C. (1990): Phatic connectives as intonational cues in contemporary spoken Italian. *Journal of Pragmatics*, 14, s. 629–647.
- BLAKEMORE, D. (2002): *Relevance and Linguistic Meaning: The Semantics of Discourse Markers*. Cambridge: Cambridge University Press.
- BRINTON, L. J. (2012): “The ghosts of old morphology”: Lexicalization or (de)grammaticalization? In: K. Davise – T. Breban – L. Brems – T. Mortelmans, *Grammaticalization and Language Change*. Amsterdam – Philadelphia: John Benjamins, s. 135–166.
- DANEŠ, F. (1968): Typy tematických posloupností v textu. *Slovo a slovesnost*, 29, s. 125–141.
- DANEŠ, F. (1979): O identifikaci známé (kontextově zapojené) informace v textu. *Slovo a slovesnost*, 40, s. 257–270.
- DANEŠ, F. (1985): *Věta a text*. Praha: Academia.
- DANLOS, L. – RYSOVÁ, K. – RYSOVÁ, M. – STEDE, M. (2018): Primary and secondary discourse connectives: definitions and lexicons. *Dialogue and Discourse*, 9, s. 50–78.
- FISCHER, K. (2006): Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume. In: K. Fischer (ed.), *Approaches to Discourse Particles*. Oxford: Elsevier, s. 1–20.
- GREIMAS, A.-J. (1966): *Sémantique structurale: recherche de méthode*. Paris: Larousse.
- HAIČOVÁ, E. (1975): *Negace a presupozice ve významové stavbě věty*. Praha: Academia.
- HAIČOVÁ, E. – PARTEE, B. H. – SGALL, P. (1998): *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*. Dordrecht – Boston: Kluwer Academic Publishers.
- HALLIDAY, M. A. K. – HASAN, R. (1976): *Cohesion in English*. London: Longman.
- HARPER, D. (2001–2021): *Online Etymology Dictionary*. Dostupné online na adrese: <<http://www.etymonline.com>>. Cit. 31. 3. 2021.
- HAUSENBLAS, K. (1962): O studiu syntaxe běžně mluvených projevů. In: J. Bauer – M. Grepl – J. Chloupek – R. Mrázek – R. Večerka – S. Žaža (eds.), *Otázky slovanské syntaxe: sborník brněnské syntaktické konference, 17.–21. 5. 1961*. Praha: Státní pedagogické nakladatelství, s. 313–325.
- HAUSENBLAS, K. (1984): Text, komunikáty a jejich komplexy (Zamyšlení pojmoslovně). *Slovo a slovesnost*, 45, s. 1–7.
- HAVRÁNEK, B. – JEDLIČKA, A. (1960): *Česká mluvnice*. Praha: Státní pedagogické nakladatelství.
- HOFFMANNOVÁ, J. (1983): *Sémantické a pragmatické aspekty koherence textu*. Praha: Ústav pro jazyk český ČSAV.
- HOFFMANNOVÁ, J. (1984): Typen der Konnektoren und deren Anteil an der Organisierung des Textes. In: J. Kořenský – J. Hoffmannová (eds.), *Text and the Pragmatic Aspects of Language*. Praha: Ústav pro jazyk český ČSAV, s. 101–140.
- HRBÁČEK, J. (1994): *Nárys textové syntaxe spisovné češtiny*. Praha: Trizonia.
- JÍNOVÁ, P. (2012): Nejčastější konektivní prostředky kauzálního vztahu v Pražském závislostním korpusu. *Studie z aplikované lingvistiky / Studies in Applied Linguistics*, 3 (1–2), s. 35–52.
- KARLÍK, P. (1995): *Studie o českém souvětí*. Brno: Masarykova univerzita.
- KARLÍK, P. (2016): Spojka. In: P. Karlík – M. Nekula – J. Pleskalová (eds.), *Nový encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny, s. 1736–1737. Dostupné online na adrese: <<https://www.czechency.org/slovník/SPOJKA>>.
- KRAUS, J. (1981): *Rétorika v dějinách jazykové komunikace*. Praha: Academia.
- MATHESIUS, V. (1939): O tak zvaném aktuálním členění větěném. *Slovo a slovesnost*, 5, s. 171–174.
- MATHESIUS, V. (1943): Jazykozpytné poznámky k řečnické výstavbě souvislého výkladu. *Slovo a slovesnost*, 7, s. 114–129.
- MISTRÍK, J. (1968): *Kompozícia jazykového prejavu*. Bratislava: Slovenské pedagogické nakladateľstvo.
- NEKULA, M. (1996): *System der Partikeln im Deutschen und Tschechischen*. Tübingen: Max Niemeyer Verlag.
- PEŠEK, O. (2011): *Argumentativní konektory v současné francouzštině a češtině: Systémové srovnání a analýza okurenční response*. České Budějovice: Editio Universitatis Bohemiae Meridionalis.

- POLÁKOVÁ, L. – JÍNOVÁ, P. – ZIKÁNOVÁ, Š. – BEDŘICHOVÁ, Z. – MÍROVSKÝ, J. – RYSOVÁ, M. – ZDEŇKOVÁ, J. – PAVLÍKOVÁ, V. – HAJIČOVÁ, E. (2012b): *Manual for Annotation of Discourse Relations in the Prague Dependency Treebank*. Technická zpráva. Praha: Ústav formální a aplikované lingvistiky MFF UK.
- PRASAD, R. – DINESH, N. – LEE, A. – MILTSAKAKI, E. – ROBALDO, L. – JOSHI, A. – WEBBER, B. (2008): The Penn Discourse Treebank 2.0. In: N. Calzolari – K. Choukri – B. Maegaard – J. Mariani – J. Odijk – S. Piperidis – D. Tapias (eds.), *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*. Marrakech: European Language Resources Association, s. 2961–2968.
- PRASAD, R. – JOSHI, A. – WEBBER, B. (2010): Realization of discourse relations by other means: Alternative lexicalizations. In: Ch. Huang – D. Jurafsky (eds.), *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*. Beijing: Tsinghua University Press, s. 1023–1031.
- PRASAD, R. – MILTSAKAKI, E. – DINESH, N. – LEE, A. – JOSHI, A. – ROBALDO, L. – WEBBER, B. (2007): *The Penn Discourse Treebank 2.0 Annotation Manual*. Pennsylvania: University of Pennsylvania.
- REDEKER, G. (1991): Linguistic markers of discourse structure. *Linguistics*, 29, s. 1139–1172.
- ROUCHOTA, V. (1996): Discourse connectives: what do they link? *UCL Working Papers in Linguistics*, 8, s. 199–214.
- RYSOVÁ, M. (2018): *Diskurzni konektory v češtině: Od centra k periférii*. Praha: Ústav formální a aplikované lingvistiky MFF UK.
- RYSOVÁ, M. – RYSOVÁ, K. (2014): The centre and periphery of discourse connectives. In: W. Aroonmanakun – P. Boonkwan – T. Supnithi (eds.), *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation (PACLIC 2014)*. Bangkok: Department of Linguistics, Faculty of Arts, Chulalongkorn University, s. 452–459.
- RYSOVÁ, M. – RYSOVÁ, K. (2018): Primary and secondary discourse connectives: Constraints and preferences. *Journal of Pragmatics*, 130, s. 16–32.
- SGALL, P. – HAJIČOVÁ, E. – BURÁŇOVÁ, E. (1980): *Aktuální členění věty v češtině*. Praha: Academia.
- SCHIFFRIN, D. (1987): *Discourse Markers*. Cambridge: Cambridge University Press.
- SVOBODA, K. (1972): *Souvětí spisovné češtiny*. Praha: Universita Karlova.
- ŠTĚPÁN, J. (1977): *Složitě souvětí s řetězcovou závislostí*. Praha: Univerzita Karlova.
- ŠTĚPÁN, J. (2013): *Souvětí současné češtiny*. Liberec: Technická univerzita v Liberci.
- ŠTÍCHA, F., A KOL. (2013): *Akademická gramatika spisovné češtiny*. Praha: Academia.
- TÁRNYIKOVÁ, J. (1998): Discourse markers in cross-cultural perspective. *Perspectives*, 8, s. 37–44.
- TÁRNYIKOVÁ, J. (2002): *From Text to Texture: An Introduction to Processing Strategies*. Olomouc: Univerzita Palackého v Olomouci.
- URGELLES-COLL, M. (2010): *Continuum Studies in Theoretical Linguistics: Syntax and Semantics of Discourse Markers*. London: Continuum.
- VAN DIJK, T. A. (1979): Pragmatic connectives. *Journal of Pragmatics*, 3, s. 447–456.
- ZIKÁNOVÁ, Š. – HAJIČOVÁ, E. – HLADKÁ, B. – JÍNOVÁ, P. – MÍROVSKÝ, J. – NEDOLUZHKO, A. – POLÁKOVÁ, L. – RYSOVÁ, K. – RYSOVÁ, M. – VÁCL, J. (2015a): *Discourse and Coherence: From the Sentence Structure to Relations in Text*. Praha: Ústav formální a aplikované lingvistiky MFF UK.
- ZIKÁNOVÁ, Š. – POLÁKOVÁ, L. – JÍNOVÁ, P. – NEDOLUZHKO, A. – RYSOVÁ, M. – MÍROVSKÝ, J. – HAJIČOVÁ, E. (2015b): Zachycení výstavby textu v Pražském závislostním korpusu. *Slovo a slovesnost*, 76, s. 163–197.
- ZIPF, G. K. (1949): *Human Behavior and the Principle of Least Effort*. Cambridge: Addison-Wesley Press.
- ZWICKY, A. M. (1985): Clitics and particles. *Language*, 61, s. 283–305.
- ŽAŽA, S. (1958): Složené spojky ve spisovné češtině. In: *Studie ze slovanské jazykovědy*. Praha: Státní pedagogické nakladatelství, s. 119–132.

ZUSAMMENFASSUNG

Primär- und Sekundärkonnectoren im Diskurs

Unsere Studie präsentiert die Korpusforschung über Konnectoren und andere Diskursphänomene im Tschechischen. Wir konzentrieren uns auf die Definition der Konnectoren aus funktionaler Sicht und beschreiben ihre weitere Klassifizierung in Primär- und Sekundärkonnectoren, vor allem nach ihrem Grammatikalisierungsgrad. Ein besonderes Augenmerk wird auf die Variabilität von Sekundärkonnectoren gelegt, die in verschiedenen lexikalischen Realisationen und grammatischen Varianten auftreten. Wir stellen auch die Häufigkeit der einzelnen Konnectoren in schriftlichen Texten dar und analysieren ihr syntaktisches Verhalten. Anschließend diskutieren wir das Verhältnis der Konnectoren zur Referenz (eine spezielle Gruppe von Verbindungsmitteln wird durch die anaphorischen und kataphorischen Konnectoren gebildet), weiter die Modifikationsmöglichkeiten der Konnectoren und auch die Konnectorenkumulation (an einer Stelle im Text). Für die Forschung werden insbesondere die Sprachdaten des Prague Discourse Treebank 2.0 verwendet. Das Ziel des Artikels ist es, eine detaillierte Analyse über Konnectoren im schriftlichen Tschechischen zu präsentieren und hauptsächlich Konnectoren als offene Klasse von Verbindungsmitteln anzuzeigen, die sowohl grammatikalisierte Ausdrücke als auch mehr freie lexikalische Phrasen enthält.

Ústav formální a aplikované lingvistiky MFF UK

Malostranské náměstí 25, Praha 1

<magdalena.rysova@ufal.mff.cuni.cz>

<rysova@ufal.mff.cuni.cz>