

Univerzita Karlova v Praze
Filozofická fakulta
Ústav českého jazyka a teorie komunikace

BAKALÁŘSKÁ PRÁCE

Jiří Pergler

**Koreferenční řetězce s nespécifickou a generickou
referencí v češtině**

Coreferential Chains with Non-specific and Generic Reference in Czech

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně a že jsem uvedl všechny použité prameny a literaturu.

V Praze dne 28. 6. 2010

Jiří Pergler

Abstrakt

Práce se zabývá vztahy koreference v textu a jejich klasifikací založenou na typech reference. Ve 2. kapitole uvádím základní principy anotace textové koreference v Pražském závislostním korpusu s důrazem na uplatňované rozdělení koreferenčních vztahů na typ 0 a typ NR, ve 3. kapitole pak upozorňuji na některé problematické jevy spojené s tímto rozdělením, které vyplývají především z příliš široce vymezeného typu NR.

Ve 4. a 5. kapitole představuji některé významné práce věnující se problematice reference a generičnosti. V 6. kapitole následně navrhuji vlastní typologii referenčních vztahů s rozdělením reference na čtyři typy podle kritéria generičnosti (reference individuální a generická) a kritéria specifčnosti (reference specifická a nespecifická). Věnuji se též možnosti jednotlivých referenčních výrazů účastnit se vztahů koreference a dokládám, že navržené schéma je pro popis některých složitějších jevů vhodnější než typologie uplatňovaná v Pražském závislostním korpusu. V 7. kapitole pak ukazuji použití navrženého schématu na rozboru koreferenčních vztahů v několika konkrétních textech.

Klíčová slova: koreference, reference, generičnost, specifčnost, Pražský závislostní korpus

Abstract

The thesis dwells on the coreference relations in the text and their classification based on the reference types. In Chapter 2, I quote the main principles of the annotation of the textual coreference in the Prague Dependency Treebank, focusing especially on the division of the types of the coreferential relations, the type 0 and the type NR. In Chapter 3, I point out some problems connected with this division, which result particularly from the too broad definition of the type NR.

In Chapters 4 and 5, I introduce some of the important literature in the field of reference and genericity. In Chapter 6, I propose my own typology of the referential relations which consists of four types based on the criterion of genericity (individual and generic reference) and the criterion of specificity (specific and non-specific reference). I also deal with the possibility of the referential expressions to participate in the coreferential relations and demonstrate that the proposed scheme is more suitable for the description of some complicated phenomena than the scheme employed in the Prague Dependency Treebank. In Chapter 7, I show how the proposed scheme can be used by the analysis of the coreferential relations in several texts.

Keywords: coreference, reference, genericity, specificity, Prague Dependency Treebank

Obsah

1. Úvod	1
2. Zpracování koreference v Pražském závislostním korpusu	3
2.1 Pojetí koreference	3
2.2 Zachycení koreference	5
2.3 Typy koreference	9
2.4 Koreference u abstrakt a dějových deverbativ	11
3. Problémy spojené s rozdělením typů reference v PDT	13
3.1 Nespecifická reference s typem 0	13
3.2 Reference k neohrazeným množinám prvků	14
3.3 Referent vymezený podmínkovou větou	15
4. Vybrané české přístupy k referenci	16
4.1 Palek, 1988	16
4.2 Hlavsa, 1975	18
4.3 Adamec, 1980	21
4.4 Shrnutí popsaných koncepcí	23
5. Pojetí generičnosti v The Generic Book	25
5.1 Generičnost propozic	25
5.2 Generičnost jmenných frází	27
5.3 Kombinace obou typů generičnosti	31
6. Model reference a koreference	33
6.1 Typy reference	33
6.2 Specifická generická reference	36
6.2.1 Forma generických jmenných frází	36
6.2.2 Význam generických jmenných frází	37
6.3 Nespecifická individuální reference	41
6.3.1 Jmenné fráze s rysem hypotetičnosti	41

6.3.2 Jmenné fráze s rysem nevymezenosti	44
6.4 Nespecifická generická reference	47
6.5 Možnosti koreference	48
6.6 Řešení případů uvedených v kapitole 3.	49
7. Analýza korpusových textů	51
7.1 Analýza textu č. 1	51
7.2 Analýza textu č. 2	53
7.3 Analýza textu č. 3	54
8. Závěry	57
Literatura	59

1. Úvod

Tato práce se zabývá vztahy koreference v češtině a jejich klasifikací založenou na typech reference. Inspiraci k výběru tématu mi poskytl projekt anotace rozšířené textové koreference a asociační anafory v Pražském závislostním korpusu (PDT),¹ na kterém se od roku 2008 podílím jako anotátor textové koreference a anaforických vztahů. Kromě odborné literatury a jazykového materiálu češtiny mi tak budou cenným zdrojem informací také zkušenosti získané během procesu anotace. Cílem práce je upozornit na některé problematické případy spojené s typologií koreferenčních vztahů, která je ve zmíněném projektu uplatňována, a navrhnout možné řešení těchto problémů.

Základním východiskem v projektu anotace v PDT, které přebírám i pro tuto práci, je důsledné oddělení vztahů koreference od jiných textových vztahů, především od vztahu anafory (resp. obecně endoforického odkazování). Endofora je chápána jako odkaz jednoho výrazu textu na jiný výraz v tomtéž textu, jde tedy o vztah asymetrický, neboť pokud výraz A odkazuje na výraz B, pak neplatí, že by také výraz B odkazoval na výraz A. Anafora je pak zvláštním případem endofory, pro který platí, že odkazující výraz je v textu předcházen výrazem, na který je odkazováno. Naopak koreference označuje vztah mezi výrazy, které mají vztah reference k téže entitě, a proto se jedná o vztah symetrický. Vztahy koreference a endofory se sice často vyskytují současně (jako např. ve větě *Máme dodavatele, s nimiž spolupracujeme*), avšak existují též případy, kdy mezi dvěma výrazy ve vztahu endofory vztah koreference není (např. ve větě *Máme již nějaké dodavatele, ale potřebovali bychom jich více²*). Kromě toho existují též koreferenční vztahy bez endofory, o tom blíže v části 2.1.

Pokud jde o uspořádání práce, nejprve podám stručný přehled anotace koreferenčních vztahů v Pražském závislostním korpusu (kapitola 2). Kromě vlastních zkušeností s projektem budu čerpat zejména z disertační práce Nedoluzhko (2010), kde je toto téma zpracováno podrobněji, a také z publikace Mikulová et al. (2005), kde jsou popsány principy anotace na tektogramatické rovině Pražského závislostního korpusu. Pozornost budu věnovat jak formální stránce zachycování koreferenčních vztahů (2.2), tak (především) jednotlivým typům koreference (2.3). Po tomto představení upozorním na některé problémy spojené s popsáním schématem (kapitola 3), o jejichž řešení budu v práci usilovat.

¹ Tento projekt je podrobně představen v disertační práci Nedoluzhko (2010).

² Zatímco spojení *nějaké dodavatele* referuje ke konkrétním dodavatelům, které již máme, referent výrazu *jich* zahrnuje i hypotetické dodavatele, které bychom teprve *mohli* mít, a jde tedy o jinou entitu než v případě výrazu *nějaké dodavatele*.

Ve 4. kapitole představím některé inspirativní práce českých lingvistů týkající se reference a jejích typů, konkrétně půjde v části 4.1 o monografii B. Palka (1988), v části 4.2 o publikaci Z. Hlavsy (1975) a v části 4.3 o článek P. Adamce (1980). V části 4.4 budou tyto práce navzájem porovnány, přičemž si budu všimnout též jejich vztahu k typům koreference uplatňovaným v PDT. Jelikož značnou část zkoumané problematiky představuje reference generická, bude 5. kapitola věnována otázce generičnosti, jak je zpracována v *The Generic Book* Carlsona et al. (1995)³, důraz bude kladen zvláště na rozlišení dvou různých typů generičnosti.

V 6. kapitole, která představuje vlastní cíl práce, pak navrhnu alternativní model jednotlivých typů reference a koreference, který bude zohledňovat poznatky získané z uvedených odborných prací. V tomto modelu hraje zásadní roli především jasné oddělení rysů generičnosti a specifičnosti, na základě kterých budou vymezeny čtyři typy reference. V části 6.5 shrnu, jaké koreferenční vztahy (z hlediska těchto typů) mohou v českých textech vznikat, a v části 6.6 vysvětlím, jak je v rámci tohoto nového schématu možné popsat problematiku jevy vyličené ve 3. kapitole.

Při výkladu zmíněného modelu v 6. kapitole se budu pro větší názornost opírat zejména o vlastní příkladové věty. Aby bylo zřejmé, jak tento model funguje na konkrétních existujících textech, provedu v 7. kapitole analýzu koreferenčních vztahů ve třech textech vybraných z Pražského závislostního korpusu. Tyto texty jsou vybrány tak, aby na nich bylo možné demonstrovat co největší množství relevantních jevů.

³ Na tuto publikaci budu dále odkazovat jako na GenBook (1995).

2. Zpracování koreference v Pražském závislostním korpusu

V této kapitole shrnu základní principy zpracování koreference v Pražském závislostním korpusu (PDT). Vztah koreference je zde anotován ve dvou fázích. V první fázi byly anotovány vztahy tzv. gramatické koreference a některé vztahy koreference textové, této anotaci je v současné době podroben celý korpus PDT a její zásady jsou shrnuty v publikaci *Anotace na tektogramatické rovině Pražského závislostního korpusu* (Mikulová et al., 2005). Druhá fáze, jejíž dokončení je plánováno do konce roku 2010, je podrobně popsána v disertační práci Anny Nedoluzhko *Rozšířená textová koreference a asociační anafora (Koncepte anotace českých dat v Pražském závislostním korpusu)* (Nedoluzhko, 2010).

2.1 Pojetí koreference

Koreference je v PDT chápána jako vztah mezi „výrazy, které odkazují na tentýž objekt skutečnosti, pojem nebo situaci“ (Nedoluzhko, 2010, s. 2), těmito výrazy jsou zpravidla jmenné fráze. Termín koreference je odvozen od termínu reference, přičemž referencí se v PDT rozumí „vztah výrazů k předmětům nebo situacím reálného světa“ (Mikulová et al., 2005, s. 932). Vztahy koreference tvoří v PDT součást tektogramatické roviny.⁴

Koreference je vztah symetrický, tj. pokud je výraz A koreferenční s výrazem B, pak je také výraz B koreferenční s výrazem A, a zároveň vztah tranzitivní, tj. pokud výraz A je koreferenční s výrazem B a výraz B je koreferenční s výrazem C, pak je též výraz A koreferenční s výrazem C (op. cit., s. 95).⁵ Těmito vlastnostmi se vztah koreference liší od vztahu anaforického, který je asymetrický a není tranzitivní.⁶

Rozlišovány jsou dva druhy koreference: gramatická a textová. Gramatická koreference je takový koreferenční vztah, u kterého „je možné určit koreferovaný člen na základě gramatických pravidel“ (Mikulová et al., 2005, s. 935). Vztahy gramatické koreference jsou v citované publikaci (op. cit., s. 935–1037) rozděleny do šesti skupin:

⁴ K pojetí a struktuře tektogramatické roviny viz např. Mikulová et al. (2005).

⁵ O výjimce z tohoto pravidla viz níže.

⁶ Anaforický vztah je chápán jako „odkaz na to, co bylo již v předchozím kontextu zmíněno“ (Nedoluzhko, 2010, s. 67). Nedoluzhko (2010, s. 68) v této souvislosti uvádí tři podmínky, při jejichž splnění se často setkáváme s koreferencí bez přítomnosti anaforického vztahu: členy koreferenčního páru se nacházejí v rématu, vyskytují se v různých diskurzivních jednotkách nebo jsou od sebe ve velké textové vzdálenosti.

1. *Koreference zvratných zájmen*, tj. vztah mezi osobním či přivlastňovacím zvratným zájmenem a jeho antecedentem, např. vztah mezi slovem *sebe* a *práce* ve větě *Žádná práce by neměla být konána pro sebe samu, ale pro potěšení vlastní a druhých.*⁷

2. *Koreference vztažných prostředků*, tj. vztah mezi vztažným zájmenem či příslovcem a jeho antecedentem, např. vztah mezi *v městečku* a *kam* ve větě *Film se odehrává na venkově, v městečku Sardent, kam se po letech vrací – aby si tam léčil zdraví – tamější rodák.*

3. *Koreference u doplnění s dvoji závislostí vyjádřených slovesnou formou*, totiž vztah, do kterého vstupuje aktant slovesného doplnění vyjádřeného trpným participiem, přechodníkem, infinitivem nebo určitým slovesným tvarem, pokud má toto doplnění funkci doplňku nebo patientu či efektu se shodou s jiným doplněním řídicího slovesa. Antecedentem je v tomto případě jiné doplnění řídicího slovesa. Např. vztah mezi slovem *mužstvo* a nevyjádřeným patientem slovesa *porazit* ve větě *Mužstvo zůstává neporaženo.*⁸

4. *Kontrola*, tedy koreferenční vztah, ke kterému dochází u tzv. sloves kontroly. Antecedentem je aktant slovesa kontroly, anaforem je povrchově nevyjádřený aktant infinitivu či deverbativního substantiva, které je součástí valenčního rámce slovesa kontroly. Např. vztah mezi slovem *mi* a nevyjádřeným aktorem slovesa *dostat se* ve větě *Navíc mám Spartě co odvádět za to, že mi před lety umožnila dostat se do velkého fotbalu.*

5. *Kvazikontrola* je vztah podobný vztahu kontroly, liší se však tím, že příslušným aktantem řídicího slovesa je substantivum (mající valenci). Např. vztah mezi slovem *Janovi* a nevyjádřeným patientem substantiva *ochranu* ve větě *Poskytl Janovi ochranu.*

6. *Koreference v recipročních konstrukcích*. Reciproční konstrukce jsou založeny na tom, že „dvě různá valenční doplnění (aktanty, popřípadě obligatorní volná doplnění) jsou postavena do symetrického vztahu“ (Mikulová et al., s. 158). V tektogramatické struktuře jsou tato doplnění zachycena pouze v jedné syntaktickosémantické funkci, pro druhou z nich je vytvořen nový uzel. Koreference pak nastává mezi symetrickými doplněními a tímto novým uzlem. Např. ve větě *Sultáni se vystřídali na trůnu* má uzel zastupující výraz *sultáni* funkci aktora řídicího slovesa, zatímco funkci patientu má s ním koreferenční nově vytvořený uzel.

Textová koreference je takový koreferenční vztah, u kterého určit koreferovaný člen na základě gramatických pravidel nelze. V první fázi anotace je tento vztah anotován u takových

⁷ Věty sloužící jako příklad vět s přítomností gramatické koreference přebírám z publikace Mikulová et al. (2005).

⁸ Pokud je aktant řídicího slovesa vyjádřen infinitivem, je takový vztah podobný vztahu kontroly. Liší se od něj tím, že „pozice infinitivu zde může být realizována i závislou slovesnou klauzí se závaznou gramatickou koreferencí subjektu řídicího slovesa této klauze“ (Mikulová et al., 2005, s. 954).

dvojic uzlů, kde koreferujícím členem je osobní či přivlastňovací zájmeno pro třetí osobu, ukazovací zájmeno *ten* nebo při elipse, pokud elidovaným výrazem není zájmeno pro první či druhou osobu (Mikulová et al., s. 1037–1038). Ve druhé fázi jsou anotovány všechny ostatní vztahy textové koreference.

2.2 Zachycení koreference

Koreference je v PDT zachycována prostřednictvím orientovaných hran mezi jednotlivými uzly tektogramatické roviny. Vztah koreference je přitom realizován mezi úplnými podstromy uzlů spojených těmito hranami (tedy mezi celými jmennými frázemi), nikoli mezi samostatnými uzly.^{9,10} Hrana je orientována směrem od jmenné fráze, která je v povrchovém slovosledu umístěna více vpravo, ke jmenné frázi, která stojí více vlevo (v případě, že příslušný uzel nemá protějšek v povrchové podobě věty, je rozhodující jeho pozice ve slovosledu hloubkovém¹¹). Výjimkou jsou případy, kdy je kromě koreference mezi příslušnými uzly i vztah katafory, v těchto případech je hrana orientována ve smyslu kataforického odkazování, tedy obráceně.

Při zachycování koreference se dále uplatňuje tzv. *princip dodržování maximálního koreferenčního řetězce* (Nedoluzhko, 2010, s. 61–63). Je založen na tom, že koreferenční hrana není konstruována u všech dvojic uzlů ve vztahu koreference, nýbrž jen mezi těmi uzly, které stojí v posloupnosti koreferenčních uzlů (uspořádané podle povrchového slovosledu) bezprostředně vedle sebe. Např. pokud A, B a C jsou jmenné fráze, které se v povrchovém slovosledu vyskytují v tomto pořadí, přičemž všechny dvojice vybrané z těchto uzlů se nacházejí ve vztahu koreference, jsou v tektogramatickém stromě vytvořeny pouze dvě koreferenční hrany, jedna od uzlu reprezentujícího jmennou frázi B k uzlu reprezentujícímu A, druhá od uzlu reprezentujícího C k uzlu reprezentujícímu B.¹²

V disertační práci Nedoluzhko (2010) je dále uvedena klasifikace výrazů (z hlediska slovnědruhového), které mohou vstupovat do vztahu koreference (s. 73–90). Jelikož je relace

⁹ Jde o tzv. princip maximální velikosti koreferující jednotky, viz Nedoluzhko (2010, s. 63–64).

¹⁰ Účast jmenné fráze na koreferenčním vztahu však nevyklučuje, aby se některá část této jmenné fráze rovněž účastnila (jiného) koreferenčního vztahu. Děje se tak tehdy, pokud jmenná fráze obsahuje takovou část, která má vlastní referenci (tj. referenci odlišnou od reference celé jmenné fráze). Např. jmenná fráze „ředitel tohoto podniku“ obsahuje část „tohoto podniku“, která má vlastní referenci a může tak samostatně vstupovat do vztahu koreference. Naopak fráze „ředitel Novák“ žádnou takovou část nemá.

¹¹ K hloubkovému slovosledu viz např. Mikulová et al. (2005, s. 183).

¹² Tento princip však není dodržován vždy. Výjimku mohou tvořit jednak hrany pro koreferenci spojenou s kataforou, které jsou povinně orientovány ve smyslu kataforického odkazování, a dále některé vztahy

koreference součástí tektogramatické roviny, je v této klasifikaci využito rozdělení autosémantik na tzv. sémantické slovní druhy, které popisuje např. Mikulová et al. (2005, s. 37–89). Těmito sémantickými slovními druhy jsou sémantická substantiva, sémantická adjektiva, sémantická adverbia a sémantická slovesa.

Vztahů koreference se účastní především sémantická substantiva, mohou do nich vstupovat (s některými omezeními) všechny jejich typy:

1. pojmenovací sémantická substantiva (např. *stůl, Petr, matčin*)
2. pojmenovací sémantická substantiva s odděleně reprezentovaným příznakem negace (tj. některá deverbativní substantiva, např. *nakupování, zralost*; jde o podtyp typu 1)
3. určitá pronominální sémantická substantiva ukazovací (tedy ukazovací zájmena v pozici syntaktického substantiva, např. *ten, toto*)
4. určitá pronominální sémantická substantiva osobní (osobní a přivlastňovací zájmena, např. *on, její*). Výjimku v této skupině tvoří zájmena pro první a druhou gramatickou osobu, která do vztahů koreference nevstupují, resp. tento vztah u nich není anotován.
5. neurčitá pronominální sémantická substantiva (vztažná zájmena, např. *kdo, co*, zájmena od nich odvozená, např. *někdo, něčí*, a zájmena *každý, všechen*; všechna tato zájmena musí být v pozici syntaktického substantiva)
6. určitá kvantifikační sémantická substantiva (číslovky základní v pozici syntaktického substantiva, např. *tři*, a číslovky dílové, např. *třetina*)

V menší míře se koreferenčních vztahů anotovaných v PDT mohou účastnit i ostatní sémantické slovní druhy. V případě sémantických adjektiv jde především o typ sémantických adjektiv pojmenovacích, a to v následujících případech:

1. adjektivum je odvozeno od vlastního jména (do koreferenčního vztahu tak vstupuje např. dvojice *Slovensko – slovenský*)
2. adjektivum má přivlastňovací význam; děje se tak především v takových případech, kdy od příslušného substantiva nelze vytvořit přivlastňovací adjektivum (v PDT chápané jako sémantické substantivum), např. *dětský* ve spojení *dětská mysl*, nikoli však ve spojení *dětské hřiště*, kde nejde o přivlastňovací význam
3. jde o adjektivum s významem odkazovacím (např. *tamní, zdejší*)

gramatické koreference, jejichž antecedent je dán gramatickými pravidly, a může proto dojít k situaci, že koreferenční hrana je od dvou různých uzlů vedena ke stejnému antecedentu.

4. není splněna žádná z předcházejících podmínek, avšak adjektivum má výraznou referenční schopnost (např. adjektivum *státní* ve spojení státní instituce)

Pokud jde o sémantická adverbia, účastní se koreferenčních vztahů sémantická adverbia pronominální, např. *tady*, *tehdy*, především jde tedy o adverbia referující k místu nebo času. Co se týče výrazu *tak*, zapojuje se do koreferenčních vztahů pouze tehdy, je-li na tektogramatické rovině reprezentováno komplexním uzlem (tj. uzlem pro autosémantické výrazy), nikoli uzlem s funktoem PREC¹³.

Zvláštní situace nastává v případě sémantických sloves, která mají, jak uvádí Nedoluzhko (2010, s. 84), ze všech anotovaných slovních druhů „nejmenší, ba žádný referenční potenciál“. Vztah koreference se proto v případě sloves zpravidla neanotuje. Výjimku však tvoří případy, kdy je sloveso (včetně svých doplňení) antecedentem anaforické jmenné fráze, v těchto případech se jako koreference označuje příslušný anaforický vztah. Koreferenční hrana tak nikdy nemůže být orientována směrem od slovesa, může však být orientována ke slovesu.¹⁴ Jako příklad uvádí Nedoluzhko (2005, s. 84) dvojici vět *Jistotu v tomto směru dávají nejnovější kroky vlády SR, která se rozhodla zavést již před časem avizovanou desetiprocentní dovozní přírážku na zboží zahraniční provenience. Byť má na tento krok určité právo, v daném okamžiku však vyznívá jako tvrdé politické rozhodnutí vlády, která se snaží velice rezolutními administrativními kroky zredukovat mnohamilionové pasívum v obchodní výměně s ČR.* Jmenná fráze *tento krok* se v pojetí PDT chápe jako koreferenční se slovesem *zavést* a jeho doplňeními.

Zvláštní pozornost vyžaduje zachycení koreference u koordinačních a apozičních skupin. V případě, že se vztahu koreference účastní celá koordinační skupina, zapojuje se do řetězce koreferenčních hran řídicí uzel koordinačního spojení. Podobná situace nastává v případě vztahu apozice: obě části apozičního vztahu jsou chápány jako vzájemně koreferenční,¹⁵ a proto se do koreferenčních řetězců zapojuje vždy řídicí uzel apozičního vztahu, nikoli jednotlivé nominální fráze v tomto vztahu.

Pokud jde o dvojice nominálních frází vstupujících do vzájemného vztahu koreference, mohou z hlediska jejich lexikálního obsazení nastat následující případy:

¹³ Komplexním uzlem je výraz *tak* tehdy, pokud je chápán jako slovesné doplnění vyjadřující způsob. Funktor PREC má přiřazen takový uzel, „který reprezentuje výraz signalizující návaznost klauze na předcházející kontext“ (Mikulová, 2005, s. 534). Pokud má uzel reprezentující výraz *tak* funktoem PREC, není tedy tento výraz chápán jako určení způsobu, ale jako navazovací výraz. Hranice mezi oběma případy však není zcela zřetelná.

¹⁴ Tyto případy tak tvoří výjimku z pravidla, že koreference je relací symetrickou a tranzitivní.

1. Antecedent nebo anafor (případně obojí) má podobu zájmena nebo uzlu zastupujícího povrchově nevyjádřený výraz (do této skupiny patří mimo jiné všechny koreferenční vztahy anotované v první fázi anotace).

2. Opakování stejného pojmenování (např. *společnost – společnost*)

3. Opakování stejného pojmenování s determinátorem (např. *žena – ta žena*)

4. Stejný řídicí uzel, ale rozdílné podstromy (např. *Ministerstvo financí ČR – ministerstvo*)

5. Různá pojmenování (synonyma, hyponymum – hyperonymum, apelativum – proprium, jiná různá pojmenování)

6. Antecedentem je sloveso.

Závěrem této části je třeba uvést, jaké dvojice výrazů se v PDT za koreferenční nepovažují. Jde o případy, které popisuje Nedoluzhko (2010, s. 99–102):

1. koreference otázkového slova a odpovědi – např. u dvojice vět *Kdo přišel? Přišel Petr* se dvojice výrazů *kdo* a *Petr* nepokládá za koreferenční

2. zájmena pro první a druhou osobu – u těchto zájmen není koreference anotována, neboť tato zájmena nevstupují do anaforických vztahů v textu

3. jmenné fráze v pozici přísudku – funkcí těchto jmenných frází je predikace, nikoli reference, proto se nepokládají za referenční a nevstupují do vztahů koreference (např. ve větě *Petr je programátor* není zaznamenán vztah koreference mezi výrazy *Petr* a *programátor*)

4. jmenné fráze v apozici – jak bylo uvedeno výše, v případě vztahu apozice se do koreferenčního řetězce zapojuje pouze řídicí uzel celé apozice

5. výrazy reprezentované uzlem s funktorem ID – jde o substantiva, „která označují metajazykově užitá slova a vlastní jména a názvy zvířat, předmětů a jevů“ (Mikulová et al., 2005, s. 555), tedy o výrazy, které nereferují, nýbrž pojmenovávají entitu, ke které referuje jejich řídicí uzel. Do koreferenčních vztahů se tedy zapojuje pouze tento jejich řídicí uzel (např. v případě jmenné fráze *město Praha* může do vztahů koreference vstupovat pouze výraz *město*).

¹⁵ Tento předpoklad však lze považovat za poněkud problematický, neboť v apozičních vztazích typu „u některých druhů zboží, například okurek“ (PDT) si je možno koreferenci mezi oběma částmi vztahu představit jen obtížně.

2.3 Typy koreference

Vztahy koreference se v PDT dělí do dvou typů, a to na základě typu reference výrazů, které do těchto vztahů vstupují. Jde o *koreferenci nominálních frází se specifickou referencí* (tzv. typ 0)¹⁶ a *koreferenci nominálních frází s generickou nebo nespecifickou referencí*¹⁷ (typ NR) (Nedoluzhko, 2010, s. 108). Přestože specifická či nespecifická reference je charakteristikou jmenných frází, které do koreferenčních vztahů vstupují¹⁸, nejsou tyto typy v PDT přiřazeny jmenným frázím, nýbrž přímo jednotlivým relacím koreference, a to na základě typu reference nominálních frází, které se příslušného vztahu účastní.¹⁹ Jak však uvádí Nedoluzhko (2010, s. 105), „hranice mezi NP se specifickou a nespecifickou referencí není vždy úplně zřetelná“. V takových případech je pro vztah koreference u jmenných frází s primárně předmětným významem preferován „defaultní“ typ 0 (Nedoluzhko, 2010, s. 108).²⁰

Pokud jde o vztah typu 0, je jeho prototypickým případem „koreference dvou jmenných frází se specifickou referencí“ (Nedoluzhko, 2010, s. 109). Specifickou referencí se přitom rozumí „odkazování ke konkrétnímu existujícímu, reálnému referentu a objektu skutečnosti“ (op. cit., s. 109), referent tedy musí být vybrán z množiny objektů, ke kterým daný výraz může potenciálně odkazovat.²¹ Tento typ reference (a koreference) je snadno identifikovatelný především u jmenných frází s konkrétním významem, jako např. výrazů *maminka* ve dvojici vět *Helena poprosila maminku, aby na ni počkala. Maminka však řekla, že nemůže* (Nedoluzhko, 2010, s. 102).

Koreference typu 0 se však může (v rozporu s výše uvedeným vymezením) vyskytovat i v případech, kdy antecedent má nespecifickou a zároveň negenerickou referenci, tedy referuje na nevybraný objekt. Děje se tak tehdy, pokud je součástí anaforu určitý identifikátor

¹⁶ Ne všechny dvojice nominálních frází ve vztahu koreference tohoto typu však musí mít specifickou referenci, o tom viz dále.

¹⁷ Vzájemný vztah mezi referencí generickou a nespecifickou zde není zcela objasněn, oba pojmy Nedoluzhko (2010) dále používá (s ne zcela jasnou distribucí) pro označení reference výrazů, u nichž je anotována koreference typu NR.

¹⁸ Pochopitelně mohou mít specifickou či nespecifickou referenci i jmenné fráze, které se žádného koreferenčního vztahu neúčastní.

¹⁹ Jak bude uvedeno dále, mohou do vztahu koreference vstupovat i dvojice nominálních frází navzájem se lišících co do typu koreference. V takovém případě je pro typ koreferenčního vztahu rozhodující typ reference anaforu. Takové řešení však pochopitelně vede k tomu, že v anotaci není zaznamenán typ reference antecedentu, pokud tento sám není anaforem jiného koreferenčního vztahu.

²⁰ Toto řešení není zřejmých důvodů není zcela ideální, je však voleno s ohledem na tzv. „princip důslednosti“ (Nedoluzhko, 2010, s. 60–61), jehož smyslem je formulace maximálně přesných pravidel tak, aby byla výsledná anotace maximálně nezávislá na subjektivním názoru anotátora. Cílem je dosažení co největší jednotnosti anotace.

²¹ V případě, že tyto podmínky nejsou splněny, se jedná o referenci nespecifickou.

(Nedoluzhko, 2010, s. 112), neboť v takových případech se reference anaforu chápe jako specifická. Nedoluzhko (2010) uvádí jako příklad věty *Přesto si značky mohl všimnout jen někdo velice pozorný [...] a ani ten velice pozorný člověk by jim patrně nepřikládal žádný význam* nebo *Poprosím o to některého kolegu a ten mi to řekne*. Podtržené koreferenční páry jsou chápány tak, že antecedent má nespecifickou negenerickou referenci, anafor má specifickou referenci (jeho reference je specifikována kontextem a podmíněna určitým identifikátorem) a vztah koreference náleží k typu 0.²²

Druhým typem koreference je typ NR, takových vztahů se účastní především jmenné fráze s generickou referencí. Generickou referencí rozumí Nedoluzhko (2010, s. 103) referenci na typ objektů, přičemž od reference na konkrétní vybrané objekty se liší „tím, že se nemusí vztahovat na všechny objekty daného typu“. Pro rozhodování o rysu generičnosti v případě jednotlivých jmenných frází jsou zavedeny dva testy. Podle prvního z nich „jmenná fráze *X* je generická, pokud *X* může být použit v konstrukci „*X* jako (typický) *Y*“, „*X* jako druh (forma) *Y*“ (Nedoluzhko, 2010, s. 114).²³ Druhý test je založen na možnosti použít jmennou frázi predikativně, tedy nahradit jmennou frázi *X* spojením „to/ten, co/kdo je *X*.“ Jako generické jsou chápány též uzly závislé na kontejnerech²⁴ (např. *mléko* ve spojení *sklenice mléka*). Dvě generické jmenné fráze jsou koreferenční (v rámci typu NR), „pokud referují ke stejnému typu objektů stejného rozsahu“ (Nedoluzhko, 2010, s. 114).²⁵

Koreference generických jmenných frází je prototypickým případem koreference typu NR. Tento typ se však vyskytuje i v některých případech nespecifické reference negenerické. Jde o případy, kdy se v textu vyskytují dvě jmenné fráze s nespecifickou referencí, které mají stejnou extenzi, ale bez determinátoru (Nedoluzhko, 2010, s. 117).²⁶ Jako příklad uvádí autorka dvojici jmenných frází *rodina* ve větách *Když si dítě bude přát, aby se o jeho problému nikdo z rodiny nebo školy nedozvěděl, musíme to respektovat, vysvětluje Jana Drtilová . [...] Většinou se stává, že dítě ani nechce, aby se rodina dozvěděla, že se nám ozval*. Negenerické nespecifické jmenné fráze však nemohou vstupovat do vztahu koreference se jmennými frázemi generickými. Pokud by se v příslušném textu za uvedenými větami

²² Vzhledem k tomu, že typ koreference vyplývá z typu reference anaforu (bez ohledu na typ reference antecedentu), v anotaci v těchto případech není zachyceno, že anafor má nespecifickou referenci. V anotaci pak tyto koreferenční řetězce nejsou nijak odlišeny od „běžných“ koreferenčních řetězců se specifickou referencí.

²³ Tento test je převzatý z RACHILINA Ekaterina; KREJDLIN Grigory E. *Denotativnyj status otglagoľnych imen. Naučno-techničeskaja informacija*, ser. 2, roč. 12, 1981, s.17–22.

²⁴ V tektogramatické struktuře mají tyto uzly funktor MAT.

²⁵ Stejným rozsahem zde není myšlena obligatorní shoda všech objektů daného typu. Smyslem této podmínky je eliminace vztahů mezi páry typu „lidé“ a „bohatí lidé“, které nejsou chápány jako koreferenční.

²⁶ V případě, že by druhá z nominálních frází obsahovala determinátor, šlo by o koreferenci typu 0.

vyskytovala věta *Linka by neměla rodinu nahrazovat, ale doplňovat*, výraz *rodinu* by nevstupoval do koreferenčního vztahu se žádnou z nominálních frází dříve uvedených vět.²⁷

Jak již bylo uvedeno, může v některých případech dojít k tomu, že součástí jednoho koreferenčního řetězce jsou uzly s různým typem reference. Stává se tak především tehdy, mají-li příslušné jmenné fráze různé lexikální obsazení. Nedoluzhko (2010, s. 119) uvádí např. dvojici jmenných frází *Romové – tento národ*, kde první fráze má referenci generickou, druhá specifickou. V těchto případech se typ koreferenčního vztahu řídí typem reference anaforu.

2.4 Koreference u abstrakt a dějových deverbativ

Rozlišení koreference na typ 0 a typ NR se dodržuje nejen u substantiv s konkrétním významem, nýbrž také u abstrakt a dějových deverbativ, která rovněž podléhají anotaci koreference.²⁸ Tyto výrazy však „stojí na pomezí mezi referujícími jmény s předmětovým významem a predikujícími slovními druhy, jako jsou např. adjektiva, adverbia a slovesa“ (Nedoluzhko, 2010, s. 119), a proto není vždy snadné rozhodnout o typu jejich reference.

Při rozhodování o typu reference je ovšem možné využít toho, že abstrakta (a podobně také dějová deverbativa) mají valenci.²⁹ Pro tyto výrazy se tedy při anotaci uplatňuje další kritérium: pokud má abstraktum či dějové deverbativum sémantické valenční doplnění se specifickou referencí, má toto slovo zpravidla specifickou referenci (Nedoluzhko, 2010, s. 121). Toto kritérium však není zcela bezproblémové, neboť některá abstrakta či deverbativa mohou mít více sémantických valenčních pozic, které mohou být obsazeny výrazy s různou referencí. Např. substantivum *láska*, které má pozice pro aktora a pacienta, se může vyskytovat v různých kontextech. Ve větě *Láska je pocit* jde zřejmě o referenci generickou, ve spojení *Petrova láska k Marii* zřejmě o referenci specifickou. Není ale jasné, o jaký typ reference se jedná např. ve větě *Láska k české vlasti je důležitý cit*, kde je specifickým výrazem obsazena jen jedna z obou pozic (v tomto případě pozice pacienta).³⁰ Navíc lze brát

²⁷ V tomto ohledu se model koreference v PDT dostává do poněkud paradoxní situace, kdy dvě jmenné fráze nejsou koreferenční z důvodu odlišného typu reference, přestože obě přináležejí ke stejnému typu, tedy typu NR.

²⁸ Jak však uvádí Nedoluzhko (2010, s. 120), klasifikace lexika na abstraktní a konkrétní je u různých autorů velice odlišná, navíc mezi oběma skupinami pravděpodobně nelze vést přesnou hranici.

²⁹ Tento poznatek přebírá (Nedoluzhko, 2010, s. 121) z publikace MENDOZOVA, Imke. *Nominaldetermination im Polnischen. Die primären Ausdrucksmittel*. München, 2004. Nepublikovaná habilitační práce. Dostupné na http://www.slavistik.uni-muenchen.de/pers_pages/mendoza.htm

³⁰ První příkladová věta je převzata z Nedoluzhko (2010, s. 122), další dva příklady jsou vlastní.

v případě mnohých abstrakt a deverbativ v úvahu také doplnění časovými či místními charakteristikami, což situaci dále komplikuje.

V rámci zachování principu důslednosti formuluje Nedoluzhko (2010, s. 122) následující pravidlo: „Pokud daná NP má jasně specifickou referenci, přiřadíme jí typ 0. Pokud však rozhodování o typu reference dané jmenné fráze není jednoznačné a vyžaduje aspoň trochu přemýšlení, označíme ji jako `coref_text`, `typ=NR`.“ Proto jsou jako koreference typu NR označeny např. vztahy v koreferenčním řetězci tvořené výrazy *nezaměstnanost* ve spojeních *míra nezaměstnanosti*, *růst nezaměstnanosti v letech 1991–1993*, *zvýšení míry nezaměstnanosti z 3,5 % (...) na 5–6 %* (Nedoluzhko, 2010, s. 124). Zůstává však otázkou, zda jde v podobných případech skutečně o koreferenci.

Z uvedených skutečností je zřejmé, že problematika koreference v případě abstrakt a dějových deverbativ by měla být předmětem dalšího výzkumu. Ten však již přesahuje možnosti této práce.

3. Problémy spojené s rozdělením typů reference v PDT

Rozdělení typů reference na typ 0 a NR se nejeví jako zcela ideální, jako problematický se v této souvislosti ukazuje především příliš široce vymezený typ NR. Tento typ reference totiž mají v zásadě všechny výrazy, které nemají referenci specifickou, jež je, jak bylo uvedeno výše,³¹ vymezena jako „odkazování ke konkrétnímu existujícímu, reálnému referentu a objektu skutečnosti“ (Nedoluzhko, 2010, s. 109). Referenci typu NR tedy mají všechny výrazy, které libovolnou z uvedených charakteristik postrádají.

Jako nejzávažnější se mi jeví tři typy problematických případů, které nyní ukáži na jednotlivých modelových příkladech.

3.1 Nespecifická reference s typem 0

Jak již bylo uvedeno,³² v některých případech se v PDT připisuje reference typu 0 i takovému anaforu, jehož antecedent referuje na nevybraný objekt, a to v případech, kdy anafor obsahuje určitý identifikátor³³. Tak např. ve větách *Chtěla by mít dceru. Dala by jí jméno Natálie*. V první větě výraz *dceru* nereferuje k vybranému reálnému referentu, a proto má referenci typu NR, koreferenční výraz *jí* má již ale referenci typu 0.

Toto řešení v podobných případech považuji za problematické ze tří důvodů. První z nich představuje skutečnost, že reference typu 0 předpokládá referenta reálně existujícího v aktuálním světě, což však v těchto případech zjevně není naplněno.

Druhý důvod spočívá ve vázanosti typu reference na přítomnost určitého identifikátoru, tedy jistého formálního prostředku. Lze si však představit případy, kdy je užití takového prostředku fakultativní a nemění referenční strukturu textu. Např. na větu *Chtěli bychom s touto firmou navázat spolupráci* lze dle mého názoru navázat jak větou *Tato spolupráce by trvala jeden rok*, tak větou *Spolupráce by trvala jeden rok*, a dokonce i větou *Trvala by jeden rok*, přičemž asi není oprávněné předpokládat, že je reference podmětových výrazů v těchto větách odlišná.

Třetí důvod je spojen spíše s technickou stránkou anotace koreference v PDT. Z technických důvodů je totiž typ reference připisován nikoli jednotlivým výrazům, nýbrž

³¹ Viz s. 9.

³² Viz s. 9–10.

³³ Pojem „určitý identifikátor“ však Nedoluzhko (2010) nijak blíže nevymezuje, takže není zcela jasné, co přesně do této kategorie patří.

jednotlivým koreferenčním vztahům, přičemž se vychází z typu reference anaforu (a nepřihlíží se tedy k antecedentu). To znamená, že v uvedených případech nikde nebude uvedena informace o referenci typu NR u prvního z výrazů, a podobné vztahy se tak na základě anotace jeví jako *běžné* případy koreference s referencí typu 0, tedy jako případy reference s existujícím reálným antecedentem.

3.2 Reference k neohrazeným množinám prvků

Druhý typ problematických případů je spojen s výrazy, které referují k množinám prvků, jež neobsahují všechny exempláře příslušného druhu, avšak jejichž hranice nejsou zřetelně stanoveny. K takové množině referuje např. výraz *oči* ve větě *Obratlovci mají oči*. Oči, o kterých se v této větě mluví, nelze přesně určit žádným výčtem, zároveň však neplatí, že by se jednalo o celý druh „očí“, neboť jistě existuje nezanedbatelné množství očí, které nepatří obratlovcům. V PDT je reference takovýchto výrazů označována typem NR, neboť reference těchto výrazů samozřejmě není specifická. V zásadě se zde tedy předpokládá reference generická, což ale naráží na zmíněný problém, že totiž takové výrazy nereferují k celému druhu.

Situace se navíc zásadně komplikuje v případě, že je součástí daného textu kromě uvedené věty např. také věta *Hmyz má také oči*. Jelikož je reference obou výrazů *oči* chápána jako generická, přičemž oba výrazy referují k druhu „očí“ (a tedy referují k téže entitě), je v těchto případech nutné chápat oba výrazy *oči* jako vzájemně koreferenční. Samotný fakt, že oba koreferenční výrazy referují ke zcela jiným prvkům, by sám o sobě nemusel být na závadu (neboť se jedná o referenci k témuž druhu), problém ale nastává, když je druhá uvedená věta následována větou typu *Tyto oči však vypadají zcela jinak než oči obratlovců*. Zde se totiž pojednává o obou množinách očí samostatně, a jelikož koreference je vztahem symetrickým a tranzitivním, je zřejmé, že pokud jsou výrazy *oči* v prvních dvou větách koreferenční a pokud oba příslušné výrazy ve třetí větě jsou koreferenční se svými antecedenty ve větách předchozích, musí být nutně navzájem koreferenční i oba výrazy ve třetí větě, tedy výrazy *tyto oči* a *oči obratlovců*. Výsledná interpretace třetí věty pak zní „*x* vypadá zcela jinak než *x*“, což je ovšem interpretace nenáležitá. Rozdělení koreference na pouze dva typy, 0 a NR, neumožňuje podobné případy důsledně řešit.

3.3 Referent vymezený podmínkovou větou

Třetí okruh problémů je spojen s větami typu *Když je stůl velký, vejde se na něj spousta věcí*, konkrétně s referencí výrazů *stůl* a *něj*. Jedná se o případy, k nimž se Nedoluzhko (2010) vůbec nevyjadřuje. Částečně se podobají případům popsaným v části 3.1, jsou však spojeny ještě s dalšími problémy.

V případě výrazu *stůl* neumožňuje omezený repertoár typů reference připsat tomuto výrazu jiný typ než typ NR, neboť jistě se zde nejedná o specifickou referenci. Intuitivně se zdá, že mezi slovy *stůl* a *něj* existuje vztah koreference. Vzhledem k tomu, že je výraz *stůl* v PDT považován za generický, je ale otázka, zda je tu tento vztah vůbec na místě, neboť zájmeno *něj* referuje nikoli ke všem exemplářům druhu stůl, nýbrž pouze k těm stolům, které jsou velké.

Kromě toho je navíc možné si představit pokračování textu větou *Když je ale stůl malý, nevejde se na něj skoro nic*. Zde by bylo v rámci zachování zásad PDT nutné chápat opět výraz *stůl* jako generický, a tudíž koreferenční s výrazem *stůl* z první věty. A pokud bychom v obou větách chápali výrazy *stůl* jako koreferenční s příslušnými zájmeny, dospěli bychom k nutnosti vnímat i obě zájmena *něj* jako vzájemně koreferenční, což opět zásadně odporuje intuici (neboť to předpokládá totožnost množiny malých stolů a množiny velkých stolů). Ani tento problém tedy není možné pomocí typů 0 a NR adekvátně vyřešit.

Z těchto příkladů vyplývá, že rozdělení reference na typy 0 a NR není pro účely zachycování koreferenčních vztahů nejvhodnější. Při hledání adekvátního řešení nejprve v následujících dvou kapitolách představím některé české a zahraniční alternativní přístupy k referenci a generičnosti, následně se pokusím navrhnout vlastní model, který by lépe zohledňoval situaci případů výše popsaných.

4. Vybrané české přístupy k referenci

4.1 Palek, 1988

Bohumil Palek se v publikaci *Referenční výstavba textu* (Palek, 1988) zabývá textem z hlediska reference (denotace) jmenných frází, které se v textu vyskytují, a na ní založených vztahů mezi těmito jmennými frázemi.

Autor především rozděluje výrazy na tzv. *tp-výrazy* (výraz jako typ) a *tk-výrazy* (token výraz). Tp-výraz je výraz chápaný neaktuálně, bez zapojení do kontextu, zatímco tk-výraz je výraz jakožto exemplář, jedná se tedy o aktuální použití tp-výrazu v konkrétním textu (Palek, 1988, s. 15–16).

Základními pojmy, které Palek zavádí, jsou pojmy *extenze*, *reference* a *denotace*. Extenzí se rozumí „vztah mezi výrazem jako typem (tp-výrazem) a třídou objektů (event. i jednoprvkovou nebo prázdnou), jejíž prvky se mohou vyskytovat ve všech možných světech“ (Palek, 1988, s. 16). Taková třída objektů se nazývá *extendem* tp-výrazu. U apelativ jde zpravidla o víceprvkovou množinu, u proprií často o množinu jednoprvkovou.

Reference je vztah mezi tk-výrazem a objektem, který tento výraz označuje, přičemž takový objekt (označovaný jako referendum) je součástí aktuálního světa. Reference i extenze jsou tedy relacemi dvoučlennými, atranzitivními a asymetrickými (op. cit., s. 15–16).

Kromě těchto dvou vztahů zavádí Palek též vztah denotace jako vztah mezi tk-výrazem a denotátem, přičemž denotát je chápán jako „stín“ objektu“ (op. cit., s. 18). Jde tedy o relaci podobnou vztahu reference, ovšem s tím rozdílem, že denotát nemusí být součástí aktuálního světa. Jak doslova autor uvádí (op. cit., s. 19), „zatímco referendum umožňuje vztáhnout jazyk vůči světu (...), denotát naopak umožňuje ‚vidět‘ svět očima jazyka“.³⁴ Pojem denotátu je užitečný zvláště v případě takových výrazů, které nereferejí k objektu aktuálního světa (a tedy mají prázdné referendum), neboť i takové výrazy mohou například vyjadřovat gramatickou kategorii čísla. Děje se tak například v případě negace, tak ve dvojici vět *Nebyla zde žádná žena / Nebyly zde žádné ženy* mají obě jmenné fráze prázdné referendum, ovšem nikoli prázdný denotát – ten je instaurován gramatickou kategorií čísla, a je tedy v těchto dvou větách různý. Speciálním případem výrazů s prázdným referendem jsou výrazy, které mají prázdné extendum, jako např. *jednorožec*, *anděl* nebo *víla* (op. cit., s. 17).

³⁴ Denotát tedy není obrazem objektu v mysli člověka, nýbrž spíše jazykovým zpracováním objektu.

Jmenná fráze se podle Palka (op. cit., s. 29) skládá ze dvou částí, *komponentu nominálního (kvalitativního)* a *komponentu instauračního (instaurátoru)*. Funkcí nominálního komponentu (nejčastěji vyjadřovaného apelativy) je vyjadřovat kvalitu denotátu, zatímco funkcí instaurátoru je vyjadřovat jeho „kviditu“, tedy vyjadřovat, „že jmenná fráze označuje nějaký denotát“ (op. cit., s. 29). Instaurátory mohou být vyjadřovány zvláštním slovem (např. některými zájmeny) nebo i jednotlivými morfémy, např. morfémy vyjadřující gramatické číslo.

Denotáty jsou v popisovaném pojetí dvojího typu: *denotáty-konstanty* a *denotáty-proměnné*. Jako denotáty-konstanty mohou figurovat jevy či situace časoprostorově určené, zatímco denotáty-proměnné jsou „pouze součástí intence mluvčího“ (op. cit., s. 78) a vystupují ve významu „nějaký jeden předmět“, nebo ve významu „kterýkoli“ (op. cit., s. 78).

Centrem pozornosti autora jsou tzv. *denotační vztahy* v textu, které jsou založeny na vzájemném (množinovém) poměru mezi denotáty jednotlivých jmenných frází, ovšem uplatňuje se při nich tzv. *denotační tolerance*, na základě které „vztahy k ostatním výrazům v daném textu jsou pouze takové, jak je stanoveno užitými instaurátory“ (op. cit., s. 177).³⁵ Základními denotačními vztahy jsou vztahy identity, inkluze, členství (v množině) a rozdílu (tj. různosti denotátů).³⁶

Palek pracuje též s pojmy *anafora* a *katafora*. Jde o takové denotační vztahy, u kterých jedna z dvojice jmenných frází obsahuje identifikátor. Pokud obsahuje identifikátor jmenná fráze umístěná v textu více napravo, jde o anaforu, v opačném případě jde o kataforu. V tomto pojetí je tedy anafora (a katafora) podmíněna existencí jednoho ze zmíněných denotačních vztahů. Do denotačních vztahů však mohou vstupovat jen takové dvojice jmenných frází, které se neliší co do typu denotátu, není tedy možné, aby byly ve vzájemném denotačním vztahu výraz denotující konstantu a výraz denotující proměnnou. Proto např. mezi zájmenem *ho* a výrazem *revolver nebo brovnink* ve větě „*revolver nebo brovnink, to je zbraň, na kterou doplatí každej, my jako kluci jsme si ho vypůjčili...*“ (op. cit., s. 80) není vztah anafory, neboť antecedentem je výraz denotující proměnnou, zatímco zájmeno *ho* denotuje konstantu.

³⁵ Jde především o dva typy tzv. závislých instaurátorů: identifikátory, jejichž funkcí je denotovat již zmíněný denotát (např. osobní zájmena), a alterátory, které denotují jiný denotát než ten, který byl zmíněn. Pokud tedy žádná jmenná fráze z dvojice jmenných frází neobsahuje identifikátor nebo alterátor, neexistuje v rámci této dvojice denotační vztah.

4.2 Hlavsa, 1975

Vztahem výrazu ke skutečnosti tímto výrazem označované se podrobně zabývá Zdeněk Hlavsa v publikaci *Denotace objektu a její prostředky v současné češtině* (Hlavsa, 1975).

Výchozími pojmy jeho analýzy jsou pojmy *designace* a *denotace*. Designací se rozumí vztah, „v němž lexikálně systémová jednotka (substantivní znak) je vztahována (skrže svůj každý význam) k celým univerzům objektů vyhovujících jejím významům“ (Hlavsa, 1975, s. 11). Jde tedy o vztah, jehož se účastní výrazy, jež nejsou aktuálně užití.³⁷ Naopak denotace je vztah, „v němž je zmíněná jednotka vztažena k objektu, o kterém mluvčí vypovídá“ (op. cit., s. 11). Denotátem výrazu je tak vždy podmnožina (nikoli nutně vlastní) jeho designátu, vymezení této podmnožiny je zajišťováno prostřednictvím operátoru denotace (op. cit., s. 13). Základní formou denotátoru je denotační fráze, přičemž ta její část, která realizuje operátor denotace, se nazývá aplikátor.^{38,39}

Důležitým poznatkem je, že ne všechna substantiva plní v textu denotační funkci. Nedenotačním užitím substantiva „je užití kvalifikující, objevující se především ve jmenném predikátu se sponou“ (op. cit., s. 15). Jde o věty typu *Jan je mechanik*, kde výraz *mechanik* není denotátorem.

Hlavsa dále (op. cit., s. 18) provádí klasifikaci denotátorů podle typu (tzv. determinace) denotace. Denotátory dělí na dvě základní skupiny: denotátory jedinečné a nejedinečné, přičemž rozdíl mezi nimi je založen na protikladu určitosti. Jedinečné denotátory mají rys určitosti a operátor jedinečnosti (Un). Denotátory nejedinečné (s rysem [– určitost]) dále dělí na denotátory s operátorem existence (Ex) a denotátory s operátorem proměnnosti (Var).⁴⁰

Při jedinečné denotaci lze denotát přímo identifikovat. Podle způsobu identifikace denotátu lze rozlišit několik typů jedinečnosti. Jedinečnost ostensivní je zabezpečena exoforicky, např. prostřednictvím aplikátorů *tenhle*, *tamtén* apod. Jedinečnost kontextová je umožněna vnitrotextovým odkazováním, např. za použití aplikátoru *ten*. U některých výrazů je jedinečnost zabezpečena tím, že má daný výraz pouze jediný designát, je tomu tak zejména

³⁶ Vztah identity je tedy blízký vztahu koreference v pojetí PDT. Rozdíl spočívá v tom, že (jak bylo zmíněno) v Palkově přístupu musí být tento vztah stanoven příslušným instaurátorem.

³⁷ Vztah designace tedy v zásadě odpovídá Palkovu pojmu extenze. Palkovu extendu zde odpovídá pojem designát.

³⁸ Jak autor upozorňuje, odpovídá tento pojem přibližně Palkovu pojmu instaurátor.

³⁹ Hlavsa připomíná v souvislosti s denotací Palkův pojem reference, avšak jeho zavedení v protikladu k denotaci neshledává nutným, neboť mluvčí „neusiluje o to, aby každá jeho výpověď byla úplným obrazem reality, nýbrž jen aby splnila jeho komunikační záměr“ (op. cit., s. 17), a proto není potřeba odlišovat objekt od „stínu objektu“.

v případě některých proprií (např. *Asie*), ale i apelativ (např. *gotika*). V některých případech vyplývá jedinečnost denotace ze „vztahu k nějakému [jedinečnému] předmětu“ (op. cit., s. 54), přičemž jedinečnost tohoto předmětu je zabezpečena některým z dříve uvedených způsobů (např. *současný prezident ČR*).

Zvláštním případem jedinečnosti je tzv. kontextová jedinečnost, která se uplatňuje např. u jmenné fráze *tu knihu* ve dvojici vět *Půjč mu nějaký pěkný román. Vezme si tu knihu na dovolenou*. Tato jmenná fráze je jedinečná (neboť má rys určitosti), přestože denotát v tomto případě nelze jednoznačně identifikovat. Jak Hlavsa uvádí, sám se dříve přikláněl k názoru, že v podobných případech o jedinečnou denotaci nejde, avšak vzhledem k tomu, že by se tak „funkční pojetí determinace příliš vzdálilo formálně syntaktickým ukazatelům“ (op. cit., s. 57), preferuje chápání těchto denotátorů jako denotátorů jedinečných.⁴¹

Determinace denotace prostřednictvím operátoru Ex se uplatňuje např. u jmenné fráze *jakási kořenářka* ve větě *To léčení mu poradila jakási kořenářka* (op. cit., s. 58). Její podstatou tedy je, že denotátor má rys [–určitost], avšak jeho identifikace je možná (v citované větě je možné kořenářku pojmenovat, ovšem pokud zjistíme její jméno). V rámci tohoto typu je ještě možné rozlišovat výrazy podle významového rysu známosti, který „se týká především mluvčího“ (op. cit., s. 58). Kladná hodnota tohoto rysu se uplatňuje např. ve větě *Chtěl by si někdo dopisovat s jedním Angličanem?*, záporná hodnota např. u jmenné fráze *jakási kořenářka* v dříve citované větě ve větě.

Jako sekundární reprezentaci operátoru Ex navrhuje Hlavsa chápat také tázací slovo v doplňovacích otázkách, neboť tyto otázky „předpokládají, že můžeme k nim utvořit denotačně ekvivalentní větu oznamovací, kde na místě tázacího slova bude existenční denotátor“ (op. cit., s. 61). Např. věta *Kdo mi půjčil hodinky?* obsahuje předpoklad, že někdo mi hodinky skutečně půjčil a že tohoto člověka je možné identifikovat. Výjimkou jsou ovšem případy, kdy jde o „zakuklenou“ otázku zjišťovací“ (op. cit., s. 62), jako např. v otázce *Kdo mi přinesl ty hodinky?*, pokud je tato otázka vyslovena ve významu „přinesl mi někdo z vás ty hodinky?“ (v takovém případě je možné na tuto otázku odpovědět výrazem *nikdo*). Druhou výjimkou jsou otázky týkající se situace dosud nerealizované, např. *Kdo mi chce půjčit hodinky?*

Denotace s operátorem Var se vyznačuje tím, že denotát není možné identifikovat, neboť „volba označeného objektu není omezena jinak než oborem proměnnosti (...) daného

⁴⁰ Operátory Un, Ex, Var Hlavsa zavádí nejen pro jmenné fráze, ale též pro jednotlivé věty, podle toho, jaké situace pojmenovávají.

⁴¹ Hlavsa se tak v tomto bodě shoduje s pojetím koreference v PDT, kde jsou podobné případy chápány jako specifická reference.

designátoru“ (op. cit., s. 18). Tento typ denotace se podle Hlavsy uplatňuje buď v případech, kdy má také celá věta operátor Var (např. *kteřoukoli tramvaji ve větě Tatínek jezdí ktereoukoli tramvaji*), nebo ve větách, které „mají jiný modalitní význam než reálnost“ (op. cit., s. 66; např. *Může jet ktereoukoli tramvaji*).⁴²

V některých případech může docházet ke dvojznačnosti co do determinace denotace. Může se tak dít v případech, kdy je součástí jmenné fráze aplikátor *nějaký*, který je možné použít jak při determinaci typu Ex, tak při typu Var. Jako příklad dvojznačné věty uvádí autor (op. cit., s. 59) větu *Chce si koupit nějakou knihu*, kde není zřejmé, zda jde o konkrétní (ale neurčenou) knihu, nebo zda jde o knihu libovolnou.

Pokud jde o generické jmenné fráze, nejsou v rámci uvedené klasifikace chápány jako zvláštní skupina, nýbrž patří mezi denotátory s operátorem Var. Hlavsa zde přímo o generičnosti nemluví, uvádí však (op. cit., s. 68), že u některých případů lze „dost oprávněně pochybovat o tom, že patří k probraným typům determinace“. Má tím na mysli jmenné fráze ve větách *Studuje staré kroniky* nebo *Rád loví ryby*, které se vztahují „k determinačně (i kvantifikačně) nevymezené množině prvků“ (op. cit., s. 68). Autor v této souvislosti zmiňuje možnost chápat tyto jmenné fráze jako nedenotační.

Při procesu denotace (tedy vymezení denotátu jako podmnožiny designátu) se však kromě determinace uplatňuje ještě operace kvantifikace, která „z oboru proměnnosti designátoru nebo z nějaké předem vymezené množiny vyčleňuje množství entit, ke kterým má výpověď denotační vztah“ (op. cit., s. 72). Kvantifikace může být určitá nebo neurčitá a parciální nebo totalitní. Operace kvantifikace je paralelní operací k operaci determinace, přičemž obě operace na sobě nejsou závislé. Kvantifikován tak může být denotátor s operátorem Un (např. *tyto dvě vázy*), s operátorem Ex (např. *nějaké dvě vázy* ve větě *Byly tam nějaké dvě vázy*) i s operátorem Var (např. *mnohé dvě vázy* ve větě *Vázy se brousí jako originály, ale přesto jsou si mnohé dvě vázy k nerozeznání podobné*⁴³).

Zvláštní pozornost je věnována problematice distributivního přiřazování množin, tedy jevu, kdy „jisté množině o *k* členech přiřazujeme určitým způsobem jinou množinu s týmž, nebo jiným počtem prvků“ (op. cit., s. 94). Autor zde rozebírá jednotlivé způsoby tohoto přiřazení, které jsou závislé na různých faktorech, zejm. lexikálním obsazení (op. cit., s. 96). Z hlediska determinace denotace je podstatné, že autor navrhuje chápat denotátory

⁴² Vzhledem k tomu, že do kategorie Var patří mimojiné i jmenné fráze, které se např. v (GenBook, 1995) pokládají generické, je zřejmé, že se denotátory s operátorem Var mohou vyskytovat i v jiných kontextech, tedy ve větách, které referují k jedinečné reálné události. Např. výraz „brambory“ ve větě *Brambory byly do Evropy dovezeny v 16. století*.

pro distribuovanou množinu jako denotátory s operátorem Var, a to z důvodů hlavně formálně jazykových (op. cit., s. 98). Např. ve větě *Každé dítě splnilo nějaký úkol* má jmenná fráze *nějaký úkol* operátor Var, neboť v případě každého dítěte se jedná o úkol jiný. Takové řešení však nepokládám zcela neproblematické, neboť veškeré splněné úkoly je možné přesně identifikovat.

4.3 Adamec, 1980

Přemysl Adamec v článku *K vyjadřování referenční určenosti v češtině a v ruštině* (Adamec, 1980) konstruuje podrobnou klasifikaci jmenných frází z hlediska typu jejich reference a následně porovnává možnosti realizace těchto typů v češtině a v ruštině.

V úvodu autor představuje tři kritéria, která se v souvislosti s typem reference zpravidla uvádějí. Prvním je kritérium generičnosti, podle kterého se jmenné fráze dělí na jmenné fráze generické (referující k celé třídě objektů) a jmenné fráze singulativní (referující „k jednotlivému konkrétnímu objektu, event. ke skupině jednotlivých konkrétních objektů“ /Adamec, 1980, s. 257/). Kritérium určenosti umožňuje rozlišit jmenné fráze na určené a neurčené podle „sepjatosti nebo nesejpjatosti v povědomí účastníků komunikativního aktu s jedním známým objektem, event. s jednou skupinou známých objektů“ (op. cit., s. 257). Třetí kritérium danosti / novosti spočívá v tom, zda je příslušná jmenná fráze zapojená, či nezapojená do předcházejícího kontextu.

Na základě těchto i některých dalších kritérií pak provádí Adamec klasifikaci typů reference. Podle kritéria generičnosti dělí jmenné fráze na *generické* a *singulativní*. Kromě těchto dvou typů však zmiňuje ještě třetí typ, tzv. referenci *podmíněně singulativní*, která „tvoří jakýsi přechod mezi oběma těmito typy“ (op. cit., s. 257). Jde o případy typu *Půjč mi nějakou knihu* či *Poprosím o to některého kolegu*, kde příslušné jmenné fráze mají „naději stát se singulativními teprve při realizaci obsahu predikátu“ (op. cit., s. 257–258).

Pokud jde o referenci singulativní, lze ji dále rozdělit podle kritéria určenosti na referenci neurčenou a určenou.⁴⁴ Pokud jde o referenci neurčenou, lze příslušné jmenné fráze rozdělit na ty, u nichž jde o neidentifikovanost objektu pouze z hlediska posluchače (např. *Mluvil jsem o tom s jedním lékařem*), nebo z hlediska mluvčího i posluchače (např. *Na chodbě*

⁴³ Podobné případy jsou v PDT chápány jako generické, což je vzhledem k přítomnosti kvantifikace předpoklad poněkud problematický.

⁴⁴ Jak uvádí autor, v případě reference podmíněně singulativní jde vždy o referenci neurčenou. V případě reference generické autor k další klasifikaci nepřistupuje, i když i tam by byla podle některých autorů možná (srov. zde část 5.2).

jsem potkal jakéhosi pána).⁴⁵ V případě reference určené může jít buď o určenost „zvnějšku“ (na základě endoforického či exoforického odkazování), nebo o určenost „zevnitř“ (založenou na jednoznačnosti pojmenování).

Celou klasifikaci je pak možné znázornit takto:⁴⁶

- Reference:
1. generická
 2. podmíněně singulativní (neurčená)
 3. singulativní
 - a. neurčená
 - i. pro oba partnery
 - ii. jen pro posluchače
 - b. určená
 - i. zvnějšku
 - ii. zevnitř

Zvláštní pozornost je však třeba věnovat větám, „v nichž substantiva stojící v rematické části mají jakoby pouze signifikativní funkci a jejich konkrétně předmětový charakter, a tedy i ta či ona referenční platnost jsou jakoby zatlačovány do pozadí“ (Adamec, 1980, s. 259). Jde o věty typu *Dědeček čte noviny* nebo *Ten dopis psalo dítě*. V těchto případech je z komunikativního hlediska reference jmenných frází *noviny* a *dítě* skutečně zatlačena do pozadí, z hlediska logického však zde denotátová vztaženost „přece jen existuje a je de facto singulativní“ (op. cit., s. 259). V některých případech může být i podmíněně singulativní (např. *Pojedeme tam autobusem*).

Tyto případy je možné rozdělit na dva typy z hlediska aktuálního členění. Buď je v takových větách v rématu příslušná jmenná fráze samotná (např. *Ten dopis psalo dítě*), nebo je kromě jmenné fráze v rématu též predikát (např. *Dědeček čte noviny*, *Jirku štípla vosa*). Jak autor uvádí, jsou v prvním případě jmenné fráze blízké predikátovému charakteru, zatímco ve druhém případě tvoří tyto jmenné fráze spolu se slovesem „jakési souslovné pojmenování činnosti“ (op. cit., s. 260).

Od vyjadřování určenosti je podle autora třeba důsledně oddělit dva případy, které jsou vyjadřovány podobným způsobem. Prvním z nich je „dovolávání se společné zkušenosti mluvčího i adresáta“ (op. cit., s. 262), vyjadřované pomocí zájmena *ten*, které se však v tomto případě „nevztahuje pouze ke svému substantivu, nýbrž k celému obsahu věty“ (op. cit., s. 262). V těchto případech může mít příslušná jmenná fráze jakoukoli referenci⁴⁷, tj. jak

⁴⁵ Podobné rozdělení v podstatě navrhuje i Hlavsa (1975, s. 58) včetně distribuce aplikátorů „jeden“ a „jakýsi“.

⁴⁶ Schéma je převzato z Adamcova článku (1980, s. 259).

⁴⁷ Z této skutečnosti je zřejmé, že funkcí zájmena *ten* zde není vyjadřování určenosti.

singulativní (např. *Do toho Brna jsme nakonec nejeli*), tak podmíněně singulativní (např. *Tak třeba pojďme do toho kina*) a generickou (*Tak jsem si ověřil, že ti pavouci skutečně mívají těch osm očí*).

Druhým případem, který přímo nesouvisí s kategorií určenosti, je použití zájmena *ten* ve funkci vyjadřování expresivnosti. Takto se zájmeno *ten* používá především ve spojení se singulativními jmennými frázemi určenými (např. *To sluníčko dnes ale pálí!*) a se jmennými frázemi generickými (např. *Pořád jen ten big-beat!*).

4.4 Shrnutí popsaných koncepcí

V publikaci Palkově a Hlavsově jsou podrobně rozebírány jednotlivé typy jmenných frází a vztahů mezi nimi a tím, co tyto jmenné fráze označují. Obě práce se navíc zabývají strukturou jmenných frází. Následující tabulka znázorňuje vztahy mezi termíny, které oba autoři užívají:

	Palek	Hlavsa
výraz aktuálně neužitý v textu	tp-výraz	designátor
výraz aktuálně užitý v textu	tk-výraz	denotátor ⁴⁸
vztah neaktuálně užitého výrazu k označovanému	extenze	designace
označované aktuálně neužitého výrazu	extendum	designát
vztah aktuálně užitého výrazu k označovanému „stínu objektu“	denotace	denotace ⁴⁹
„stín objektu“ označovaný aktuálním výrazem	denotát	denotát
vztah aktuálně užitého výrazu k označovanému objektu reálného světa	reference	–
objekt reálného světa označovaný aktuálním výrazem	referendum	–
část jmenné fráze zajišťující referenci / denotaci	instaurátor	aplikátor

Ve všech třech citovaných pracích jsou předloženy více či méně podrobné klasifikace typů vztahů mezi výrazem v textu a objektem, který tento výraz označuje. Vzájemné vztahy těchto typů v rámci všech popsaných přístupů zachycuje následující tabulka:⁵⁰

⁴⁸ Platí pouze pro případ, kdy daný výraz plní denotační funkci.

⁴⁹ Hlavsa ovšem termínu „stín objektu“ neuzívá.

⁵⁰ Z citovaných prací není zcela zřejmé, do jaké skupiny patří typ, který Hlavsa nazývá kontextovou jedinečností (příkladem je zmiňovaná dvojice vět *Půjč mu nějaký pěkný román. Vezme si tu knihu na dovolenou.* V Hlavsově pojetí má výraz *tu knihu* operátor Un, v PDT má v případě koreferenčního vztahu typ 0. V ostatních citovaných pracích není uvedený typ explicitně zmíněn. V Adamcově koncepci by byl pravděpodobně (na základě rysu určenosti) zařazen k referenci singulativní určené zvnějšku. Co se týče pojetí Palkova, mohl by být tento výraz chápán jako konstanta, pokud by byl kontext daný první větou uznán jako dostatečný základ

Palek	Hlavsa	Adamec	Příklad	PDT	
konstanta	Un	jedinečnost ostensivní	singulativní urč. „zvnějšku“	<i>Podej mi <u>tamty papíry</u>.</i>	typ 0
		jedin. kontextová		<i>Na stole leží tužka. Podej mi <u>ji</u>.</i>	
		jediný designát	singulativní urč. „zevnitř“	<i>Zítřa pojedeme do <u>Brna</u>.</i>	
		vztah k jedin. předm.		<i>Viděli jsme <u>prezidenta USA</u>.</i>	
	Ex	singulativní neurč. pro oba	<i>Potkal jsem <u>jakéhosi pána</u>.</i>		
		singulativní neurč. pro	<i>Mluvil jsem s <u>jedním lékařem</u>.</i>		
proměnná	Var	podmíněně singulativní		<i>Chci si koupit <u>auto</u>.</i>	typ NR
		generická		<i><u>Kolegové</u> si musejí pomáhat.</i>	

Kromě těchto klasifikací je velmi důležitý Adamcův poznatek, že u některých jmenných frází v rématu dochází k upozadění jejich referenční schopnosti, a typologie těchto jmenných frází. Důležité také je, že navzdory této skutečnosti je možné chápat tyto jmenné fráze jako referenční.

Poměrně zásadní je rovněž Hlavsovo vymezení operace kvantifikace jakožto paralelní operace k operaci determinace při procesu denotace objektu. Z hlediska reference a koreference zasluhuje tato operace pozornost především v případě uplatnění determinačního denotačního operátoru Var, neboť v takových případech je zřejmě nutné odlišovat takové jmenné fráze od jmenných frází generických.⁵¹

časoprostorového určení; v opačném případě by se na základě jeho povahy jakožto pouhé „součásti intence mluvčího“ mohlo jednat o proměnnou.

⁵¹ Není například jasné, do jakého typu by v Adamcově koncepci bylo lze zařadit jmennou frázi *mnohé dvě vázy* v již zmiňované větě *Vázy se brousí jako originály, ale přesto jsou si mnohé dvě vázy k nerozeznání podobné* (případně zda by v případě zařazení této jmenné fráze do skupiny jmenných frází generických bylo možné chápat tuto jmennou frázi jako koreferenční s výrazem *vázy* v první klauzi).

5. Pojetí generičnosti v The Generic Book

V této části nastíním pojetí generičnosti popsané v publikaci *The Generic Book* (GenBook, 1995). Toto pojetí je zpracováno především s ohledem na jazykový materiál angličtiny. Rozlišují se zde dva typy generičnosti, které je třeba důsledně oddělovat (i když spolu navzájem úzce souvisí): generičnost propozic a generičnost jmenných frází.

5.1 Generičnost propozic

Generičnost propozic chápou autoři publikace jako jisté zobecnění nad stavy, procesy či událostmi vyjadřovanými příslušnými propozicemi. Rozlišují se tak dva typy propozic: propozice epizodické (tj. s příznakem [- generičnost]) a propozice generické (charakterizační). Tento protiklad je vidět na příkladu dvojice vět *A dog is barking. / A dog barks.*⁵² Zatímco první věta je jazykovým zachycením jednoho konkrétního procesu psího štěkání, druhá věta vyjadřuje zobecnění nad všemi procesy štěkání libovolných psů.⁵³

Jako pomocné kritérium pro rozhodování, zda jsou, nebo nejsou konkrétní propozice generické, je navrhováno několik testů. Jeden z testů spočívá v ověření možnosti rozvíjet řídicí sloveso adverbii jako *usually* nebo *typically*. Pokud takové rozvíjení zásadně nemění význam propozice, je tato propozice pravděpodobně generická. (Příkladem může být dvojice vět *A lion (usually) has a bushy tail. / A lion (usually) stood in front of my tent.* První věta je generická, druhá nikoli.) Další test využívá skutečnosti, že význam predikátů v generických propozicích je chápán jako statický. Věty, jejichž predikát je ve tvaru, který statickou interpretaci vylučuje (v angličtině jde o slovesné tvary obsahující sufix *-ing*), tedy jistě generické nejsou. Generické přitom pravděpodobně nejsou ani takové věty, jejichž význam je možné pomocí vět obsahujících zmíněný typ predikátu alternativně vyjádřit.

Opozice generičnosti u propozic je v této publikaci součástí širší klasifikace vět. Věty jsou zde především rozděleny na statické a dynamické (aniž by byl tento protiklad nějak blíže vymezen), přičemž je konstatováno, že rys [+ generičnost] implikuje statičnost příslušné věty. V rámci vět generických se dále rozlišuje mezi větami habituálními („habitual sentences“) a lexikálně charakterizačními („lexical characterizing sentences“). Rozdíl spočívá v tom, že

⁵² Příkladové věty použité v této části přebírám z GenBook (1995).

⁵³ Opozice epizodičnost / generičnost tedy v podstatě odpovídá opozici aktuálnost / neaktuálnost, která je zmiňována např. v *Mluvnici češtiny 2* (Komárek et al., 1986, s. 165). V této publikaci ovšem tato opozice není přesně definována.

predikáty habituálních vět mohou být derivovány od predikátů vět epizodických, a habituální věty tak mohou mít svůj epizodický protějšek (např. *John smokes.* / *John is smoking*), zatímco lexikálně charakterizační věty epizodický protějšek nemají v důsledku sémantického významu svých predikátů (jako příklad se zmiňují predikáty typu *know*, *have*⁵⁴ atp.).

V dalším výkladu je proveden podrobnější rozbor sémantiky generických propozic. Jak již bylo zmíněno, podstatou významu těchto propozic je zobecnění (generalizace). Toto zobecnění se může realizovat buď nad objekty⁵⁵, nebo nad situacemi, nebo nad obojím. Zobecnění nad objekty se uplatňuje např. na větě *A lion has a bushy tail*, kde se vlastnost vyjádřená predikátem přisuzuje celé množině lvů, zobecňuje se tedy nad reprezentacemi druhu „lev“. K zobecnění nad situacemi dochází např. ve větách *Mary smokes* a *Mary smokes when she comes home*. V prvním případě jde o zobecnění nad situacemi, ve kterých se obvykle kouří⁵⁶, ve druhém případě o zobecnění nad situacemi charakterizovanými vedlejší větou.

Význam generických propozic navrhuji autoři zapisovat pomocí generalizačního operátoru **GEN**. Tento operátor má dva argumenty: *restriktor* („restrictor“), který vyjadřuje předmět zobecnění, a *matici* („matrix“), která vyjadřuje predikaci nad předmětem zobecnění. Významový zápis první příkladové věty předchozího odstavce tak bude vypadat takto: **GEN[x;y](x is a lion; y is a bushy tail, x has y)**⁵⁷. Význam druhé věty lze zapsat ve tvaru **GEN[s,x;](x = Mary & s is a normal situation with respect to smoking & s contains x; x smokes in s)**.⁵⁸

Operátor **GEN** tedy zajišťuje generalizaci nad objekty nebo situacemi. Problémem, který se nepodařilo uspokojivě vyřešit, je otázka kvantifikace těchto generalizovaných objektů či situací. Generalizaci totiž rozhodně nelze zaměňovat s univerzální kvantifikací, neboť např. větu *A cat has a tail* považujeme za pravdivou i přes to, že velmi pravděpodobně na světě existují i takové kočky, které z nějakého důvodu o ocas přišly. V publikaci *GenBook* (1995,

⁵⁴ Sloveso *have* patří do této kategorie, pokud je použito v posesivním významu. V jiných případech může toto sloveso tvořit i predikát epizodických propozic (např. predikát *have an affair*).

⁵⁵ Zde nikoli jen ve smyslu gramatických objektů.

⁵⁶ Situace, nad kterými se zobecňuje, zde tedy není explicitně pojmenována, nýbrž se vyrozumívá z kontextu na základě pragmatických zkušeností se světem.

⁵⁷ Část předcházející středníku je restriktor, část za středníkem je matice.

⁵⁸ Zápis pomocí tohoto operátoru je výhodný v tom, že umožňuje zachytit rozdíly v aktuálním členění vět. Např. výpověď *Typhoons arise in this part of the Pacific* může být realizací dvou vět lišících se v aktuálním členění (v jednom případě je základem věty výraz *typhoons*, ve druhém *in this part of the Pacific*). Významový zápis obou vět se bude lišit, neboť v restriktoru se vždy objeví základová část věty. Tento významový zápis tedy umožňuje rozlišit i známé dvojice vět *Česky se mluví na Moravě.* / *Na Moravě se mluví česky.* či *Málo knih čte mnoho lidí.* / *Mnoho lidí čte málo knih.*

s. 45–63) je zmíněno několik teorií generalizační kvantifikace, z nichž žádná se autorům nejeví jako zcela bezproblémová. Jako příklad zde některé z teorií uvedu.

Jednou z teorií je teorie *relevantní kvantifikace*. Na základě této teorie je generalizační operátor chápán jako kvantifikace nad relevantními objekty. Např. ve větě *Whales give birth to live young*. se generalizuje pouze nad samicemi velryb, nikoli nad samci. Kritéria pro stanovení relevance, které je třeba provést pro každou propozici zvlášť, však nejsou jasně daná. Podle této teorie by tak musela být za pravdivou generickou větu označena i věta *Whales are sick* s tím, že relevantními objekty jsou ty velryby, které jsou nemocné, což je zcela v rozporu s intuicí.

Další teorií je teorie *prototypů*, podle které je podstatou generalizační kvantifikace univerzální kvantifikace nad prototypickými objekty. Vychází se při tom z předpokladu, že pro každou třídu objektů lze nalézt množinu takových objektů, které lze označit za prototypy. Na problém však narazíme u dvojice vět *A duck has colorful feathers* a *A duck lays whitish eggs*. Zatímco v první větě jde o zobecnění pouze nad samci, ve druhé větě pouze nad samicemi. Prototypický případ kachny tedy není možné hledat v množině existujících kachen. Pokud by tato teorie měla být adekvátní, bylo by třeba prototypické případy „konstruovat“ (jedním z kroků této konstrukce by pak mohla být abstrakce od pohlaví jedince). Je však zřejmé, že problém generalizační kvantifikace by tím nebyl vyřešen, nýbrž pouze nahrazen problémem jiným.

5.2 Generičnost jmenných frází

Druhým typem generičnosti je generičnost jmenných frází, která konstituuje opozici jmenných frází *objektových* („object-referring NPs“, rys [– generičnost]) a *generických* („kind-referring NPs“). Rozdíl mezi nimi je založen na tom, že zatímco objektové jmenné fráze mají referenci k jednotlivým objektům, generické referují k druhům. Jako klíčová zde tedy vyvstává otázka podstaty druhu. Podle autorů musí jít o tzv. „well-established kind“ (op. cit., s. 11), tedy o takový pojem, který je všeobecně jako druh chápán. Např. anglické spojení *Coke bottle* je mluvčími jako druh chápáno, zatímco *green bottle* nikoliv. Jazykovým odrazem této skutečnosti je fakt, že generickou frází *the Coke bottle* konstruovat lze, zatímco analogická fráze *the green bottle* generický význam mít nemůže⁵⁹. Rovněž spojení jako *lions*

⁵⁹ Pochopitelně však může fungovat jako objektová fráze, tj. referovat ke konkrétnímu objektu s rysem určitosti.

that have toothaches se nepovažují za generické, neboť nelze analogicky tvořit jmenné fráze s generickým významem.

Rovněž pro rozlišení generických frází od objektových je zavedeno několik pomocných testů. Kromě testu na všeobecné chápání příslušného pojmu jako druhu jde o test s predikáty, které jako jeden ze svých argumentů obligatorně vyžadují jmennou frázi označující druh. Jde o predikáty jako *die out* či *be extinct*, které vyžadují generickou frázi v pozici gramatického subjektu, a predikáty typu *invent* nebo *exterminate*, které ji vyžadují v pozici gramatického objektu. Na základě tohoto testu se ukazuje, že ne všechny jmenné fráze, které figurují jako subjekty v generických propozicích, musí být nutně generické. Tak např. v generické větě *A dog barks* subjekt není vyjádřen generickou frází, protože fráze *a dog* nemůže fungovat ve spojení s uvedenými predikáty.⁶⁰

Kromě opozice generičnosti se v případě jmenných frází v publikaci pracuje též s opozicí *specifičnosti*. Specifické jmenné fráze jsou takové, které referují k jistým způsobem dané entitě („particular entity“), zatímco nespecifické k takto dané entitě nereferují. Nespecifické jmenné fráze mají v angličtině zpravidla podobu spojení neurčitého členu a substantiva, jako např. *a lion* ve větě *A lion has a bushy tail*. Jmenná fráze *a lion* je nespecifická, protože (jakožto objektová fráze) nereferuje k danému lvu, nýbrž k celé množině lvů, která je vymezena generalizačním operátorem **GEN**. Neplatí však, že by každá takto konstruovaná jmenná fráze byla nespecifická, např. ve větě *A lion stood in front of my tent* fráze *a lion* je specifická, neboť označuje jednoho vybraného lva (přestože má díky přítomnosti členu rys neurčitosti). Naopak ve větě *The lion has a bushy tail* je subjektová fráze specifická, protože (jakožto generická fráze) referuje ke konkrétnímu, vybranému druhu. Rysy [\pm generičnost] a [\pm specifičnost] na sobě tedy nejsou přímo závislé, přítomnost či nepřítomnost jednoho z těchto rysů neimplikují přítomnost či nepřítomnost druhého.

Autoři se v publikaci (op. cit., s. 64–77) podrobně věnují formě, jaké mohou nabývat generické jmenné fráze. Jak upozorňují, co do formy těchto jmenných frází se jednotlivé jazyky značně liší. Hlavním zkoumaným jazykem je zde angličtina, upozorňováno je však též na některé odlišnosti v případě několika jiných jazyků, např. francouzštiny, němčiny či čínštiny.⁶¹

⁶⁰ Zároveň pochopitelně rovněž platí, že ne všechny věty obsahující generickou jmennou frázi musí být generické. Příkladem může být věta *Potatoes were introduced into Ireland by the end of the 17th century*, která není generická, přestože jejím gramatickým subjektem je generická fráze *potatoes*.

⁶¹ Např. ve francouzštině nacházejí autoři značné odlišnosti v případě používání členu u generických nominálních frází. Jak upozorňují, ve francouzštině je u těchto frází obligatorně přítomen člen, a to i v případě substantiv v plurálu a látkových jmen, kde se v angličtině člen nepoužívá (např. *Les pandas sont éteint, L'or prend de la valeur*).

Při zkoumání formy generických jmenných frází se jeví jako užitečné rozlišovat běžnou (netaxonomickou) interpretaci jmenné fráze od taxonomické interpretace. Při běžné interpretaci může mít generická jmenná fráze v angličtině čtyři základní formy: spojení singulárové podoby substantiva s určitým členem (*The panda will become extinct soon.*), plurálová podoba substantiva bez členu (*Pandas will become extinct soon.*), singulare tantum (*Gold has the atomic number 79.*) a tzv. vlastní jména druhů (*Ailuropoda melanoleuca will become extinct soon*). Všem těmto formám lze kromě rysu [+generičnost] připsat též rys [+specifičnost].

Poněkud jiná je situace v případě taxonomické interpretace jmenných frází. Taxonomicky interpretovaná jmenná fráze referuje k taxonomické jednotce, která je v taxonomické hierarchii na nižší pozici, než by byl referent příslušné jmenné fráze při netaxonomické interpretaci. Např. substantivum *whale* při taxonomické interpretaci referuje ke druhu velryby. Je zřejmé, že každá jmenná fráze mající taxonomickou interpretaci je generická. Generické fráze s taxonomickou interpretací mohou tedy nabývat libovolných forem obvyklých pro jmenné fráze. Generickou frází tak může být i spojení neurčitého členu se substantivem, jako např. ve větě *A whale, namely the blue whale, is nearly extinct*. Taxonomická interpretace je také nutnou podmínkou k tomu, aby mohla být jmenná fráze nespécifickou a zároveň generickou (např. *A predatory cat (e. g., the leopard) is exterminated when it is dangerous to people*).

Pokud jde o význam generických nominálních frází, je jejich základní charakteristikou reference k druhu. Ta však může být interpretována různým způsobem a jednotlivé interpretace se od sebe mohou poměrně výrazným způsobem lišit. V publikaci (op. cit., s. 78–85) je popsáno následujících 8 různých interpretací reference k druhu:

1. *Vlastní reference k druhu*. V tomto případě jmenná fráze referuje pouze k druhu jako celku, nikoli k jednotlivým reprezentacím druhu. Tato interpretace se uplatňuje ve spojení s predikáty vyžadujícími pro příslušnou pozici obligatorní použití generické fráze, tedy těmi predikáty, které byly zmíněny jako jeden z prostředků testování generičnosti jmenných frází. Příkladem může být věta *The dodo is extinct*. Význam věty se pochopitelně týká i jednotlivých reprezentací (dozvídáme se z něj, že žádné reprezentace neexistují), nicméně příslušný predikát nemůže být použit ve spojení s objektovou jmennou frází.

2. *Souhrn reprezentací*. Predikát mající jako argument jmennou frází s touto interpretací vypovídá o souhrnu všech reprezentací příslušného druhu. Např. *Linguists have more than 8000 books in print*. Tuto interpretaci však není možno uplatnit ve všech případech

vypovídání o vlastnosti souhrnu reprezentací druhu, např. větu *Rabbits have a weight of more than one million tons* hodnotí autoři jako neobvyklou.

3. *Průměrná reprezentace*. Jmenná fráze referuje k průměrné reprezentaci druhu, jde o věty typu *The American family contains 2.3 children*. Formálně se věty uplatňující tuto interpretaci generických jmenných frází neliší od vět, v nichž se uplatňuje předchozí interpretace, rozhodnutí o tom, o kterou interpretaci jde, je závislé na znalostech světa (např. je zřejmé, že významem poslední věty není fakt, že by měly všechny americké rodiny 2,3 dětí). V některých případech však může docházet ke dvojznačnostem, např. *The American family used less water this year than the last year* (pravdivostní podmínky této věty se při změně interpretací mohou lišit, např. pokud by během roku došlo ke zvýšení počtu amerických rodin).

4. *Reprezentace druhu*. Při této interpretaci se uplatňuje reference k jednotlivým reprezentacím příslušného druhu, jde o věty typu *The potato contains vitamin C*.

5. *Jmenné fráze s predikací vyjadřující odlišující vlastnost*. Tato interpretace se uplatňuje, pokud predikát, jehož je generická fráze argumentem, vyjadřuje vlastnost, která odlišuje příslušný druh od jiných druhů. Příkladem je věta *Dutchmen are good sailors*. Takové věty se na první pohled podobají větám předchozího typu, rozdíl však spočívá v tom, že ve větách předchozího typu lze bez změny pravdivostních podmínek zaměnit generickou frází za nespecifickou frází objektovou, např. *A potato contains vitamin C*. V případě této interpretace podobná záměna není možná, neboť věta *A Dutchman is a good sailor* se v pravdivostních podmínkách liší: vyjadřuje, že vlastnost „být dobrý námořník“ platí pravděpodobně pro většinu náhodně vybraných Holanďanů, zatímco analogická věta s generickou frází znamená, že celý druh (tj. národ Holanďanů) je znám tím, že má dobré námořníky.

6. *Reference k reprezentativním objektům*. Při této interpretaci nominální fráze referuje pouze k omezenému množství reprezentací druhu, jako ve větě *In Alaska, we filmed the grizzly*. Jak autoři uvádí, jednotlivé jazyky se liší v tom, do jaké míry tuto interpretaci připouštějí, např. v němčině je tato míra větší než v angličtině.

7. *Průkopnická reprezentace*. Uplatňuje se např. ve větě *Man set foot on the Moon in 1969*. Podstatou této interpretace je, že vlastnost reprezentanta příslušného druhu, která je pro tento druh netypická, se připisuje celému druhu. Taková vlastnost však musí být chápána jako

relevantní pro celý druh, proto není možné interpretovat tímto způsobem např. větu **Man ate 128 pretzels in one hour in 1976*⁶².

8. *Porovnání reprezentantů*. Ve větách obsahujících nominální fráze s touto interpretací dochází k porovnání jednotlivých objektů náležejících ke stejnému druhu. Jde o věty typu *The wolves are getting bigger as we travel north*. Rovněž pro věty tohoto typu platí, že generickou frází nelze zaměnit za nespécifickou frází objektovou: věta *A wolf is getting bigger as we travel north* je sice gramaticky možná, ale její význam je odlišný a neodpovídá našim zkušenostem o vnějším světě.

5.3 Kombinace obou typů generičnosti

V předchozích částech jsem popsal dva typy generičnosti, které autoři publikace rozlišují. Jak však vyplývá z některých příkladových vět, mohou se někdy oba typy generičnosti kombinovat.

Tuto skutečnost je třeba zohlednit při konstrukci významových zápisů takových vět. Popsaný způsob zápisu významu generických vět je nutné určitým způsobem modifikovat, neboť generalizační operátor **GEN** provádí operaci kvantifikace (byť její přesné vymezení se autorům nepodařilo exaktně definovat), přičemž tato operace se může uplatňovat pouze na množině vyjádřené nespécifickou jmennou frází. Kvantifikovány jsou navíc ve většině případů jednotlivé reprezentace druhů (objekty), nikoli jednotlivé druhy.

Z těchto důvodů zavádějí autoři tzv. relaci **IS**. Jde o vztah mezi dvěma proměnnými, pro které platí, že buď jsou identické, nebo jsou ve vztahu reprezentace, tj. jedna proměnná představuje objekt, který je reprezentantem druhu představovaného druhou proměnnou. Formálně lze relaci **IS** definovat takto: $IS(x,y) \leftrightarrow (x=y \vee R(x,y))$.

Významový zápis věty *The lion roars when it smells food* s uplatněním relace **IS** vypadá takto: $GEN_{[x,s;]}(IS(x, \mathbf{Leo\ leo}) \& x \mathbf{smells\ food\ in\ s}; x \mathbf{roars\ in\ s})$ ⁶³. Tento zápis má vzhledem k definici relace **IS** dvě různé interpretace: při první interpretaci se za **IS(x, Leo leo)** dosadí relace **R(x, Leo leo)**, při druhé interpretaci se za ni dosadí vztah $x = \mathbf{Leo\ leo}$. Zmíněná věta je tedy dvojnásobná, přičemž pokud bychom v ní generickou frází nahradili nespécifickou frází objektovou *a lion*, takto získaná věta by vyjadřovala pouze význam první interpretace věty s generickou frází. Zmíněné dvě interpretace se přitom skutečně liší svými pravdivostními podmínkami. Zatímco v první interpretaci musí skutečně dojít k situaci, že lev

⁶² Tato věta je negramatická, neboť nominální fráze „Man“ je v angličtině povinně generická.

ucítí jídlo, načež tento lev zařve, ve druhé interpretaci k pravdivosti věty postačuje, že jeden lev ucítí jídlo a jiný lev zařve (což je situace, kterou lze v reálném světě vysvětlit např. tím, že samice mají silnější čich než samci, avšak pouze samci mohou řvát).

Další rozdíl v distribuci generických a nespecifických objektových frází lze pozorovat u tzv. kolektivních predikátů, jejichž subjektivním argumentem nemůže být výraz referující k množině objektů s kardinalitou 1. Např. ve větě *The lion gathers near acacia trees when it is tired* nelze namísto generické fráze *the lion* použít nespecifickou objektovou frázi *a lion*, protože agentem predikátu *gather* může být pouze skupina individuí. Generickou frázi naopak použít lze, protože při uplatnění relace **IS(x, Leo leo)** může být *x* interpretováno jako více než jeden reprezentant druhu „lev“.

Je tedy zřejmé, že protiklad generičnosti u nominálních frází má v angličtině závažné sémantické důsledky nejen v případě použití těchto frází v tzv. epizodických propozicích, ale i v propozicích generických.

⁶³ Leo leo je označením druhu „lev“.

6. Model reference a koreference

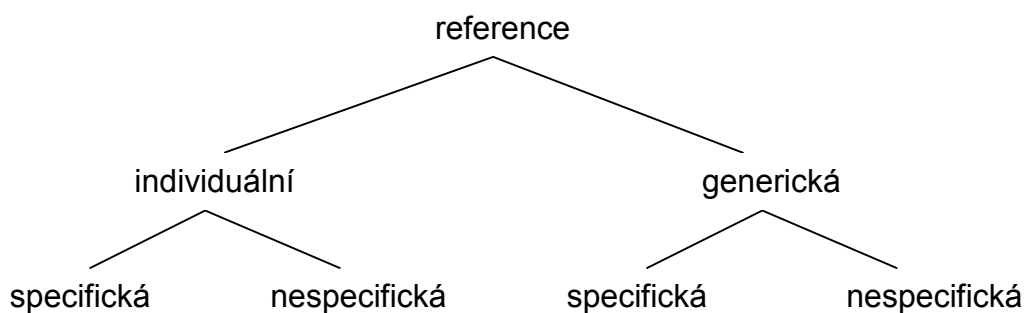
6.1 Typy reference

Na základě klasifikace typů reference a rozboru dvou typů generičnosti v citovaných studiích je možné přistoupit ke konstrukci alternativního modelu reference a koreference, který by zohledňoval poznatky obsažené v těchto studiích.

V typologii reference považují za výhodné pokládat za základní opozici mezi jmennými frázemi referujícími k objektu (v GenBook (1995) označovanými jako „object-referring NPs“) a jmennými frázemi referujícími k druhu (tamtéž nazývanými „kind-referring NPs“). Důvodem tohoto rozhodnutí je, že na rozdíl od některých jiných protikladů (např. protikladu specifičnosti či určenosti) se tato opozice dotýká přímo ontologické povahy referentu, nikoli např. „jen“ jeho vymezenosti či známosti. Referenci spojenou s prvním typem jmenných frází budu dále označovat jako *referenci individuální*, referenci spojenou s druhým typem jako *referenci generickou*.

Z publikace GenBook (1995) přejímám též striktní oddělení opozice specifičnosti od opozice generičnosti. Za jmenné fráze *specifické* (se specifickou referencí) budu pokládat ty, které referují k jasně vymezené entitě jakožto součásti aktuálního světa, zatímco jmenné fráze, které takto vymezený referent nemají, budu označovat jako jmenné fráze *nespecifické* (s nesespecifickou referencí). Rysu specifičnosti i nesespecifičnosti mohou nabývat jak jmenné fráze s individuální referencí (individuální jmenné fráze), tak jmenné fráze s referencí generickou (jmenné fráze generické).

Popsané rozdělení typů reference podle kritéria generičnosti a specifičnosti znázorňuje následující schéma:



Takto konstruované rozdělení typů reference v zásadě koresponduje s rozdělením typů reference podle GenBook (1995). Pokud jde o jeho vztah k dalším zmiňovaným modelům reference, lze konstatovat, že:

1. specifická individuální reference odpovídá Palkovým (1988) denotátům-konstantám, Hlavsovým (1975) jmenným frázím s operátory Un a Ex a Adamcově (1980) singulativní referenci,

2. nspecifická individuální reference odpovídá části Palkových denotátů-proměnných, části Hlavsových jmenných frází s operátorem Var a Adamcově podmíněně singulativní referenci,

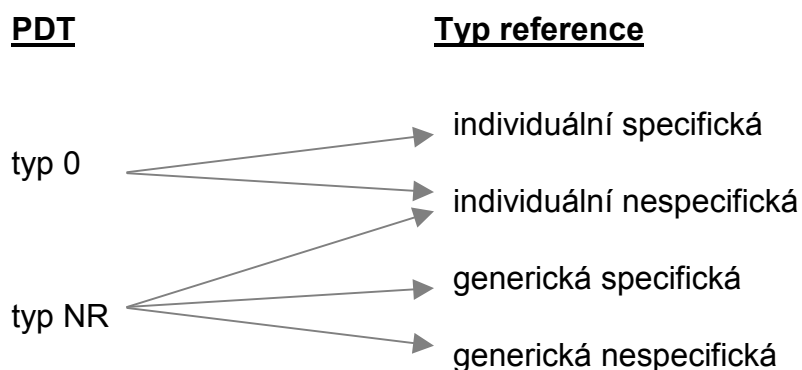
3. specifická generická reference koresponduje s částí Palkových denotátů-proměnných, částí Hlavsových jmenných frází s operátorem Var a většinou Adamcovy generické reference,

4. nspecifická generická reference odpovídá části Palkových denotátů-proměnných, části Hlavsových jmenných frází s operátorem Var a části Adamcovy generické reference.⁶⁴

Dále je třeba zmínit vzájemné vztahy mezi tímto modelem a typy reference a koreference, které jsou rozlišovány v rámci anotace koreference v Pražském závislostním korpusu. Především v pojetí PDT není přesně vymezen vztah mezi referencí generickou a nspecifickou, zatímco zde jsou rysy generičnosti a specifičnosti jasně popsány jako nezávislé. Většina případů koreference typu 0 spadá v navrhovaném pojetí do individuální specifické reference. Výjimkou jsou však případy, kdy jde o takový vztah koreference typu 0, kde antecedentem je negenerická jmenná fráze s nspecifickou referencí⁶⁵ – tyto případy navrhuji chápat jako případy nspecifické individuální reference. Jmenné fráze vstupující do vztahů anotovaných jako typ NR pak v popisovaném schématu mohou realizovat nspecifickou individuální referenci nebo jeden z typů reference generické. Vztahy mezi navrhovaným modelem a typy 0 a NR používanými v PDT jsou zachyceny následujícím schématem:

⁶⁴ Adamec se však v citovaném článku generickou referencí v zásadě nezabývá, takže jeho vymezení této reference není z článku zcela zřejmé. Je proto možné, že by v jeho pojetí do reference generické spadaly i některé případy, které v této práci budu řadit k nspecifické referenci individuální.

⁶⁵ Viz zde s. 9–10.



Pokud jde o koreferenci, je zřejmé, že koreferenční vztahy se vyskytují především (ovšem nikoli výhradně) mezi dvojicemi jmenných frází, z nichž buď mají obě individuální referenci, nebo obě mají generickou referenci.⁶⁶ V této souvislosti je třeba upozornit na významný rozdíl mezi koreferencí u těchto dvou typů reference. V případě koreference s individuální referencí jde u jmenných frází ve vztahu koreference o referenci k týmž objektům, zatímco v případě koreference s generickou referencí příslušné jmenné fráze referují ke stejnému druhu, přičemž tato skutečnost nevypovídá nic o objektech, které lze označit za reprezentanty tohoto druhu. Tento fakt je možné demonstrovat na dvojici vět převzatých z GenBook (1995):⁶⁷ *Kachna má barevné peří. Kachna snáší bílá vejce.* V obou těchto větách má jmenná fráze *kachna* specifickou generickou referenci, a jelikož obě tyto jmenné fráze referují ke stejnému druhu (druhu „kachna“), existuje mezi nimi koreference, a to přesto, že první věta vypovídá pouze o samčích reprezentantech příslušného druhu, zatímco druhá věta pouze o samičích.

V následujících částech se budu postupně zabývat jednotlivými druhy reference a možnostmi koreference v případě těchto druhů, konkrétně specifickou generickou referencí, nespecifickou individuální referencí a nespecifickou generickou referencí. Specifické individuální referenci, která není hlavním tématem této práce, zde zvláštní kapitolu nevěnuji, neboť její pojetí se v popisovaném modelu prakticky neliší od jejího zpracování v koncepci anotace koreference v Pražském závislostním korpusu.

⁶⁶ O případech koreference, kdy se příslušné jmenné fráze liší co do typu reference z hlediska generičnosti viz části 6.2.1 a 6.5.

⁶⁷ Zde uvádím jejich české ekvivalenty.

6.2 Specifická generická reference

V souladu s chápáním specifické generické reference v GenBook (1995) nabývají této reference ty jmenné fráze, které referují k vymezenému druhu jakožto součástí aktuálního světa.

6.2.1 Forma generických jmenných frází

GenBook (1995) vymezuje pro specifické generické jmenné fráze v angličtině čtyři formy při jejich základní (netaxonomické) interpretaci: spojení singulárového substantiva s určitým členem, plurálové substantivum bez členu, *singulare tantum* bez členu a tzv. vlastní jméno druhu. V případě taxonomické interpretace je možná i podoba spojení substantiva s neurčitým členem.⁶⁸

V češtině, kde neexistuje určitý a neurčitý člen, je možná generická interpretace jmenné fráze tvořené jak substantivem v singuláru, tak substantivem v plurálu (např. ve větách *Lev je nebezpečné zvíře* / *Lvi jsou nebezpečná zvířata*). Možná je v češtině rovněž taxonomická interpretace jmenných frází, např. existuje možnost substituovat jmenné fráze *lev* a *lvi* v uvedených větách frázemi *jedna šelma (totiž lev)* a *některé šelmy (např. lvi)*. Jako součástí generických jmenných frází vystupují v češtině často též substantiva hromadná či látková (např. *lidstvo, listí, zlato* apod.). Existují však kontexty, ve kterých mají jmenné fráze s těmito substantivy jinou referenci než specifickou generickou. Je tak např. nutné rozlišovat mezi kontexty *Na podzim opadává listí ze stromů* (specifická generická reference) a *Odnes prosím ze zahrady to listí* (specifická individuální reference) nebo *Zlato je drahý kov* (specifická generická reference) a *Nalezené zlato pochází ze 16. století* (specifická individuální reference).

Zvláštní případ představují substantiva, která referují k druhu již na základě svého lexikálního významu, jako je tomu např. u substantiv *druh, typ* apod. Tyto jmenné fráze (jako např. *this kind of tiger*) jsou v GenBook (1995, s. 6) chápány jako generické.⁶⁹ Argumentem proti jejich generickému chápání jsou však dle mého názoru dvojice typu: *Tento druh lva žije v Africe* / *Zoologie studuje druhy zvířat*. Pokud by jmenná fráze *tento druh lva* byla označena za generickou, vzniká otázka, jak její referenci odlišit od reference jmenné fráze

⁶⁸ Viz zde s. 29.

⁶⁹ Tento přístup ale není samozřejmý, např. v Pražském závislostním korpusu je v těchto případech anotována koreference typu 0.

druhy zvířat, která je jistě generická. Proto budu referenci typu *tento druh lva* pokládat za specifickou individuální referenci a referenci typu *druhy zvířat* za specifickou generickou referenci.

Podobným případem jsou pak substantiva, jejichž lexikální význam má rys hromadnosti, např. *národ*. Možnost koreference např. jmenné fráze *český národ* se zcela jistě generickou jmennou frází *Češi* by mohla vést k domněnce, že i první z těchto jmenných frází má generickou referenci. I v tomto případě si je ale možné představit dvojice kontextů, v nichž nabývají jmenné fráze s rysem hromadnosti různé reference, např. *Český národ žije v ČR / Národy mívají vlastní jazyk*. Proto budu jmenné fráze typu *český národ* pokládat za specifické individuální, zatímco jmenné fráze typu *národy* za specifické generické.

V případě substantiv, která svým lexikálním významem označují druh, i v případě substantiv s rysem hromadnosti může docházet k existenci koreferenčních vztahů mezi jmennými frázemi lišícími se co do typu reference z hlediska kritéria generičnosti, tj. např. existuje koreferenční vztah mezi jmennými frázemi *Češi* a *český národ* (v jakémkoli pořadí), kde první jmenovaná fráze má generickou referenci, druhá referenci individuální. Situace, kdy jsou druh a individuum označeny za vzájemně koreferenční, je paradoxní pouze na první pohled. Lze ji vysvětlit existencí taxonomické hierarchie, kterou např. v uvedeném případě tvoří pojmy *Čech – Češi / český národ – národ*. V této hierarchii označuje výraz *Čech* individuum, které je reprezentací druhu „český národ“, přičemž tento druh je individuem ve vztahu k druhu „národ“ (je jeho reprezentací).

6.2.2 Význam generických jmenných frází

Významy specifických generických jmenných frází v češtině v zásadě odpovídají významům, které jsou popsány v GenBook (1995, s. 78–85).⁷⁰ Uvádím zde proto výčet těchto významů s příkladovými větami, které jsou volnými překlady vět, s nimiž pracuje citovaná publikace:

1. vlastní reference k druhu – např. *Blboun nejapný je vyhynulý.*
2. souhrn reprezentací – např. *Němečtí zákazníci si loni koupili 11 000 nových BMW.*
3. průměrná reprezentace – např. *Mladí lidé se dívají na televizi šest hodin denně.*
4. reprezentace druhu – např. *Brambora obsahuje vitamin C.*
5. predikace odlišující vlastnosti – např. *Holand'ané jsou dobří námořníci.*

6. reference k reprezentativním objektům – např. *Na Aljašce jsme natáčeli medvědy grizzly.*

7. průkopnická reprezentace – např. *Člověk poprvé stanul na Měsíci v roce 1969.*

8. porovnání reprezentantů – např. *Cestou na sever jsou vlci větší a větší.*

Jako poněkud problematický se jeví čtvrtý z uvedených významů. Referentem jmenných frází s touto interpretací generické reference jsou totiž ve skutečnosti reprezentace příslušného druhu, tedy jednotlivé objekty, v důsledku čehož se reference těchto jmenných frází přibližuje individuální (nespecifické) referenci.

Tento význam se vyskytuje pouze v generických větách, v nichž dochází ke generalizaci nad objekty, které stojí ve vztahu reprezentace k příslušnému druhu. Význam těchto jmenných frází je tedy součástí restriktoru v sémantickém zápise příslušných vět, přičemž právě tato skutečnost (tedy generalizace zajišťovaná generalizačním operátorem **GEN**) umožňuje generickou interpretaci těchto jmenných frází. V angličtině je přitom při obsazování příslušných pozic v generických větách možné volit mezi jmennou frází generickou a negenerickou, neboť zde existuje forma, která vylučuje generickou interpretaci – spojení substantiva s neurčitým členem. V tomto jazyce je proto možné vytvořit dvojici navzájem různých vět: *The potato contains vitamin C* a *A potato contains vitamin C*. V první větě jde o jmennou frází generickou, ve druhé větě o nespecifickou jmennou frází individuální.

V češtině nelze tento rozdíl v generických větách vyjádřit, neboť specifické generické a nespecifické individuální jmenné fráze od sebe není možné formálně odlišit. Zdá se přitom, že tato skutečnost není nijak omezující, neboť pravdivostní podmínky obou anglických vět jsou stejné: to, co lze označit za (v daném ohledu) typickou bramboru, obsahuje vitamin C. Je známo, že generická jmenná fráze v této interpretaci nemusí referovat nutně ke všem existujícím reprezentacím, k nim ovšem v daném kontextu nerefereje ani nespecifická jmenná fráze individuální, neboť její kvantifikace je zajišťována generalizačním operátorem **GEN**.

Existují ovšem případy, kdy má volba mezi generickou a individuální (nespecifickou) jmennou frází v podobných kontextech v angličtině důsledky i co do pravdivostních podmínek věty. Autoři (GenBook, 1995) to dokládají na již zmiňované dvojici vět *A lion roars when it smells food* / *The lion roars when it smells food*.⁷¹ Rozdílné pravdivostní podmínky těchto vět je možné vysvětlit na základě již zmiňovaných poznatků o koreferenci –

⁷⁰ Viz též zde s. 29–31.

⁷¹ Viz zde s. 31–32.

zatímco v první větě jde o koreferenci mezi dvěma jmennými frázemi s individuální referencí (*a lion – it*), ve druhé větě jsou obě části koreferenčního vztahu generické, a proto nejde o identitu jednotlivých objektů, nýbrž o identitu druhů, z nichž jsou objekty vybírány.

V případech typu *The / A lion roars when it smells food* tedy mezi oběma variantami existuje nejen rozdíl formální (který není možné v češtině vyjádřit), ale též rozdíl významový. Je přitom třeba předpokládat, že významový rozdíl se v těchto případech uplatňuje i v češtině, resp. že v češtině existují analogické případy, v nichž jde o koreferenci nikoli mezi druhy, nýbrž mezi objekty. Příkladem těchto případů jsou následující věty (převzaté z PDT): *Pokud daňové přiznání zpracovává daňový poradce, musí ho podat do 30. 6. a Důvodů je mnoho, primární ten, že v kandidátovi politické strany volí občan nejen konkrétní osobu, ale i politickou stranu, která je jeho zázemím.* V případě obou vět jde jistě o koreferenci mezi objekty, nikoli „jen“ mezi druhy: daňový poradce musí podat to přiznání, které zpracovává, a pokud člověk volí kandidáta jisté politické strany, tak tím volí také tuto stranu, nikoli stranu jinou.

V podobných případech budu proto u příslušných jmenných frází předpokládat individuální (nespecifickou) referenci. V ostatních případech (tedy pokud v případě koreference nejde nutně o shodu objektů, nýbrž o identitu druhů) budu jmenné fráze tvořící restriktor významového zápisu generických vět pokládat za generické.

Z právě uvedeného pravidla vyplývá, že jako generické jmenné fráze by měly být chápány též jmenné fráze typu *lvi, které bolí zuby*⁷². Tyto jmenné fráze nejsou v GenBook (1995) pokládány za generické, neboť nereferují k „well-established kind“, tedy k entitě, která je všeobecně chápána jako druh. V angličtině je toto rozhodnutí motivováno nemožností konstrukce jmenné fráze s určitým členem a substantivem v singuláru s generickou interpretací – taková konstrukce je možná jen v případě, že jde o „well-established kind“. Např. jmenná fráze *the lion that has toothache* nemůže mít generickou interpretaci (na rozdíl od jmenné fráze *the lion*). V češtině však zřejmě možnost generického chápání příslušných jmenných frází omezena není, jak ukazují následující příklady: *Lev, kterého bolí zuby, je nebezpečný / Lvi, které bolí zuby, jsou nebezpeční.* Navíc se tyto jmenné fráze chovají z hlediska reference stejně jako „běžné“ generické jmenné fráze, tj. referují k otevřené množině prvků, které mají shodné relevantní vlastnosti. Proto budu v této práci i tyto jmenné fráze pokládat za generické.

Jmenné fráze, které získávají rys generičnosti na základě přítomnosti generalizačního operátoru GEN, se mohou vyskytovat též v adnominálních pozicích. Jde o ty případy, v nichž

je vztah mezi závislým a řídicím substantivem možno vyjádřit významovým zápisem, ve kterém závislé substantivum tvoří součást restriktoru. Generickou referenci má tak např. výraz *čokolády* ve spojení *milovník čokolády*, neboť příslušný vztah (tedy skutečnost, že milovník čokolády je člověk, který miluje čokoládu) je možno vyjádřit takto: **GEN[x;y](IS(x, čokoláda); (IS(y, milovník čokolády) & y miluje x)).**⁷³ Tyto případy je potřeba odlišit od podobných případů nespecifické individuální reference.⁷⁴

Zvláštní případ představují jmenné fráze referující k entitám explicitně kvantifikovaným⁷⁵ (např. jmenná fráze *dva vlaky*). Tyto jmenné fráze mohou běžně nabývat specifické i nespecifické individuální reference, možnost jejich generické interpretace (která se nabízí např. ve větě *Dva vlaky jsou lepší než jeden vlak*) je však poněkud problematická, neboť kvantifikovány jsou v těchto případech jednotlivé reprezentace druhu, tedy objekty.⁷⁶ Při širším chápání druhu, tedy chápání druhu jako otevřené množiny prvků se shodnými relevantními vlastnostmi, je ale generické chápání těchto jmenných frází možné, např. jmenná fráze *dva vlaky* tak (poněkud neintuitivně) referuje k druhu, jehož reprezentacemi jsou všechny množiny sestávající z právě dvou vlaků.

V případech, kdy jsou podobné jmenné fráze součástí restriktoru generických vět (jako je tomu ve větě *Dva vlaky jsou lepší než jeden vlak*), budu tyto jmenné fráze pokládat za generické, a to z několika důvodů. Zaprvé fungují v generických větách podobně jako typické generické jmenné fráze, tak např. reference jmenných frází ve zmiňované větě je analogická referenci ve větě *Vlak je lepší než auto*. Generické chápání jmenné fráze *dva vlaky* navíc umožňuje chápat tuto jmennou frázi jako koreferenční s případným dalším výskytem formálně shodné jmenné fráze (např. ve větě *Dva vlaky ovšem nejsou lepší než dvě letadla*), neboť obě jmenné fráze při tomto chápání referují ke stejnému druhu.⁷⁷ Dalším argumentem pro chápání těchto jmenných frází v příslušném kontextu jako generických je jejich analogie (a potenciální koreference) se jmennými frázemi typu *dvojice vlaků*, jejíž generické interpretaci ve zmiňovaných kontextech nebrání žádné překážky.

⁷² Jde o volný překlad jmenné fráze, která je jako příklad uvedena v GenBook (1995).

⁷³ Jde tedy o to, že milovníkem čokolády je člověk, pro kterého platí, že miluje **všechno** to, co lze označit za (relevantní) čokoládu.

⁷⁴ O tom viz kapitola 6.3.2, s. 45–46.

⁷⁵ Jak již bylo upozorněno (viz s. 20), Hlavsa (1975) považuje v procesu denotace kvantifikaci za operaci paralelní k operaci determinace.

⁷⁶ Pochopitelně zde mám na mysli pouze případy, kdy jsou skutečně kvantifikovány jednotlivé objekty. Jmenné fráze referující ke kvantifikovaným druhům (např. *dva lvi*, *lev jihoafrický* a *lev konžský*) představují jiný případ (jde jednoznačně o specifickou generickou referenci).

⁷⁷ V případě individuálního chápání těchto jmenných frází by byla jejich koreference poněkud problematická, neboť se v obou případech nemusí jednat o tytéž dva vlaky.

6.3 Nespecifická individuální reference

Nespecifickou individuální referenci mají ty jmenné fráze, které referují k objektu (nikoli k druhu), který není vymezen (není tedy identifikovatelný, netvoří konečnou množinu prvků) nebo vůbec není součástí aktuálního světa. Tyto jmenné fráze je možné snadno odlišit od jmenných frází s individuální referencí specifickou (neboť ty referují k vymezeným objektům aktuálního světa), jak však bude uvedeno dále, poněkud problematičtější může být jejich zřetelné odlišení od (specifických) jmenných frází generických.

Pro popis nespecifické reference postupně zavedu pojmy *rys hypotetičnosti* a *rys nevymezivosti*. Za nespecifické pak budu pokládat ty jmenné fráze, u kterých bude přítomen alespoň jednoho z těchto dvou rysů (oba rysy se mohou u jedné jmenné fráze vyskytovat též současně).

6.3.1 Jmenné fráze s rysem hypotetičnosti

Rys hypotetičnosti má taková individuální jmenná fráze, která referuje nikoli k objektu aktuálního světa, nýbrž k objektu světa potenciálního, tj. světa, který je v textu přítomen pouze jako možnost.⁷⁸ Takový objekt budu nazývat hypotetickým objektem. „Hypotetičnost“ objektů je zajišťována příslušným predikátem, který referuje k eventualitě, jejíž realizace v aktuálním světě je buď nějakým způsobem podmíněna, nebo je přímo vyloučena (dále budu tyto eventuality označovat termínem *hypotetické eventuality*). Pokud by přítom k realizaci takové eventuality v aktuálním světě došlo, měla by příslušná jmenná fráze specifickou individuální referenci.

Jmenné fráze s rysem hypotetičnosti se mohou vyskytovat především v následujících kontextech:

1. kondicionál – např. *Karel by si koupil auto, kdyby měl peníze.*
2. imperativ – např. *Kup si auto!*
3. futurum – např. *Koupím si auto.*
4. negace – např. *Nekoupil jsem si auto.*
5. epistémická modalita – např. *Možná si koupím auto.*

⁷⁸ Jde o jmenné fráze, které Adamec (1980) pokládá za jmenné fráze s podmíněně singulativní referencí.

6. deontická modalita – např. *Musím si koupit auto*.⁷⁹

Jak ukazují data z Pražského závislostního korpusu, hypotetické eventuality jsou v češtině často vyjadřovány slovesem v neurčitém tvaru, které je ve funkci valenčního doplnění jiného slovesa, případně predikátovým slovesem závislé klauze. Nespecifickou referenci pak může mít doplnění závislého slovesa (příp. predikátu závislé klauze), např. ve větách *Svolat morální tribunál doporučil ředitel Střediska pro dokumentaci protiprávnosti komunistického režimu* či *Slovenská vláda vyzvala ministerstvo zahraničních věcí SR, aby (...) začalo připravovat mezinárodní smlouvu, která by upravila tento problém*. Hypotetická eventualita pochopitelně může být v podobných kontextech vyjádřena též dějovým substantivem, které tak získává nespecifickou individuální referenci, např. ve větách *Spolupráci nabízí rovněž dalším českým společnostem* či *Předseda ODA Jan Kalvoda považuje za velmi liché návrhy KDS na změnu zákona o nutné obraně*.

Vztah koreference mezi dvěma jmennými frázemi s rysem hypotetičnosti navrhuji předpokládat tehdy, pokud tyto jmenné fráze referují k objektům, které by v případě realizace příslušných hypotetických eventualit byly totožné. Tak např. ve dvojici vět (převzatých z PDT) *V nádvoří paláce Smiřických by měla být zřízena dokonce kavárna pro veřejnost (...) Část poslaneckou a veřejnou však nelze nijak oddělit, proto vždy po zavírací době kavárny projde ochranka se psem celé přízemí a prohlédne je* by se v případě, že by ke zřízení kavárny skutečně došlo, jednalo v obou větách o tutéž existující kavárnu, proto jsou obě nespecifické individuální jmenné fráze koreferenční.⁸⁰

Jak již bylo uvedeno výše,⁸¹ při anotaci koreference v PDT se v těchto případech postupuje tak, že první z dvojice jmenných frází je pokládána za nespecifickou, ovšem druhá je, má-li určitý identifikátor, chápána jako specifická a koreference mezi těmito jmennými frázemi je typu 0 (např. v již zmiňované větě *Poprosím o to některého kolegu a ten mi to řekne*). Takové řešení nepokládám za náležité, neboť v těchto kontextech určitý identifikátor nezpůsobuje, že by jmenná fráze referovala k reálně existujícímu (identifikovatelnému)

⁷⁹ Ve všech těchto kontextech se pochopitelně může vyskytovat též jmenná fráze se specifickou individuální referencí, např. při substituci příslušných jmenných frází spojením *toto auto* (v takovém případě by jmenná fráze referovala ke konkrétnímu autu).

⁸⁰ Naopak koreferenční nejsou např. podtržené jmenné fráze v následujících větách: *Petr si chce koupit auto. Karel si chce také koupit auto. / Karel si ho také chce koupit*. Důvodem je, že pokud by v případě Petra i Karla k nákupu auta došlo, šlo by o dvě různá auta. Na této skutečnosti nic nemění ani možnost pronominalizace (v tomto případě jde o nekoreferenční anaforu).

⁸¹ Viz s. 9–10.

objektu (např. ve zmíněné větě nelze identifikovat referenta výrazu *ten*⁸²), pouze poukazuje na existenci koreferenčního vztahu. Vázanost typu reference na přítomnost určitého identifikátoru je navíc problematická v tom, že v mnohých kontextech je jeho přítomnost možná, ale nikoli nutná. Tak ve zmíněném příkladu s kavárnou je možné jmennou frází ve druhé větě substituovat jmennou frází *této kavárny*, aniž by se změnil význam.

Kromě dvou jmenných frází s nespécifickou individuální referencí však dle mého názoru do koreferenčního vztahu mohou vstupovat i jmenná fráze s nespécifickou individuální referencí a jmenná fráze se specifickou individuální referencí. Děje se tak tehdy, pokud by při realizaci příslušné hypotetické eventuality obě jmenné fráze referovaly k témuž objektu. Příkladem jsou věty *Museli vyhotovit architektonickou studii stavby a projekt na její využití. To squateři již jako nadace Ladronka učinili. Stavební část projektu prý pro ně vypracovali studenti 3. ročníku Fakulty architektury ČVUT. Zatímco v první větě jde o nespécifickou individuální referenci (mluví se zde pouze o nutnosti vyhotovení projektu, tedy k jeho vyhotovení v příslušné časové rovině ještě nedošlo, a projekt proto nebyl součástí aktuálního světa), ve třetí větě má již *projekt* specifickou referenci, neboť již bylo konstatováno, že projekt byl vyhotoven. Obě jmenné fráze jsou tedy koreferenční, co do typu reference podle specifičnosti se ovšem liší. Podobná je situace i v následujícím příkladě, kde naopak předchází specifická jmenná fráze před nespécifickou: *Setkání premiérů Maďarska a Slovenska, které se uskutečnilo v neděli večer, mělo komplikovaný úvod. (...) Proto začal premiér Boross projevovat zájem o setkání s Moravčíkem.**

Mezi jmenné fráze s rysem hypotetičnosti řadím i některé jmenné fráze nacházející se v restriktoru generických vět, jde o případy vět odpovídajících anglické větě *A lion roars when it smells food.*⁸³ Od jmenných frází s generickou referencí se tyto jmenné fráze odlišují tím, že netvoří restriktor významového zápisu samostatně, nýbrž jsou v restriktoru doplněny nějakou specifikací. Tak např. český ekvivalent zmíněné věty *Lev zařve, když ucítí potravu* má významový zápis **GEN[x,s;](x je lev & x ucítí potravu v s; x zařve v s)**, tj. v restriktoru se nachází nejen informace o lvu, nýbrž i informace o situaci, kdy lev cítí potravu. Intuitivně lze tuto skutečnost interpretovat tak, že příslušná jmenná fráze nerefereuje k celému druhu, tj. ke všem relevantním lvům, nýbrž pouze k těm lvům, pro které platí, že se nacházejí v určité situaci. Z toho je zřejmé, že v těchto případech nemůže jít u takové jmenné fráže o generickou referenci.

⁸² Jeho identifikace pomocí antecedentu tu není možná, protože antecedent zde identifikovatelného referenta nemá, neboť má sám nespécifickou referenci.

⁸³ O nich bylo částečně pojednáno již v části 6.2.2, s. 38–39.

Příklady tohoto typu nespecifické individuální reference (převzaté z PDT) jsou např. věty *A když si na jednu lod' zvyknu, nerad ji měním* nebo *Víte, když se někde dejme tomu unese letadlo se čtyřmi sty pasažéry, svět se okamžitě postaví za jejich záchranu, uzavírá doktor Kulenovič* (podtržené jmenné fráze mají nespecifickou individuální referenci a jsou ve vztahu koreference). Tento typ ovšem není omezen jen na souvětí se spojkou *když* a nemusí se týkat jen jmenných frází v rámci jedné věty, jak ukazují další příklady z PDT: *Zaměstnanci přecházejí do české či slovenské části v poměru 2:1 a tzv. poznatkové fondy se předávají té republice, která je vytvořila*⁸⁴ / *Pokud se nějaká země usnese na ústavním zákonu, pak se to těžko mění. Znamenalo by to pro tu zemi stabilitu a nebylo by ji možné ničit neustálým schodkovým rozpočtem, řekl Mandelík.*

6.3.2 Jmenné fráze s rysem nevymezenosti

Rys nevymezenosti je založen na tom, že jmenná fráze referuje k množině prvků, která nemá jasně stanovené hranice, tj. jde o množinu otevřenou, tj. potenciálně nekonečnou.⁸⁵ Tato situace může nastávat v generických větách a je způsobena tím, že generické věty referují k nevymezenému množství eventualit.

Jmenné fráze s rysem nevymezenosti jsou charakteristické tím, že se v generických větách nacházejí mimo restriktor významového zápisu. Příkladem takové jmenné fráze je jmenná fráze *auta* ve větě *Lidé mají auta*. Její sémantický zápis vypadá takto: **GEN[x;y](IS(x, lidé); y je auto, x má y)**⁸⁶. Jmenná fráze *lidé* je v restriktoru, protože součástí významu věty je zobecnění nad všemi relevantními lidmi, a proto má tato jmenná fráze specifickou generickou referenci. Jmenná fráze *auta* je mimo restriktor a označuje otevřenou množinu objektů, proto jde o nespecifickou individuální frázi s rysem nevymezenosti. Tento příklad také ukazuje, proč není možné pokládat jmenné fráze mimo restriktor za generické – nedochází zde totiž ke generalizaci, tj. jmenná fráze v této větě nereferuje ke všem autům (s možnými výjimkami), nýbrž pouze k některým autům, tj. k těm, která jsou vlastněna lidmi.

⁸⁴ V tomto případě je v restriktoru nejen informace o poznatkových fondech, ale i informace o jejich vytvoření příslušnou zemí.

⁸⁵ Důsledkem této skutečnosti se jmenné fráze s rysem nevymezenosti podobají generickým jmenným frázím, neboť ty referují k druhu, který je rovněž chápán jako otevřená množina prvků. Je tedy nutné rozlišit, které otevřené množiny je možno chápat jako druh.

⁸⁶ Tento zápis odpovídá tomu významu věty, podle kterého jsou auta lidem distributivně přiřazena tak, že každý člověk, který má auto, má právě jedno auto.

Jak již bylo uvedeno výše,⁸⁷ model významu generických jmenných frází zachyceného pomocí operátoru **GEN** souvisí s aktuálním členěním věty, neboť dokáže rozlišit dvojice vět, které se liší pouze co do rozdělení na základovou část a ohnisko. Jelikož interpretace reference některých jmenných frází závisí na jejich pozici v sémantickém zápisu věty, je zřejmé, že v popisovaném modelu může mít aktuální členění věty vliv na jejich referenci. Jako příklad může posloužit dvojice vět *Ševci prodávají boty* a *Boty prodávají ševci*.⁸⁸ V případě první věty je v restriktoru jen výraz *ševci*, v případě druhé věty jen výraz *boty*. Z toho vyplývá, že v první větě má výraz *ševci* specifickou generickou referenci a výraz *boty* nespecifickou individuální referenci, zatímco v druhé větě má výraz *ševci* individuální nespecifickou referenci a výraz *boty* specifickou generickou referenci. Tyto odlišnosti se pochopitelně projevují i v rozdílu mezi intuitivními porozuměními významu těchto vět. Zatímco první věta vypovídá o všech relevantních ševcích (tedy o ševcích jakožto o druhu), ale nemusí vypovídat o všech relevantních botách (nevyplyvá z ní, že by všechny relevantní boty prodávali ševci, tedy věta připouští možnost, že většinu bot prodává někdo jiný než ševci, a je tedy zřejmé, že nevypovídá o celém druhu „boty“), druhá věta vypovídá o botách jakožto druhu (tj. o všech relevantních botách), ale nevyplyvá z ní, že by všichni relevantní ševci prodávali boty.

Vztah mezi aktuálním členěním věty a významovým zápisem generických vět ovšem není přímočarý, neplatí tedy, že by základ věty byl vždy v restriktoru. V některých případech je v restriktoru významového zápisu výraz, který se nachází v ohnisku věty. Příkladem je věta *Karel má rád auta*, jejíž významový zápis vypadá takto: **GEN**[x;y](**IS**(x, **auta**); y = **Karel** & y **má rád** x). V tomto případě je tedy v restriktoru výraz *auta* (neboť dochází ke zobecnění nad všemi relevantními auty), a proto má tento výraz specifickou generickou referenci. Podobným případem je věta *Lidé mají rádi auta*, v jejímž významovém zápisu je v restriktoru jak výraz *lidé*, tak výraz *auta* (a obě jmenné fráze jsou tudíž generické).

Jmenné fráze s rysem nevymezenosti se mohou vyskytovat též v adnominálních pozicích, pokud vztah mezi řídicím a závislým substantivem není možné vyjádřit tak, aby závislé substantivum (s rysem nevymezenosti) tvořilo součásti restriktoru významového zápisu (konstruovaného pomocí operátoru **GEN**). Jde např. o výraz *bot* ve spojení *výrobce bot* nebo ve spojení *výroba bot* ve větě *Předmětem našeho podnikání je výroba bot*. Tyto případy se tedy liší od případů typu *milovník čokolády*,⁸⁹ neboť nelze např. tvrdit, že výrobce bot je

⁸⁷ Viz část 5.1.

⁸⁸ Počítám zde s jejich bezpříznakovou intonací s intonačním centrem na posledním slově.

⁸⁹ K těmto případům viz s. 39–40.

člověk, který vyrábí všechno to, o čem lze říci, že je to bota (nevyrábí tedy všechny relevantní reprezentanty druhu „bota“, pouze nevymezenou část těchto reprezentantů). Rovněž ve spojení *výroba bot* v uvedené větě má výraz *bot* nespécifickou individuální referenci.⁹⁰

Pokud jde o koreferenci výrazů s tímto typem nespécifické reference, mohou tyto výrazy vstupovat do koreferenčního vztahu, pokud referují ke stejné množině objektů. Např. na zmíněnou větu *Lidé mají auta* může v konkrétním textu navazovat věta *Jezdí s nimi do práce* nebo *Tato auta jim slouží k cestě do práce*, kde výrazy *nimi* a *tato auta* jsou koreferenční s jmennou frází *auta* a rovněž mají nespécifickou individuální referenci.

V případě pokračování *Tato auta jim slouží k cestě do práce* je situace poněkud složitější, neboť zde není na první pohled jasné, k jakému zobecnění zde dochází, tj. zda dochází ke zobecnění nad auty, nebo nad lidmi. Na první pohled by se dokonce mohlo zdát, že jde o zobecnění nad auty, což by znamenalo, že by jmenná fráze *tato auta* musela mít generickou referenci. Takové řešení však dle mého názoru není nutné, neboť v podobných případech je oprávněné předpokládat, že věty v textu tvoří významově koherentní celek, a že proto u dvojice vět typu *Lidé mají auta. Tato auta jim slouží k cestě do práce* nedochází ke změně perspektivy zobecnění. I ve druhé větě se mi proto jeví oprávněné chápat jako součást restriktoru pouze výraz referující k lidem (tj. *jim*), zatímco výraz *tato auta* je možné chápat jako jmennou frází s nespécifickou individuální referencí.⁹¹

Pokud dvě jmenné fráze s nespécifickou individuální referencí referují k různým množinám objektů, vztah koreference mezi nimi nenastává, a to ani při pronominalizaci. Tuto situaci ilustruje dvojice vět *Psi mají ocas. Kočky jej mají také*. Výrazy *ocas* a *jej* nejsou koreferenční, neboť nereférují ke stejným množinám objektů.⁹²

Závěrem této části se ještě pro úplnost zmíním o nespécifických individuálních jmenných frázích, které mají rys hypotetičnosti i nevymezenosti současně. Takové jmenné fráze se nacházejí mimo restriktor generických vět a zároveň referují k objektům potenciálních světů. Příkladem může být jmenná fráze *auta* ve větě *Lidé chtějí mít auta*.

⁹⁰ Jinak by tomu ale bylo např. ve větě *Výroba bot je složitý proces*, neboť tato věta vypovídá o výrobě jakýchkoli (tj. všech relevantních) bot, a proto zde jde o spécifickou generickou referenci (je to podobný případ jako typ „milovník čokolády“).

⁹¹ Zcela jinou situaci je ovšem pochopitelně dvojice vět typu *Lidé mají auta. Auta jsou neekologická*. V tomto případě je výraz *auta* ve druhé větě součástí restriktoru, a proto má generickou referenci (ke koreferenci zde tak samozřejmě nedochází). Je to dáno tím, že jde o referenci na všechna (relevantní) auta, nikoli pouze na ta auta, která jsou vlastněna lidmi.

⁹² Jelikož zde nejde o generickou, nýbrž o individuální referenci, není relevantním kritériem shoda druhu, ale shoda objektů.

6.4 Nespecifická generická reference

Výrazy s nespecifickou generickou referencí podobně jako výrazy se specifickou generickou referencí referují k druhu. Jejich reference je ovšem nespecifická, tyto jmenné fráze tedy mají rys hypotetičnosti nebo nevymezenosti. Lze očekávat, že jmenné fráze s tímto typem reference se v reálných textech budou vyskytovat spíše zřídka, neboť jejich referentem je již pojem značně abstraktní, pro úplnost modelu je však myslím třeba i s tímto typem počítat.

Pokud jde o formu nespecifických generických jmenných frází, je možné předpokládat jak existenci jmenných frází s interpretací taxonomickou, tak těch s interpretací netaxonomickou. Příkladem fráze s taxonomickou interpretací je jmenná fráze *nějaké nové zvíře* ve větě *Chtěl bych objevit nějaké nové zvíře*,⁹³ frází s netaxonomickou interpretací je např. fráze *příslušníci nějakého druhu* ve větě *Když jsou příslušníci nějakého druhu ohroženi, začnou být chráněni*.⁹⁴

Nespecifičnost generických jmenných frází, stejně jako možnost koreference takových frází, se může uplatňovat v podobných kontextech a za podobných podmínek, jako je tomu v případě nespecifické reference individuální. Tyto podmínky byly podrobně popsány v části 6.3, zde proto uvedu pouze příklady jednotlivých typů:

1. jmenná fráze s rysem hypotetičnosti – *Možná objevím nějakou novou sloučeninu*.
2. jmenná fráze s rysem nevymezenosti – *Lidé obvykle mají své oblíbené jídlo*.
3. jmenná fráze s oběma rysy – *Lidé by chtěli objevit nějakou novou rostlinu*.

Jak již bylo uvedeno,⁹⁵ některé nespecifické individuální jmenné fráze jsou referenčně blízké generickým specifickým frázím – pokud se nespecifická individuální jmenná fráze, která figuruje v generické větě mimo restriktor, dostává do restriktoru, zpravidla nabývá specifické generické reference, přičemž množina nevymezených objektů je takto „povýšena“ na druh. V případě nespecifických generických jmenných frází je proto možné očekávat podobné chování. Např. ve větě *Ohrožená zvířata jsou chráněna* má jmenná fráze *ohrožená*

⁹³ Ve významu „nějaký dosud neznámý druh zvířete“.

⁹⁴ Zde nejde o nespecifickou individuální referenci, jak by se mohlo na první pohled zdát, protože v případě této jmenné fráze jde o tzv. vlastní referenci k druhu (viz s. 29), tj. predikát *být ohrožen* v příslušném významu vyžaduje, aby byla jeho argumentem generická fráze (podobně jako např. predikát *vyhynout*). Koreference této jmenné fráze s nevyjádřeným subjektem hlavní věty navíc nevyžaduje totožnost individuí, nýbrž pouze totožnost druhu.

⁹⁵ Viz s. 44.

zvířata specifickou generickou referenci, neboť označuje druh, jehož reprezentanty jsou všechny relevantní možné ohrožené druhy zvířat. Tento druh je tedy otevřenou množinou druhů. Bylo by tak možné uvažovat o jakémisi „druhu druhého stupně“, něco takového ovšem dle mého názoru není nutné, neboť od podobných jmenných frází zřejmě není možné očekávat zvláštní referenční chování.

6.5 Možnosti koreference

V této části postupně shrnu, jaké páry jmenných frází (z hlediska typu reference) mohou vstupovat do vzájemného vztahu koreference. Na základě toho, co bylo již uvedeno o jednotlivých typech reference, je zřejmé, že koreferenční vztahy mohou existovat mezi jmennými frázemi se všemi kombinacemi popsaných typů reference, v některých případech je ovšem taková možnost vázána na určité lexikální obsazení.

Jako nejtypičtější se jeví takové koreferenční vztahy, do kterých vstupují jmenné fráze se stejným typem reference. Příkladem koreference jmenných frází se specifickou individuální referencí je např. vztah mezi výrazy *Karel* a *ho* ve větách *Karel byl dnes ve škole. Viděl jsem ho na chodbě*. Koreference jmenných frází s nespécifickou individuální referencí se uplatňuje např. mezi výrazy *dceru* a *jí* ve větách *Chtěla by mít dceru. Dala by jí jméno Natálie*. Pokud jde o specifickou generickou referenci, najdeme koreferenci např. mezi výrazem *čokoládu* a nevyjádřeným subjektem druhé věty ve větách *Mám rád čokoládu. Je nejen dobrá, ale také zdravá*. Konečně koreference mezi výrazy s nespécifickou generickou referencí se uplatňuje např. mezi frázemi *nějaké nové zvíře* a *ho Petr by chtěl objevit nějaké nové zvíře. Pojmenoval by ho podle sebe*.

Dále se vztah koreference může uplatňovat u dvojic jmenných frází, které jsou obě buď generické, nebo individuální, a které se liší co do typu reference podle kritéria specifičnosti. Může se tak ovšem dít pouze tehdy, má-li příslušná nespécifická jmenná fráze rys hypotetičnosti. V případě, že má jmenná fráze rys nevymezenosti, nemůže vstupovat do vztahu koreference se jmennou frází specifickou, protože označuje otevřenou množinu prvků, zatímco specifická jmenná fráze označuje uzavřenou množinu (a tyto dvě množiny tedy nemohou být totožné). Příklad koreference dvou individuálních jmenných frází (lišících se co do specifičnosti) najdeme v již citovaných větách *Museli vyhotovit architektonickou studii stavby a projekt na její využití. To squateři již jako nadace Ladronka učinili. Stavební část projektu prý pro ně vypracovali studenti 3. ročníku Fakulty architektury ČVUT*. Koreference

nastává mezi výrazem *projekt* (který má nespécifickou individuální referenci) a výrazem *projektu* (který má specifickou individuální referenci). Podobný vztah koreference si lze představit i mezi jmennými frázemi s generickou referencí, např. mezi slovy *přístroj* a *něhož* (reference generická nespécifická) a *ho* (reference generická specifická ve větách *Bell se dlouho snažil vymyslet přístroj, pomocí něhož by spolu lidé mohli komunikovat na dálku. Vymyslel ho v roce 1876.*

V některých případech může docházet též ke koreferenci mezi jmennými frázemi, z nichž jedna má individuální a druhá generickou referenci.⁹⁶ Je to však možné pouze v případech (popsaných v části 6.2.1), kdy je individuální jmenná fráze představována výrazem, který označuje druh svým lexikálním významem nebo který má rys hromadnosti. Příkladem jsou výrazy *snovačka jedovatá* (specifická generická reference) a *tento druh* (specifická individuální reference) ve větách *Snovačka jedovatá je velmi nebezpečné zvíře. Tento druh ale u nás nežije.* Pochopitelně i v těchto případech se mohou příslušné jmenné fráze lišit co do rysu specifčnosti a je tak možné představit si koreferenční vztah mezi všemi zbývajícími typy reference. Příklady je možné získat substitucí generických frází v dříve uvedených větách individuálními jmennými frázemi obsahujícími např. slovo *druh* (např. *Petr by chtěl objevit nějaké nové zvíře. Pojmenoval by tento druh podle sebe.*)

6.6 Řešení případů uvedených v kapitole 3.

Ve 3. kapitole byly představeny tři okruhy problematických případů, které jsou spojené s anotačním schématem PDT, založeným na dvou typech reference, typu 0 a typu NR. V tomto modelu je všechny tyto případy možné efektivně vyřešit.

V případech popsáných v části 3.1, tedy případech typu *Chtěla by mít dceru. Dala by jí jméno Natálie*, budou mít oba koreferenční výrazy nespécifickou individuální referenci (s rysem hypotetičnosti). Tím se odstraní jak sporná vázanost typu reference na určitý identifikátor, tak i nenáležitě chápání reference anaforu jakožto reference specifické a nelogická neshoda typů reference u obou koreferenčních výrazů.

Část 3.2 pojednávala o větách typu *Obratlovci mají oči*. V tomto případě bude mít výraz *oči* nespécifickou individuální referenci (s rysem nevymezenosti), čímž se odstraní jeho generická interpretace, a tedy i nutnost koreference s výrazem *oči* ve větě *Hmyz má také oči*. Pokud jde o uváděnou větu *Tyto oči však vypadají zcela jinak než oči obratlovců*, zde již

budou mít oba příslušné výrazy generickou referenci (neboť jsou v dosahu operátoru **GEN**), každý z nich však bude referovat k jinému druhu očí. Výrazy ve třetí větě tedy nebudou koreferenční s výrazy ve větách předcházejících (neboť ty nejsou generické), přičemž v případě jmenné fráze *tyto oči* půjde o nekoreferenční anaforu.⁹⁷

Třetím případem byly věty typu *Když je stůl velký, vejde se na něj spousta věcí* (část 3.3). Situaci zde opět řeší kategorie nespecifické individuální reference, která umožňuje chápat výrazy *stůl* a *něj* jako vzájemně koreferenční a zároveň nás zbaví nutnosti uvažovat o koreferenci těchto výrazů s výrazy ve větě *Když je ale stůl malý, nevejde se na něj skoro nic*.

⁹⁶ Příslušné jmenné fráze se pochopitelně nemusí vyskytovat v tomto pořadí, neboť koreference je relací symetrickou.

⁹⁷ Absence koreference tu není nikterak na závadu, neboť nekoreferenční anafora, v níž má antecedent individuální referenci, zatímco anafor referenci generickou, je poměrně běžná. Obsáhle je o ní pojednáno v publikaci GenBook (1995), kde se její častý výskyt na s. 353–354 vysvětluje takto: „any occurrence of a CN [= common noun] in discourse would simultaneously introduce the kind in the main discourse representation structure, thus (...) making it always accessible for anaphora“.

7. Analýza korpusových textů

V předchozích výkladech jsem za účelem větší názornosti používal jako textový materiál především vlastní příkladové věty nebo věty převzaté z jiných prací. V této části proto provedu analýzu koreferenčních vztahů ve třech textech vybraných z Pražského závislostního korpusu (PDT).

Výsledky této analýzy pochopitelně nemohou sloužit jako reprezentativní zdroj informací např. o zastoupení jednotlivých typů reference v češtině – za účelem jejich získání by bylo třeba analyzovat výrazně větší množství textů, což ale přesahuje možnosti a hlavní cíl této práce. Při výběru textů jsem proto usiloval především o to, aby na nich bylo možné demonstrovat co nejvíce zde popisovaných jevů. V textech se vyskytují výrazy s individuální referencí specifickou i nespecifickou a se specifickou referencí generickou. Text, který by obsahoval výrazy s nespecifickou generickou referencí, se mi zatím nepodařilo vyhledat, což potvrzuje hypotézu uvedenou v části 6.4, totiž že jmenné fráze s tímto typem reference se budou v komunikaci vyskytovat spíše zřídka.

Pokud jde o metodu analýzy, jako základ mi poslouží způsob anotace koreference v PDT, zohledněny však budou všechny návrhy změn, které byly v této práci vzneseny. Dále se od metody PDT odchýlím v tom, že do koreferenčních vztahů nebudu zahrnovat adjektiva ani slovesa, neboť jejich reference je (jak bylo výše zmíněno) problematická a vyžadovala by důkladnější studium, analyzovanými slovními druhy budou tedy pouze tzv. sémantická substantiva, popsána zde v části 2.2. V PDT probíhá anotace na tektogramatické rovině věty, která obsahuje též uzly zastupující v povrchové podobě nepřítomné výrazy, tyto uzly se tak rovněž účastní koreferenčních vztahů. V této analýze však pro jednoduchost budu z těchto povrchově nepřítomných výrazů zapojovat do koreferenčních vztahů pouze nevyjádřené subjekty jednotlivých vět.^{98,99}

⁹⁸ Případné doplnění též ostatních uzlů reprezentovaných pouze na tektogramatické rovině do koreferenčních řetězců by nemělo představovat žádný zásadní problém, proto si zde dovoluji toto zjednodušení.

⁹⁹ Nevyjádřené subjekty jsou v uvedených textech symbolizovány výrazem [subj].

7.1 Analýza textu č. 1

Text č. 1:

- (1) *Komise rozhodne o střetu zájmů*
- (2) *Praha*
- (3) *Organizační výbor Poslanecké sněmovny by měl ustavit na principu poměrného zastoupení poslaneckých klubů komisi, která by připravila komplexní návrh zákona o střetu zájmů na základě tří předloh upravujících danou problematiku.*
- (4) *Předlohy vypracovali poslanci ODA, ČSSD a vláda.*
- (5) *Usnesl se na tom včera parlamentní výbor petiční, pro lidská práva a národnosti poté, co [subj] přerušil projednávání zmíněných právních úprav.*
- (6) *Podle vládního návrhu by se střet zájmů měl týkat pouze ústavních činitelů, podle návrhů ODA a ČSSD i jejich nejbližších příbuzných a přímých podřízených.*
- (7) *Vládní návrh rovněž nepočítá s úplným zákazem přijímání darů.*

V tomto textu se vyskytuje celkem 9 koreferenčních řetězců. Platí, že v každém z těchto řetězců se všechny členy shodují co do typu reference. Koreferenční řetězce tvořené výrazy s individuální specifickou referencí jsou následující:

- tří předloh (3)¹⁰⁰ – předlohy (4) – zmíněných právních úprav (5)
- ODA (4) – ODA (6)
- ČSSD (4) – ČSSD (6)
- výbor petiční, pro lidská práva a národnosti (5) – [subj] (5)
- vládního návrhu (6) – vládní návrh (7)

Ve všech případech jde o referenci k jasně vymezeným objektům tvořícím součást aktuálního světa, proto jde o specifickou individuální referenci.

Koreferenční řetězce tvořené výrazy se specifickou generickou referencí jsou tyto:

- střetu (1) – střetu (3) – střet (6)
- ústavních činitelů (6) – jejich (6)

¹⁰⁰ Číslo v závorce udává pořadové číslo věty, ve které se příslušná jmenná fráze nachází.

Generičnost příslušných jmenných frází je zde zajišťována jejich přítomností v restriktoru generalizačního operátoru **GEN**. Např. ve spojení *zákon o střetu zájmů* jde o to, že zákon vymezuje vše, co lze označit za (relevantní) střet zájmů. Rovněž v případě spojení *ústavních činitelů*, kterých by se měl zákon týkat, je interpretace taková, že zákon by se měl týkat všech lidí, které lze označit za (relevantní) ústavní činitele.

Výrazy s nespécifickou individuální referencí jsou zastoupeny v těchto koreferenčních řetězcích:

- komise (1) – komisi (3)
- zájmů (1) – zájmů (3) – zájmů (6)

V případě prvního z řetězců jde o nespécifickou individuální referenci s rysem hypotetičnosti, neboť se referuje ke komisi, která v aktuálním světě neexistuje, pouze se připouští možnost jejího budoucího ustavení. Ve druhém řetězci jde o referenci s rysem nevymezenosti, neboť množina zájmů, o které se jedná, nemá vymezeny hranice a zároveň neplatí, že by šlo o referenci generickou, protože existuje nezanedbatelná množina zájmů, které nejsou ve střetu, a proto se jejich spojení *střet zájmů* netýká.

7.2 Analýza textu č. 2

Text č. 2:

- (1) *Pivovar Rakovník koupí český FALL*
- (2) *Rakovník*
- (3) *Jak uvedl včera na tiskové konferenci majitel firmy FALL František Urik, má jeho společnost zájem o koupi stoprocentního podílu na pivovaru Bohemia Rakovník.*
- (4) *Nová společnost ponese název Pivovar Rakovník.*
- (5) *V současné době je uzavřena dohoda o budoucí smlouvě, je složena záloha na koupi rakovnického pivovaru.*
- (6) *Je také určen harmonogram předávání pivovaru.*
- (7) *Od pondělka 20. února probíhají inventury, předávání základních prostředků, nemovitostí apod.*
- (8) *V dohodě je definovaná cena a podmínky, za jakých tento prodej proběhne.*
- (9) *Nic nebrání tomu, aby k prodeji došlo.*
- (10) *Ten by měl být uskutečněn nejpozději do konce března letošního roku.*
- (11) *Cena pivovaru je přijatelná pro obě strany.*

V tomto textu najdeme celkem 5 koreferenčních řetězců, podobně jako v případě předchozího textu se jejich členy navzájem shodují co do typu reference. Řetězce tvořené výrazy s individuální specifickou referencí jsou tyto:

- Pivovar Rakovník (1) – pivovaru Bohemia Rakovník (3) – rakovnického pivovaru (5) – pivovaru (6) – pivovaru (11)
- český FALL (1) – firmy FALL (3) – jeho společnost (3)
- majitel firmy FALL František Urik (3) – jeho (3)
- dohoda o budoucí smlouvě (5) – dohodě (8)

Podobně jako v případě předchozího textu je vymezení koreferenčních řetězců s tímto typem reference poměrně jednoduché. Za zmínku zde stojí výraz *Rakovník* (2), který sice referuje k městu Rakovník, ale nevstupuje do žádného koreferenčního vztahu, protože ostatní výrazy *Rakovník* v textu nereferují k městu, nýbrž jsou názvem pivovaru.

Řetězec s nespécifickou individuální referencí je jeden:

- koupi (3) – koupi (5) – tento prodej (8) – prodeji (9) – ten (10)

Jedná se v případě všech jmenných frází o nespécifickou individuální referenci s rysem hypotetičnosti. Na tomto příkladu je dobře vidět, jaké problémy by byly spojeny s anotací tohoto řetězce při zachování modelu anotace PDT, kde by typ reference závisel na použití určitého identifikátoru. Tak např. výrazy *tento prodej* (8) a *prodeji* (9) by se musely lišit v typu reference, přestože referují k téže entitě a přestože by vypuštění či naopak přidání identifikátoru vůbec neměnilo referenci. Dále je zde možné dokumentovat rozdíl mezi referencí a lexikálním významem, který se pochopitelně neomezuje jen na referenci specifickou. Tak zatímco lexikální význam slov *koupě* a *prodej* je různý (dokonce protikladný), referent příslušných výrazů v analyzovaném textu je totožný.

7.3 Analýza textu č. 3

Text č. 3:

- (1) *Požádá Kubánec Oropesa o politický azyl v USA?*
- (2) *Buffalo*
- (3) *Univerziálním baseballovým turnajem v americkém Buffalu se šíří zpráva, že kubánský nadhazovač Oropesa (23) chce emigrovat do USA.*
- (4) *Krátce před sobotním utkáním proti Tchaj-wanu Oropesa kvapem zdolal řetězovou zábranu na baseballovém stadionu a před zraky stovek fanoušků a policie [subj] nastoupil do čekajícího modrého vozu.*
- (5) *Zástupci kubánského týmu v neděli prohlásili, že žádný z nich neslyšel, že by Oropesa hledal v USA azyl.*
- (6) *Také podle vyjádření Johna Inghama, ředitele místního obvodního úřadu pro emigraci, zatím nikdo v Buffalu o azyl nepožádal, i když Ingham dodal:*
- (7) *„Pokud obdržíme nějakou žádost o azyl, budeme ji pochopitelně držet v tajnosti.“*
- (8) *Šéf kubánského družstva Morales Quintana řekl:*
- (9) *„Nálada v týmu nenasvědčovala, že by něco uvnitř bylo v nepořádku.*
- (10) *Pokud chce Oropesa zůstat v USA, mohl tak učinit jiným způsobem, méně nápadně.“*
- (11) *Podle zprávy amerických pořadatelů je druhým kubánským sportovcem, který se vzdálil od svého družstva, šermíř Perez.*
- (12) *Jde o kubánskou dvojku v šavli.*

Tento text obsahuje 8 koreferenčních řetězců, jejich členy se opět shodují v typu reference. Řetězce s výrazy se specifickou referencí jsou následující:

- Kubánec Oropesa (1) – kubánský nadhazovač Oropesa (3) – Oropesa (4) – [subj] (4) – Oropesa (5) – Oropesa (10)
- USA (1) – USA (3) – USA (5) – USA (10)
- Buffalo (2) – americkém Buffalu (3) – Buffalu (6)
- zástupci kubánského týmu (5) – nich (5)
- kubánského týmu (5) – kubánského družstva (7) – týmu (8)
- šermíř Perez (11) – kubánskou dvojku v šavli (12)

Za pozornost zde stojí poslední z řetězců, neboť je otázka, zda výraz *kubánskou dvojku v šavli* nemá spíše predikační než referenční funkci. Tomu by nasvědčovala významová podobnost konstrukce *jde o x* s konstrukcí *je to x*, v níž má referenční funkci pouze výraz *to*, zatímco *x* má charakter predikační. V případě jednovalenční struktury *jde o x* se však při anotaci v PDT postupuje, jako by zde „*x*“ referenční funkci mělo.

Koreferenční řetězce s nespécifickou individuální referencí:

- politický azyl v USA (1) – azyl (5)
- nějakou žádost o azyl (7) – ji (7)

V případě výrazů *azyl* je patrný rozdíl mezi zde uplatněným přístupem a přístupem PDT. V PDT by pravděpodobně výrazům *azyl* byla připsána reference typu NR a do koreferenčního řetězce by byly zapojeny i výrazy *azyl* v 6. a 7. větě. V navrhovaném pojetí je reference výrazů ve větě (1) a (5) interpretována jako nespécifická individuální, s rysem hypotetičnosti (jde o možný budoucí azyl, který je vymezen svým nositelem i místem). Výrazy *azyl* ve větě (6) a (7) mají také individuální nespécifickou referenci, která však kromě rysu hypotetičnosti má i rys nevymezenosti. Nevstupují ovšem do koreferenčního vztahu, protože v případě realizace příslušných hypotetických dějů by nemuselo jít o totožné množiny azylů. Na rozdíl od přístupu PDT zde tedy nepočítám s takovými koreferenčními vztahy, jejichž oprávněnost je spíše diskutabilní a které jsou tak zdrojem jisté vágnosti při anotaci koreference v PDT.

8. Závěry

Tématem práce byly koreferenční vztahy v češtině a jejich typy. Po představení schématu anotace koreference v Pražském závislostním korpusu a některých problémů, které s tímto schématem souvisí, jsem si kladl za cíl navrhnout takovou typologii reference, která by dokázala nastíněné problematické případy adekvátněji popisovat.

Zásady anotace koreference v Pražském závislostním korpusu byly popsány ve 2. kapitole, důraz byl kladen zejména na uplatňované typy reference a koreference. 3. kapitola následně upozornila na některé problémy spojené s dichotomickým rozdělením reference a koreference na typ 0 a typ NR, které vyplývaly především z příliš široce vymezené kategorie NR a také z nenáležitého důrazu kladeného na formální prostředek určitého identifikátoru.

Na základě odborných prací představených v kapitolách 4. a 5., které se týkaly teorie reference a problematiky generičnosti, jsem dále v kapitole 6. navrhl alternativní schéma referenčních vztahů založené na čtyřech typech reference: specifické individuální referenci, nspecifické individuální referenci, specifické generické referenci a nspecifické generické referenci. Poslední z typů je zmiňován spíše za účelem úplnosti systému, jeho frekvence v českých textech bude pravděpodobně velice nízká. V části 6.6 jsem ukázal, že tímto schématem je možné efektivně popsat i případy, v jejichž popisu model PDT selhává.

V 7. kapitole jsem na konkrétních textech vybraných z Pražského závislostního korpusu demonstroval, jak fungují koreferenční vztahy založené na vymezených typech reference ve skutečných textech. Ukázalo se, že navrhované schéma může přispět k omezení vágnosti, k níž při anotaci koreference v PDT dochází.

Tato práce si nekladla za cíl úplné vyčerpání problematiky reference a koreference, soustředila se především na základní rozvržení jejich typů. Z oblastí, jimž by bylo dále možné věnovat pozornost, považuji za vhodné zmínit zvláště následující:

1. *Problematika reference abstrakt a dějových deverbativ.* Na tuto oblast bylo upozorněno již v části 2.4. Na základě schématu popsaného v 6. kapitole je možné snadno stanovit referenci výrazů označujících konkrétní objekty. Situace u abstrakt a dějových deverbativ je výrazně složitější, avšak i zde je možné rozdělení na uvedené typy předpokládat, vycházet přitom lze např. z valence těchto výrazů.

2. *Koreference dalších slovních druhů.* V této práci jsem schopnost reference (a koreference) přisuzoval pouze výrazům se substantivní platností. O možnosti účasti na koreferenčních vztazích je však možné uvažovat i v případě jiných slovních druhů. Jak bylo uvedeno v části 2.2, v PDT mohou do koreferenčních řetězců za určitých okolností vstupovat též některá adjektiva a slovesa, situace zde však není řešena zcela důsledně.

3. *Reference k reprezentativním objektům druhu.* V části 6.2.2 byla jako jeden z typů interpretace generické reference uvedena reference k reprezentativním objektům druhu (např. ve větě *Na Aljašce jsme natáčeli medvědy grizzly*). Jak bylo uvedeno v části 5.2, jednotlivé jazyky se liší v míře, do níž tuto interpretaci připouštějí. Prozkoumání této možnosti v češtině by bylo žádoucí mimo jiné i kvůli tomu, že v těchto případech není možné použít test generičnosti založený na přítomnosti příslušného výrazu v restriktoru generalizačního operátoru **GEN**.

4. *Vztah reference ke kvantifikaci a distributivnímu přiřazování.* Na operaci kvantifikace a na distributivní přiřazování množin upozorňuje již Hlavsa (1975)¹⁰¹. Distributivně je věnována pozornost též v GenBook (1995). Bylo by užitečné prozkoumat vztah těchto jevů k referenci, zejm. pokud jde o zde navrhované typy reference.

5. *Další referenční vztahy v textu.* V této práci jsem věnoval pozornost pouze vztahu koreference. Bylo by však vhodné zkoumat též další referenční vztahy v textu. Jak již bylo uvedeno,¹⁰² Palek (1988) v této souvislosti zmiňuje vztahy inkluze, členství (v množině) a rozdílu. Některé z dalších referenčních vztahů jsou zachycovány též v Pražském závislostním korpusu v rámci anotace asociační anafory, jde např. o vztah části a celku, vztah množiny a podmnožiny či vztah sémantického a pragmatického kontrastu.¹⁰³ V souvislosti s navrhovaným rozdělením typů reference stojí za pozornost zejména vztah mezi výrazy, z nichž jeden referuje k druhu a druhý k individu, které je tohoto druhu reprezentantem.

Pokud jde o navrhované schéma typů reference, bylo by jistě užitečné ověřit jeho adekvátnost na analýze většího korpusu textů. V této souvislosti by pak bylo možné též prozkoumat frekvenci jednotlivých typů reference a možnosti jejich zapojení do koreferenčních vztahů. V případě některých typů (zejména nespecifické individuální reference) by na základě takové analýzy také bylo možno provést ještě podrobnější klasifikaci. Podobná analýza však již přesahuje možnosti této práce.

¹⁰¹ V této práci se o této problematice zmiňuji v části 4.2.

¹⁰² Viz část 4.1.

¹⁰³ Viz Nedoluzhko (2010).

Literatura

ADAMEC, P. K vyjadřování referenční určenosti v češtině a v ruštině. *Slovo a slovesnost*, 1980, roč. 41, s. 257–264.

CARLSON, G. N.; PELLETIER, F. J. *The Generic Book*. Chicago/London : The University of Chicago Press, 1995.

HLAVSA, Z. *Denotace objektu a její prostředky v současné češtině*. Praha : Academia, 1975.

KOMÁREK, M.; KOŘENSKÝ, J.; PETR, J. et al. *Mluvnice češtiny 2 : Tvarosloví*. Praha : Academia 1986.

MIKULOVÁ, M.; BÉMOVÁ, A.; HAJIČ, J. et al. *Anotace na tektogramatické rovině Pražského závislostního korpusu : Anotátorská příručka*. Praha : MFF UK, 2005.

NEDOLUZHKO, A. *Rozšířená textová koreference a asociační anafora : Koncepce anotace českých dat v Pražském závislostním korpusu*. Praha : Filozofická fakulta UK v Praze, 2010. Nepublikovaná disertační práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav českého jazyka a teorie komunikace.

PALEK, B. *Referenční výstavba textu*. Praha : UK, 1988.