

## Kapitola 2. Text

### 2.1 Co je to text?

Koncept textu je v sociálních vědách jedním z nejpoužívanějších a zároveň se jen ojediněle stává předmětem historické nebo konceptuální reflexe. Do konce 60. let minulého století bylo používání slova text omezené prakticky jen na filologické a bibliografické obory. Poté však došlo k dramatickému nárůstu jeho použití, zejména ve dvou významech: a) termínem *text* se začala označovat literární díla místo dosavadního termínu *dílo*; b) koncept *text* začal univerzálně zahrnovat všechny psané a tištěné jazykové nebo parajazykové (jazyk připomínající) projevy v nejrůznějších podobách [Wilson 2012].

Toto nové používání nemá jasnou návaznost na předchozí používání termínu *text*; podle Wilsona [2012] se zdá, že rozhodující pro jeho rozšíření v západním světě byly eseje Jacquese Derridy a Rolanda Barthesa. Derrida ve své významné knížce *O gramatologii* [Derrida 1967] formuloval maximu „není nic mimo text“ a Barthes v článku „Od díla k textu“ [Barthes 1971] zavedl nový koncept „textu“ na místo předcházejícího „díla“. Toto nové používání termínu *text* se stalo všeobecně rozšířené v 80. letech minulého století, kdy se objevuje ještě příbuzný termín „textovost“ (textuality). Význam konceptu se dále rozšiřuje prakticky na každý kulturní objekt – Paul Ricoeur tuto pozici formuloval v článku „Model textu: smysluplné jednání jako text“ [Ricoeur 1973]; v jiném kontextu viz Lotman a Uspensky [Lotman, Uspensky 1978].

Podle Hankse [Hanks 1989] reagovali sociální vědci na tuto pojmovou expanzi („všechno je text“) několika způsoby. Buď s texty (záznamy verbálních aktivit) i nadále zacházeli tradičně jako s víceméně neproblematickými objekty vhodnými k analýze a interpretaci, nebo omezili význam textů pro vysvětlení sociálního dění a snaží se využívat jiná data než texty (behaviorální orientace), nebo přijali rozšíření definice textu na veškeré kulturní artefakty, na lidský svět jako sémiotický systém.

Vraťme se ale k pojmu *text*. V rámci lingvistiky existuje množství definic textu od těch zaměřujících se vnitřní povahu textu až po ty, které považují texty za koreláty komunikace (pro přehled viz [Helbig 1991: 139–141]).

V širším sociálněvědním rámci bývá text definován pragmaticky jako takový objekt, který lze číst, přičemž čtení je míněno v širokém smyslu slova jako aktivita, jejímž výsledkem je porozumění [Wilson 2012]. To znamená, že textovost je kvalita, která je nějakému objektu přisouzena člověkem, přesněji čtenářem nebo kulturním kontextem. Jako text tak lze označit psaný dokument, kreslený komiks, film, hudební skladbu, gestikulaci nebo krajinu; za podmínky, že pro ně nalezneme možnost jejich „čtení“. Texty mají, kromě potenciálu být čteny, ještě jednu vlastnost – abstraktnost. Ačkoliv text vnímáme jako materiální záznam (písmo na papíře, hlas, pohyby určitých objektů apod.), tzn. jako vlastnost materiálního prostředí, přesto to, co vytváří text, je možné z materiality abstrahovat a reartikulovat [Wilson 2012]. Když si přečteme báseň, měli bychom ji být schopni i recitovat, opsat nebo ji rozpoznat, pokud ji najdeme jako součást jiného textu. Tento text píšu s pomocí klávesnice, vidím ho na monitoru, ale někteří čtenáři ho budou číst na papíře – a vždy půjde o jeden a tentýž text. Identitu textu tedy nehledáme v jeho materiální podobě, ale v podobě abstraktní. Na druhou stranu význam textu, jeho interpretace čtenářem, bývá materiálním prostředím textu (kontextem) vždy ovlivněna, jak si ukážeme dále. Shrnu-li to, tak textem se rozumí jakýkoliv objekt nebo soubor objektů, který lze číst, to znamená z jejich materiality abstrahovat význam, který může být znovu reartikulován, tj. převeden do materiální podoby. V této knize se ovšem budeme zabývat jen verbálními texty jako nejčastějším typem textových dat používaných ve společenských vědách.

### 2.2 Psané a mluvené texty

Jednou ze zajímavých otázek, která bývá v souvislosti s analýzou textů diskutována, je rozdíl mezi mluvenými a psanými texty [Čermák 2007: 36–39; Čmejrková, Daneš, Havlová 1994; Čmejrková, Hoffmannová 2011: 19–60]. Nemám zde na mysli lingvistiku často studovaný rozdíl mezi mluvenou (hovorovou) a psanou (spisovnou) řečí, tj. že mluvíme jinak, než píšeme, nebo mezi fonetickým a psaným záznamem (např. říkáme *gde, ftom, ňák*, ale píšeme *kde, v tom, nějak*), ale rozdíl mezi:

- promluvou, to znamená řečovým jednáním, které se odehrává v čase a prostoru a účastní se ho alespoň jeden mluvčí a jeden posluchač;
- zaznamenanou promluvou, tzn. audiovizuálním, zvukovým nebo psaným záznamem řečového jednání, který není závislý na okamžiku promluvy a lze jej reprodukovat na jiných místech a v jiný čas;
- psaným textem, tzn. řečovým jednáním prostřednictvím písma, které proběhlo za účelem jeho čtení v jiný okamžik než v době jeho vzniku

Podle některých autorů není promluva textem [Ricoeur 1973]; pro jiné textem je, protože je to materiální objekt, z něhož lze text abstrahovat [Brown, Yule 1983]. Například když A řekne „Dobry den!“ a B špatně slyšel a nerozuměl, co A řekl, může se zeptat C, co A řekl. C odpoví, že A řekl „Dobry den!“. „Dobry den“ je v tomto případě text, který byl C abstrahován z promluvy, kterou udělal A, a B abstrahován z promluvy, kterou udělal C.

Důležitost tohoto rozdílu spočívá v odlišném přístupu k textu a interakci s ním. Jak píše P. Ricoeur [1973: 91], v mluveném textu v situaci tváří v tvář se snažíme porozumět druhému, to znamená poznat, co druhý subjekt míní nebo zamýšlí na základě všech druhů znaků, kterými se psychický život projevuje. Porozumění v tomto smyslu předpokládá intersubjektivitu, schopnost v daný okamžik myšlenkově zaujmout místo toho druhého a podívat se na situaci z jeho pozice. V případě psaných textů, a může jít o přepisy promluv nebo psané texty, kde nejsme s jeho autorem v situaci tváří v tvář, se naše *interpretace* omezuje na kategorii znaků, které jsou fixovány psaním. K textu přistupujeme jako k záznamu a s jeho autorem nejsme v bezprostředním intersubjektivním vztahu; náš vztah je zprostředkován textem, ať už čteným námi samými nebo někým jiným.

## 2.2.1 Odlišnosti mluveného a psaného textu

Ricoeur [1973: 92–97] uvádí čtyři rozdíly mezi promluvou a psaným textem z hlediska komunikace významu:

1. Událost mluvené řeči (promluva) je na rozdíl od psaného textu vždy realizována v určitý časový okamžik, a pokud se jí účastníme, tak vždy v přítomnosti; je-li promluva zachycena psaným textem, pak text nekonzervuje událost promluvy, ale to, co bylo řečeno, tj. význam řečové události, nikoli událost samu. Dokládají to úvahy J. L. Austina a J. R. Searla [Austin 2000/1962; Searle 1969], podle nichž sestává řeč z mluvních aktů (promluv) vykazujících tři aspekty nebo úrovně: lokuční, ilokuční a perlokuční. Vezměme například promluvu konferenciéra k jedné ze soutěžících při příležitosti vyhlášení výsledků soutěže: „Prohlašuji vás za vítězku naší soutěže.“ *Lokuční* úroveň spočívá v pronesení smysluplné promluvy, která je správná foneticky i gramaticky a má určitý význam. *Ilokuční* úroveň spočívá ve významech, které z promluvy plynou; nejčastěji to jsou ilokuční akty rady, rozkazu nebo žádosti; v příkladu uvedeném výše je zřejmé, že se nejedná o pouhé konstataování, ale na adresátku vyvíjí sdělení v dané situaci určitý tlak, jenž posiluje sdělovaný význam, například mluvčí doprovodí svá slova úsměvem a napřažením ruky ke gratulaci a od adresátky promluvy se očekává, že gratulaci při-

jme (bylo by překvapivé, kdyby na ni nijak nereagovala); proto se někdy mluví o „ilokuční síle“, protože je to právě efekt tlaku, který z lokučního aktu činí akt ilokuční. *Perlokuční* úroveň spočívá v tom, že mluvení má nejen sílu nás ovlivnit, ale že jím i něco děláme; v našem příkladu není hlavní význam promluvy ve sdělení (informaci), nýbrž v tom, že se jejím prostřednictvím něco udělalo (tím, že byl někdo prohlášen za vítěze soutěže, se jím také skutečně stal a je možné mu předat odměnu). I když se můžeme setkat s mluvními akty, které mají pouze lokuční úroveň (oznámení: „Včera přes Českou republiku přecházela studená fronta.“), ve většině běžných promluv lze identifikovat současně ilokuci (ironicky: „To ti teda budu věřit.“) a často i perlokuci, třeba ve formě zdvořilostních ohledů na postavení adresáta sdělení (pozdrav: „Dobry den, pane doktore.“) [srovnej Helbig 1991: 159–193; Iser 2009: 146–151].

Jednotlivé úrovně promluv jsou podle Ricoeura v různé míře zachytitelné v psané podobě. Nejúplněji lze převést do psané podoby lokuční úroveň; psaným textem se promluva stane, pokud je převedena do věty nebo vět. Psaná věta může být identifikována a reidentifikována jako tatáž věta nesoucí stejný význam. Ilokuční úroveň (gesta, mimika, prózodie) je méně zachytitelná v psaném záznamu řeči než úroveň lokuční; může být reprezentována textovými prostředky (otazníkem, vykřičníkem, třemi tečkami apod.). Ještě obtížněji je zapsatelný perlokuční aspekt promluvy, který je podle Ricoeura typický pro mluvenou řeč. I když bychom mohli namítnout, že existují významné výjimky, jakými jsou písemná jmenování, rozhodnutí nebo rozsudky, jejichž perlokuční aspekt se psanou podobou nijak nevytrácí. Nicméně Ricoeur má pravdu v tom, že perlokuční úroveň je „to nejméně mluvené v promluvě,“ je to redukce promluvy na určitý stimul, který působí nikoli prostřednictvím porozumění subjektivnímu záměru komunikačního partnera, ale energeticky, přímým působením na svět, ve kterém byla promluva pronesena (jeho efekt lze přirovnat k pohlazení, pohlavku, připnutí ocenění nebo jeho stržení apod.).

Souhrnně lze říci, že význam promluvy není pouze v propoziční úrovni, ale rovněž i v ilokuční síle a perlokučním aktu; tento význam lze v různé míře zachytit v psaném textu, tzn. textově kodifikovat tak, že pomocí této kodifikace může být text identifikován a reidentifikován jako mající stejný význam.

2. Zatímco jazyk postrádá subjekt – v tom smyslu, že otázka „Kdo mluví?“ postrádá smysl [Ricoeur 1973] – promluva vždy odkazuje ke svému mluvčímu nejrůznějšími prostředky, například osobními zájmeny. Když někdo mluví, nelze rozlišit význam promluvy od záměru (intention) mluvčího, respektive na základě úsudku o komunikačním záměru mluvčího stanovujeme význam řečeného. V psaném textu se autorův záměr a význam textu přestá-

vají překrývat; pokud čteme nějaký text, už nemůžeme jednoduše uvažovat, co nám autor sděluje, ale co nám text sděluje. Při zapsání promluvy dochází k oddělení subjektivních záměrů a verbálního významu textu. Toto oddělení nepopírá ani fakt, že v běžných situacích může docházet k ambivalentním situacím, například když někdo řekne: „Co tím chceš říct?“ Mluví zde odděluje verbální význam od subjektivního záměru, aby poukázal na to, že jen těžko si dovede představit, že verbální význam promluvy je identický se subjektivním komunikačním záměrem. Na druhé straně si lze představit psaný text, například osobní korespondenci, která je čtena, jako kdyby autor promlouval (což bývá někdy ve filmech zdůrazněno tím, že text je čten nahlas hlasem autora dopisu). Tyto hraniční situace nám ukazují, že většina promluv (které nejsou čtením) spojuje mluvčího s významem výpovědi (textu), zatímco v psaných textech je význam záležitostí širší než jen to, co chtěl (mínil) autor říci; každý výklad může vytvářet jeho význam bez nutnosti brát v úvahu subjektivní (psychické) procesy autora.

3. Promluva (ať už zapsaná, nebo mluvená) je podle Ricoeura vždy o něčem; odkazuje na svět, který si nárokuje popsat, vyjádřit nebo reprezentovat. V promluvě je uskutečňována symbolická funkce jazyka, který sám, jako uzavřený systém, postrádá jakýkoli referent (přesněji: je autoreferenční). V orální promluvě je reference ostenzivní, zjevná; promluva odkazuje na situaci, kterou mluvčí i posluchač sdílejí: „Prohlašuji vás za vítězku naší soutěže.“/„Co tím chceš říct?“ Psaný text to však nedokáže. Autor píšící text a čtenář čtoucí tentýž text nejsou ve stejné situaci a proto psaný text nemůže pracovat s ostenzivní referencí; nestačí napsat „tam, kam ukazují“ bez toho, aby bylo v textu rovněž uvedeno, kam se ukazuje. Svět, ke kterému psaný text odkazuje, je prakticky světem čtenáře, nikoli autora. Ricoeur jde ještě dál, když tvrdí, že svět je pro nás souhrnem referencí otevřených přečtenými texty; teprve psaný text nám umožňuje pochopit vlastní situaci, naše postavení ve světě nezávisle na subjektivitě autora a jeho ostenzivních tady a teď [Ricoeur 1973: 96]. I zde bychom ale mohli nalézt hraniční případy, se kterými Ricoeur nepočítá; například videozáznam vykazuje jak některé rysy mluveného textu, tak rysy textu psaného. Ostenzivní reference může ve videozáznamu velice dobře fungovat; význam promluvy „Podívejte se na to, vždyť jste mi to auto úplně zničil!“ lze velmi dobře interpretovat podle toho, jak poškozený vůz vypadá. Na druhé straně z vlastní zkušenosti víme, jak jinak vnímáme situaci, které jsme byli svědky, když ji vidíme na videozáznamu. Často máme sklon záznam komentovat, doplnit, co tam chybí, nebo naopak potlačit, co je příliš výrazné. Každý videozáznam je totiž činěn z nějaké pozice, která nemusí odpovídat té, v níž byl posluchač a z níž promluvu vnímal.

4. Čtvrtý rozdíl mezi promluvou a psaným textem se projevuje v roli adresáta. Promluva je adresována bezprostředně přítomným posluchačům, psaný text je adresován publiku, které si sám vytváří. Kdokoli se může stát adresátem textu, pokud ho dokáže přečíst [Ricoeur 1973: 97]<sup>3</sup>. Zde bychom mohli opět poukázat na případy, ve kterých se vlastnosti mluveného a psaného textu prolínají: čtený text, audiozáznam nebo videozáznam mluveného textu. V prvním případě je adresátem současně čtenář implikovaný textem (např. přítomní v soudní síni i všichni ostatní, kterým se text rozsudku dostane do ruky), ale i posluchač přítomný aktuálnímu čtení (přítomní v soudní síni). Podobně v audiovizuálním záznamu konverzace jsou adresáti promluv aktéři v původní situaci, ale rovněž i všichni, kdo se na záznam dívají; i ti se mohou cítit rozveseleni vyprávěným vtípem nebo dotčení urážlivou poznámkou, jíž se cítí být adresáti.

5. Konečně posledním rysem, který sice Ricoeur neuvádí, ale který významně odlišuje promluvu od psaného textu, je linearita. Když někdo promlouvá, tok mluvy jde pouze jedním směrem, a to i v případě, kdy nemluví (ticho, resp. mlčení, je důležitá komunikační strategie). V mluvené řeči se nelze vrátit zpět, lze pouze ještě jednou zopakovat, co bylo řečeno, což není totéž. Platí to i pro posluchače, který, přeslechnel-li část promluvy, už nemá možnost si ji znovu poslechnout. Psaný text není takto lineární jako promluva, bez ohledu na to, že k němu často jako k lineárnímu přistupujeme. Když píšu tuto větu, ještě jsem nenapsal všechny věty, které jí budou v tomto textu předcházet; podobně i čtenář může začít knihu číst až od této věty a na začátek se vrátit až později. V psaném textu se lze vracet nebo číst opakovaně.

Shrnu-li rozdíly mezi promluvou a psaným textem, tak k obojímu lze přistupovat jako k textu. Nicméně promluva se od psaného textu odlišuje v několika aspektech, které z těchto dvou forem textů mohou (i když nemusí) vytvářet dva odlišné předměty analýzy. Mluvený text je bohatší svou ilokuční a perlokuční silou, poskytuje větší možnosti, jak výpověďmi něco dělat. Zároveň je úžeji vázán a) na mluvčího, jehož domnělé záměry jsou pro posluchače rozhodující pro porozumění významu promluvy, b) na situaci, ve které k promluvě dochází, prostřednictvím ostenzivního odkazování ke světu, c) na posluchače, který je promluvě simultánně účasten, d) na čas, protože promluva nezadržitelně plyne s časem a není možné znovu identicky aktualizovat již vyslovená slova.

<sup>3</sup> Ricoeur zde má pochopitelně na mysli lidského čtenáře, ale platí to i pro strojové čtení. Můžeme promlouvat k nějakému předmětu, který je v dané situaci přímým adresátem naší komunikace. Jakmile je tato naše promluva přepsána, přímý adresát se ztrácí a promluva může být předmětem psychiatrické analýzy prováděné lidským čtenářem nebo s pomocí strojového čtení.

## 2.2.2 Text jako záznam promluvy

Abychom k promluvě mohli přistupovat jako k textu, musíme ji zaznamenat. Nejpřirozenější záznam je paměť a také ji nejčastěji používáme. Paměť však zpravidla nestačí jako záznam promluvy pro odbornou analýzu, protože je nekontrolovaně výběrová; pamatujeme si jen to, co nám „utkví“ v paměti, aniž bychom mnohdy věděli, proč zrovna to a ne něco jiného. Proto při textové analýze pracujeme s technickými záznamy promluv. Audiovizuální záznam promluvy reprezentuje událost promluvy pro abstrakci textu v jeho efektivním kontextu nejlépe, ale i tak pouze částečně. V každém případě je tento záznam už určitou interpretací události, jednak perspektivou pohledu, jednak redukovanou reprodukcí kontextu a jednak – víme-li, že promluva už proběhla a jedná se o záznam – i neautentičností znovupřehrané události.

Ještě výraznější interpretací promluvy je její písemný přepis, se kterým se můžeme v sociálněvědním výzkumu setkat nejčastěji. Existují způsoby záznamu, kdy je maximum textových i kontextových skutečností promluvy zaznamenáno, nicméně při přepisu se zcela ztrácí barva a přirozená síla hlasu, jemné odstíny v intonaci a naproti tomu je zvýrazněna textová složka promluvy (co bylo řečeno). Ve většině analýz nám však jednoduchý přepis postačuje, protože verbální složka promluvy je to, co považujeme za data, a situační kontext výpovědi nejsme schopni analyticky uchopit. Ke vztahu promluv, jejich přepisů a psaných textů se ještě dostaneme v dalších kapitolách při výkladu jednotlivých přístupů a metod.

## 2.3 Transtextové vztahy

Při charakteristice textu byla uvedena abstraktnost jako jeho podstatná vlastnost. Nyní se podíváme na to, jaké důsledky má abstraktnost na vztah textu k jiným textům a k netextovému prostředí. G. Genette rozlišuje několik základních vztahů, které textovost umožňuje a které strukturují univerzum existujících textů. Obecně nazývá soubor těchto vztahů transtextovostí a definuje ji povšechně jako „vše, co dává text do vztahu (zjevně nebo skrytě) s jinými texty“ [Genette 1982: 7]. Jednotlivými specifickými transtextovými vztahy jsou intertextovost, paratextovost, metatextovost a hypertextovost.

*Intertextovost* označuje jev, kdy jsou texty v sobě navzájem přítomné [Bazerman 2004; Genette 1982: 8; Slembrouck 2011]. Může se to díť odkazy (např. nyní odkazují na Genettův text a činím ho přítomným v tomto textu) nebo doslovnými citacemi, které mohou být přiznané nebo nepřiznané (např.

plagiátorství), nebo i nejrůznějšími náznaky realizovanými volbou slov, souloví, slovních vazeb, délkou vět apod. Pro textové analýzy má intertextovost veliký význam například při hledání „zdrojového“ textu pro různá životní vyprávění, která na něj skrytě odkazují; známým příkladem „úspěšných životních příběhů“ je příběh amerického selfmademanství; jiným příkladem mohou být mediální texty nebo odborné texty, ve kterých je míra intertextovosti (vzájemného odkazování) tak vysoká, že vede některé badatele k domněnkám, že jde o sebereferenční intertextovou realitu.

*Paratext* je nejčastější v psaných textech, v nichž vytváří výpravu a prostředí textu, tzn. název, podnázev, mezititulky, obsah, rejstříky, seznamy literatury, bibliografické záznamy, předmluvy, dovětky, dodatky, protokoly, přílohy apod. Paratexty jsou pro textové analýzy výzvou, protože je vnímáme jako součást textu, ale na druhou stranu stojí jaksí vedle něj – otázka je, jsou-li textem, kotextem nebo kontextem (viz dále). Při analýze rozhovoru publikovaném v týdeníku, který je opatřen úvodem tazatele o tom, jak se rozhovor seběhl, dále pak redakčními mezititulky a rámečky, ve kterých je uvedena krátká atraktivní či provokativní citace z rozhovoru, jež se objevila i na titulní stránce, musíme tyto paratexty vzít v úvahu, protože rovněž čtenář je bere v úvahu, ale zároveň s nimi nemůžeme nakládat jako se součástí rozhovoru. Kromě paratextu existují paralingvistické atributy textu, které můžeme identifikovat prakticky v každém textu, například uspořádání řádků umožňuje pochopit, že jde o konverzaci nebo o báseň; tyto atributy textu tak vytvářejí předpoklady i pokyny pro své vlastní čtení a interpretaci.

Termínem *metatext* označuje Genette komentář, text reagující na jiný text, mluvící za jiný text, někdy dokonce bez nutnosti ho zmínit nebo citovat; „diskurz, který odkazuje, popisuje nebo rámuje interpretaci textu“ [Hanks 1989: 96]. Prototypickým metatextem je kritika. Specifickým druhem metatextu jsou databáze textů, citační indexy, které zvláště ve vědě nabývají velkého významu a bezprostředně ovlivňují tvorbu textů, například abstraktů (pisatel je instruován, aby abstrakt napsal tak, aby byl vhodným textem pro automatické generování klíčových slov a jiných metatextových strategií). Svého druhu metatextem jsou výsledky vyhledávání na internetu, které se mohou stát předmětem analýzy.

Spolu s internetem se dostáváme ke kategorii *hypertextu*. Genette definuje hypertext jako text založený na nějakém předchozím textu (hypotextu), který ale není jeho komentářem. Podle něj jsou všechny texty v určité – větší nebo menší – míře hypertexty, protože jejich tvůrci ať nápodobou nebo transformací při tvorbě textu z již existujících textů vycházejí. Základními třídami hypertextovosti jsou parodie, travestie nebo pastiš; Genette nazývá texty s vysokou mírou hypertextovosti literaturou „druhého stupně“.

Koncept hypertextu získal ještě další význam v souvislosti s internetem a World Wide Web prostředím. I když bychom mohli hledat podobnosti s Genettovým vymezením pojmu, zvláště z toho důvodu, že se jedná o transtextový vztah, jednodušší bude chápat ho v této souvislosti zcela jinak. Elektronický hypertext je určitý „nadtext“, který umožňuje čtení neomezené jedním textem. Obsahuje totiž hypertextové odkazy na jiné texty, které čtenář může při čtení použít a následně ihned začít číst jiný text, na nějž bylo odkazováno. Hypertext tak umožňuje vytvořit hustou strukturu vzájemně propojených textů, která vybízí k nelineárnímu a promiskuitnímu čtení (Facebook, Twitter). Bylo by možné říci, že www hypertext působí jako intertext, paratext a metatext současně. Převratnost elektronického hypertextu však není principiální; již před hypertextem bylo možné psát texty s odkazy, které, pokud čtenář disponoval dostatečně velkou knihovnou nebo mediátékou, vybízely k nelineárnímu čtení; vědecké texty plné odkazů a čtení v prostředí rozsáhlé knihovny lze jistě ze předchůdce internetového hypertextu považovat.

Pátým aspektem textovosti je *architextovost*, příslušnost k určité třídě textů (báseň, román, kronika, vyprávění, zpráva, rozhovor, zpověď atd.). Architext se podobá žánrovému *eidos*, esenci určité třídy textů. Architextovost má nejbližší k textovosti samotné (určuje, zda je objekt před námi text, a pokud ano, o jaký druh textu se jedná) a podobně jako ona je architextový status nejčastěji definovaný čtenářem; tím bývá ovšem také textový analytik a úvahy o třídění textů pro analýzu se často o více či méně fundované architextové soudy opírají.

Vedle přehledu transtextových vztahů je třeba popsat ještě dva vztahy: kotext a kontext. Jejich vymezení je v literatuře různé, zejména co se týká pojmu kontext [srov. Brown, Yule 1983: 35–58; Eco 2010: 25–30; van Dijk 1977: 191–195; Widdowson 2004: 36–73]). Jelikož není možné dostát všem přístupům, zvolil jsem „jednoduché“ řešení, spočívající v rozlišení do jakého sémiotického systému daný transtextový prvek patří.

*Kotext* je okolí textu patřící do téhož sémiotického systému (např. jazykového). *Kotext* tak může nabývat textové, paratextové, metatextové, intertextové nebo i hypertextové kvality a jeho identifikace závisí na čtenáři, který při čtení (poslechu) textu činí kotextové selekce. Pokud začneme transtextovými kotexty, například titulek novinového článku nebo text „Převzato z databáze ČTK“ jsou vzhledem k článku paratextovým kotextem; citace jiného autora (klasika) uvozující text, jak to mívá v oblibě romantická literatura, je intertextový kotext. Promluva „Tak já zapnu to nahrávání a zkusím, jestli to dobře nahrává. Mohl byste něco říci, abychom zkusili, jestli jste slyšet?“ je metatextový kotext rozhovoru. Při textové analýze jsou kotexty důležitou slož-

kou textových dat, například datum vydání u novinových článků. Naprostá většina kotextu je ovšem tvořena bezprostředním textovým okolím místa v textu, které je čtenářem, resp. posluchačem, právě čteno, resp. slyšeno, a umožňuje čtenáři efektivní porozumění textu. Z hlediska rozsahu můžeme rozlišovat kotext na úrovni slov, vět a nadvětných celků (odstavců, kapitol, replik apod.). Když vidíme na zdi napsáno „Radku já tě miluju,“ tak kotextem slova „miluju“ budou pro nás nejspíše slova „Radku já tě,“ která nám umožňují text smysluplně interpretovat jako milostný vzkaz. Slovo „VOLE“ napsané vedle jinou barvou pro nás žádný bezprostředně relevantní kotext mít pravděpodobně nebude, ale i tak je interpretovatelné na základě širšího transtextového kotextu jako nadávka; nebo případně textová interpretace není relevantní, protože se jedná o „poškození cizího majetku“ a kotext bude hledán v právním sémiotickém systému (přestupek, trestný čin apod.).

*Kontext* je na rozdíl od kotextu takové okolí textu, které nepatří do téhož sémiotického systému; možný kontext promluvy tvoří nejazykové prostředí, například gestikulace, podoba mluvčího, síla hlasu apod.; kontext psaného textu je grafické zpracování, obrázky doprovázející text, podoba média, které text materializuje apod. Podobně jako kotext i kontext je vztahem, který ovlivňuje interpretaci textu jako projevu komunikace. V našem výše uvedeném případě kontextem řečové výpovědi „Radku já tě miluju“ může být konkrétní zeď, na které je napsána, způsob, jakým je napsána, přítomnost čtenáře a dalších osob, kteří se mohou, ale nemusí cítit být adresáty atd. Jak poznamenává van Dijk [1977: 192], kontext je dynamický jev, je to běh událostí (course of events). A to nejen povrchově ve smyslu, že kontext se mění tím, jak se mění prostředí textu, ale i hloubkově, protože samotný čtenář se poslechem nebo čtením textu proměňuje (v uvedeném případě se kontext změní, když po přečtení dojdeme k závěru, že víme, kdo je autorem textu nebo jeho adresátem). Proto je nutné rozlišovat *možné kontexty* a *aktuální kontext*, tj. kontextové selekce. Při interpretaci textu bereme v úvahu zejména aktuální kontext, ale přihlížíme také k dalším možným kontextům, které připadají v úvahu; aktuální kontext je totiž vždy čtenářskou/posluchačskou selekcí z univerza možných kontextů. Při této příležitosti znovu připomínám, že i když by to mohlo tak vypadat, kontext není určován textem nebo situací, ve které se text nachází, ale rozhodující je vždy čtenář/posluchač, který pragmaticky z dostupných sémiotických zdrojů vytváří kontext za účelem interpretace textu. Pro zachycení tohoto dynamického procesu přiřazování kontextů se používá pojem kontextualizace [Gumperz 1982].

Role kontextu je významná pro všechny texty, ale pro některé texty je důležitější než pro jiné. Zvláště důležitý je kontext pro výpovědi, které obsahují deiktická slova jako *zde, táhle, teď, potom, včera, tohle, tenhle, ona, je-*

jich apod., protože teprve jejich aktuální (posluchačem přisouzený) kontext jim dává konkrétní význam. Výrazně kontextově závislé jsou též performativní výpovědi jako výše uvedený příklad „Radku já tě miluju,“ ale též sliby, rady, příkazy apod., kde je pro jejich porozumění důležité, kdo, kdy, komu a v jaké situaci je dává. Konečně zmíním hodnotící výpovědi, obsahující slova *správný, normální, lepší, to je k neuvěření, rozhodně, moc málo, moc silně, příliš nízký* apod., kde je kontext aktualizovaný čtenářem/posluchačem z univerza možných kontextů pro interpretaci rovněž klíčový.

Základní kotextové a kontextové vztahy zobrazuje tabulka 1:

**Tabulka 1** Základní kotextové a kontextové vztahy

	sémiotický systém	
okolí textu	stejný jako text	jiný než text
možné	kotext	kontext
aktuální	kotextová selekce	kontextová selekce

Vzhledem k výše uvedenému, tj. že význam komunikace je difuzně a dynamicky situován v interakci text (promluva) – posluchač/čtenář – transtextové vztahy, jsou tyto koncepty pro textovou analýzu navýsost důležité. Analytik musí zvážit, zda texty, které se chystá studovat, nepředpokládají určité významné transtextové vztahy a zda jsou tyto empiricky dostupné a analyticky postižitelné. Je evidentní, že nelze zachytit všechny možné vztahy zkoumaného textu nebo textů k jiným textům a kontextům. Nicméně při určitém metodologickém důmyslu lze postihnout frekventované kontextuální vztahy nebo ty vztahy, které jsou institucionálně či situačně preferované. Například když někdo vypráví svůj životní příběh a svému konverzačnímu protějšku říká: „Já jsem ten šedesátý osmý prožil na vojně, mě se to moc netýkalo.“, tak se nabízejí dva kontexty, ve kterých dává výpověď smysl: historický a identitní [Hájek, Havlík, Nekvapil 2014]. Nebo když zjistíme, že v mediálních textech feministických organizací a skupin se slova násilí a pornografie nepojí s žádnými dalšími frekventovanými slovy [Hájek, Kabele, Vojtíšková 2006a], můžeme se domnívat, že tato slova jsou používána ve variabilním kotextu, který indikuje i různé kontexty (např. publika, čtenáře), ve kterých se významy pornografie a násilí v komunikaci objevují.

Tento oddíl uzavřeme konceptem *hranice textu*. Otázka hranic je z našeho pohledu – na rozdíl od pohledu lingvistického nebo literárněvědního<sup>4</sup> – tech-

<sup>4</sup> Zajímavým případem jsou začátky (i konce) historiografických vyprávění, například jak začínají dějiny Francouzské revoluce nebo druhé světové války [srov. Richardson 2008].

nickou, nikoli teoretickou záležitostí. Při analýzách promluv a textů jsme často nuceni „vyříznout“ text z proudu komunikačního dění, a to vyvolává otázku, zda hledat nějaké „přirozené“ hranice textu nebo zda konec textu definovat arbitrárně. Text může mít určité autodefinované hranice, například věty „Byl jednou jeden... Tady příběh končí.“ nebo „Tak to je všechno, víc vám k tomu asi neřeknu.“ – „Děkuji za rozhovor.“ Hranice textu může být daná paratextově, například velkou kapitálkou na začátku a slovem KONEC za textem. Konec může být i určen kontextově, například grafickým symbolem nebo intonací. Jestliže text chápeme jako vlastnost přidělovanou nějakému objektu čtenářem, pak text končí tam, kde končí jeho čtení – ostatní je v dané situaci kontext. Textem tak může být jedno heslo v encyklopedii nebo celá encyklopedie nebo dokonce celá knihovna, pokud se někdo rozhodne, že ji celou přečte. V analýzách vycházíme nejčastěji z hranic textu definovaných paratextově (interpunkčními znaménky, odstavci, články, webovou stránkou) nebo kontextově, že začíná záznam rozhovoru, respondent ukončí hovor nebo končí přepis rozhovoru. V některých případech se však uchylujeme k arbitrárnímu rozhodnutí a text vymezujeme sémanticky (jako z našeho čtenářského pohledu významovou jednotku) nebo svévolně určeným rozsahem při studiu spoluvýskytů slov.

## 2.4 Čtení textů

Jak je z předchozích odstavců zřejmé, role čtenáře je pro textovou analýzu klíčová – čtenář přisuzuje objektu vlastnosti textu, čtenář vtahuje text do transtextových a kontextových vztahů, čtenář vytváří význam textu. To neznamená, že samotný text je ničím, to nikoli; znamená to však, že text je substrátem čtenářovy aktivity, podobně jako je půda substrátem rolníkovy práce; zda dá půda úrodu, čeho a jakou záleží z velké části na rolníkově práci, podobně jako význam textu je spoluvytvářen čtenářem.

Správně by se tento oddíl měl nazývat „Poslouchání a čtení textů“, protože texty jsou mluvené i psané. Dokonce většinu textů v životě přijímáme jako promluvy, tedy posloucháním, nikoli čtením. Mohlo by se zdát proto překvapivé, že daleko více se odborné zkoumání role recepce textů věnuje čtení textů a interakce čtenáře s textem; je to dáno pravděpodobně jednak tím, že otázky recepce textů pocházejí z pole literární vědy, která má za svůj předmět především texty psané, a jednak i tím, že v případě mluvených textů se pozornost zaměřuje na vzájemnou interakci mluvčích/posluchačů, tzn. předmětem analýzy je řečová interakce mezi lidmi a nikoli mezi lidmi a texty. Položme si proto otázku: existují nějaké rozdíly mezi posloucháním a čte-

ním textů ve vztahu k porozumění jejich významu? V oddíle věnovaném rozdílu mezi psanými a mluvenými texty jsem, vycházející z Ricoeura [1973], uvedl několik rozdílů mezi promluvou jako událostí a textem jako abstrakcí. Ricoeur se však jako i ostatní autoři soustředil na promluvu obecně, nikoli na roli posluchače. Kognitivistická literatura [např. van Dijk, Kintsch 1983; Rubin et al. 2000; Rubin 1980] považuje poslouchání a čtení jako v podstatě tytéž aktivity recepce textu; zmiňuje-li nějaké odlišnosti, tak následující:

- a) posluchač snáze než čtenář nalezne hlavní myšlenky textu a pragmatické závěry, které z něj plynou; čtenáři mají naopak snazší přístup k jazykové struktuře sdělení a snadněji činí závěry ze specifických řečových detailů; jinými slovy, posluchači z textu pochopí hlavní myšlenku a co z ní pro ně plyne, zatímco čtenáři se více soustředí na detaily textu [Rubin et al. 2000];
- b) posluchači obtížněji řídí své posluchačské strategie a jsou na textu více závislí; čtenáři mají větší možnosti s textem strategicky kooperovat [Rubin et al. 2000].

Současně je třeba si uvědomit, že v naprosté většině posloucháme texty vytvořené pro poslouchání (promluvy) a čteme texty vytvořené pro čtení. S křížením těchto modů se setkáváme ve specifických situacích, buď v institucionálních kontextech (orální prezentace psaných textů při projevech), nebo v analytické práci nebo literatuře (přepisy orálních výpovědí a konverzací; konverzace jako součást literárního díla). Proto asi nemá valný smysl přímo srovnávat poslech textu v konverzaci se čtením textu v románu. Ale má smysl se zamyslet nad formou analytické práce, kdy a) badatel čte přepisy orálně realizovaných textů; b) přepisovatelé mají tendenci překládat orální projev do psané formy projevu.

Ad a) výzkumníka by neměla překvapovat odlišnost orálních textů od textů psaných, tzn. jiný slovník a syntax, například omezený slovník, opakování slov, časté deiktické výrazy, dále zvýraznění hlavních rétorických struktur, orientace na konverzační maximy, informační redundance a další [Rubin et al. 2000: 123]. Tyto odlišnosti jsou vlastní orálnímu textu a nejsou tedy interpretačně relevantní vlastnosti; naopak, pokud by někdo mluvil „jako kniha“, tento aspekt mluvy by byl použit jako kontext pro interpretaci toho, co říká, například ve smyslu ironie. Ad b) při přepisování orálních textů mají přepisovatelé tendenci normalizovat mluvený projev, tzn. převádět foneticky neurčitě výrazy do psaných určitých výrazů, například vyslovené [pže] zapsat jako *protože*, [páč] jako *poněvadž*, rozdělovat zvuky do jednotlivých slov, vytvářet interpunkci, věty apod. [Brown, Yule 1983: 9–10]. Hodně detailní přepis, neobsahující interpunkční znaménka a obsahující čá-

sové údaje o pomlčkách, intonaci a dalších paralingvistických jevech, je pro nezkušeného badatele jen obtížně čitelný, zejména ve větších celcích. Pro některé účely (např. pro konverzační analýzu) je ovšem takto podrobný přepis zapotřebí. Na místě je proto reflexivní přístup a kritické posuzování, jestli zkušenost čtení takového složitěho záznamu může napomoci analytické rekonstrukci zkušenosti původní situace poslechu.

V následujícím textu se budeme věnovat už jen čtení, jednak z důvodu, že naprostá většina textových analýz, o kterých pojednává tato kniha, je založena na čtení a analýze čtení včetně přepisů orálních rozhovorů, a také proto, že čtení je více teoreticky rozpracováno jako modelový příklad recepce textu. Nicméně v návaznosti na van Dijka a Kintsche [1983: 22, pozn. 1] si dovoluji tvrdit, že analogickou argumentaci by bylo možné vztáhnout i na recepci orálních textů. Rovněž je třeba mít na mysli, že čtení textů se netýká jen zkoumaných subjektů, ale stejnou měrou i čtení analytického, které provádí badatel nebo badatelka v rámci výzkumu.

Tony Bennett [1985] rozděluje přístupy ke zkoumání vztahu textu a čtenáře do několika širokých kategorií, které nám vhodně poslouží k přehlednému vstupu do této problematiky:

První přístup se zaměřuje na to, jak texty organizují své vlastní čtení, jak konstruují vlastní srozumitelnost. Typickým představitelem tohoto přístupu je U. Eco a jeho koncept modelového čtenáře. Podle Eca [2010: 65] „text, jak se projevuje svým lingvistickým povrchem (neboli svou manifestací), reprezentuje řetězec výrazových dovedností, které by měl příjemce (čtenář) aktualizovat.“ Nicméně, „co se aktualizování týče, žádný text není kompletní“ [ibid.]. Text od svého čtenáře vyžaduje přinejmenším gramatickou kompetenci, tzn. čtenář by měl být schopný rozeznávat slova v textu, chápat jejich významové složky a zároveň rozumět vztahům mezi slovy v rámci věty; čtenář je v tomto smyslu Ecem pojímán jako operátor. Aby byl čtenář schopen aktualizovat obsah textu a čtení nezůstalo jen sledem mechanických operací rozeznávající povrchovou strukturu textu, musí se z povrchu textu přesunout hlouběji do jeho komplexnosti. Ta má svůj původ v nevyčteném, které musí být aktualizováno v rovině obsahu. A na této úrovni si již text vymáhá na čtenáři aktivní a vědomé kooperační činy [ibid.: 66], například aktualizaci koreferencí (o kom se píše) nebo aktualizaci vlastní encyklopedie (soubor socio-kulturních vědomostí o světě, ke kterému se text vztahuje). Modelový čtenář je podle Eca takový čtenář, který je schopný spolupracovat na aktualizaci textu podle očekávání autora: „[modelový] čtenář se bude pohybovat interpretačně tak, jak se on [autor] pohyboval generativně“ [ibid.: 71]. To však neznamená, že by existoval nějaký pravý smysl textu a čtení realizované modelovým čtenářem bylo jediné možné nebo správné čtení. Eco uvádí Proustovo

čtení jízdního řádu železnice, kdy v názvech zastávek nacházel něžné a labyrintické ozvěny hledání Sylvie. Proust, resp. jeho románový hrdina, se zde nechoval jako modelový čtenář, dokonce se ani nepokoušel o jinou interpretaci jízdního řádu; jízdní řád prostě použil ke svému vlastnímu poetickému účelu, protože jízdní řád předpokládá pouze jediného modelového čtenáře, který bude povrchově aktualizovat místa a časy s jasným lineárním pojetím časovosti.

Druhý přístup vychází z prvního a zakládá se na skutečnosti, že žádný text není čten nezávisle na zkušenosti čtení jiných textů. Podle Bennetta by se analýza měla zaměřit na to, jak některé texty organizují čtení jiných textů, vůči nimž fungují jako rámuující metatexty; „jaké jsou způsoby takového intertextového určení, které vnitřně organizuje texty pomocí specifické konstituce textu jako textu-ke-čtení (text-to-be-read)“ [Bennett 1985: 9]. Za tímto účelem zavádí koncept *čtecí formace* [reading formation; Bennett 1983, 1985]. „Čtecí formace je soubor diskurzních a intertextových determinant, které organizují a oživují čtecí praxi, spojující texty, čtenáře [a kontexty] navzájem specifickým vztahem, a utvářející čtenáře jako čtoucí subjekty určitých typů a texty jako objekty-ke-čtení určitými způsoby“ [Bennett 1985: 7]. Obsah textu zde není vlastností textu, obsah je vytvářen čtenářem ve vztahu k čtecí formaci. Příkladem může být čtení zákonů, pravidel, nařízení, vyhlášek či návodů z různých historických období. Když v současnosti čteme *Deklaraci práv člověka a občana* z roku 1789, jistě pro nás nemá takový osvobozující účinek jako pro mnohé čtenáře v době jejího vzniku; dokonce můžeme mít pocit i jisté skepse vzhledem k uvědomění si toho, co v následujících dvou stoletích následovalo.

Čtení bezpochyby ovlivňuje společenské postavení čtenáře (třída, gender, vzdělání, atd.). Proto třetí přístup k otázce čtení bere v úvahu nikoli modelového čtenáře, ale reálného čtenáře nebo přesněji kategorie čtenářů s jejich specifickými predispozicemi k určitému čtení. Vliv sociální pozice nemusí být přímý, mnohdy je jen zprostředkovaný, tzn. umísťuje člověka do určitých diskurzních formací a teprve tyto formace určují, jaké texty a jak bude číst. Problémem tohoto přístupu je podle Bennetta nejasný převod vlivu z (mimotextového) kontextu na proces čtení; jakým mechanismem diskurzní formace zajišťují (třídně, genderově, vzdělanostně, etnický) preferované texty a jejich preferované čtení? [Bennett 1985: 9–10]. S. Hall navrhl tři možnosti čtení vzhledem k sociálnímu kontextu postavení čtenáře: *dominantní čtení*, v němž je čtenářem akceptován sociálně preferovaný způsob čtení (který se nemusí překrývat se čtením modelového čtenáře implikovaného textem); *vyjednané čtení*, ve kterém čtenář částečně přijímá sociálně preferované čtení, ale v některých ohledech „čte jinak“ podle vlastních zájmů a zkušeností; *opoziční*

čtení, ve kterém čtenář text čte proti sociálně preferovanému čtení a realizuje tak kontrakulturní zájmy své sociální pozice [Hall 2005/1973]. Všechny tři druhy čtení by byly podle mého názoru přítomny při čtení Hitlerova *Mého boje* dnešní českou veřejností.

Posledním zde uvedeným přístupem je etnometodologicky orientovaný výzkum A. McHoula [1982] nebo E. Livingstona [1995], kteří se zaměřili na to, jak aktivita čtení probíhá a jak čtenáři vytvářejí a činí význam textu srozumitelný a zdůvodnitelný pro-všechny-praktické-účely („accountable“). McHoul pomocí pokusu s přeházeným textem například zjistil, jakým způsobem čtenáři dokážou dávat významy i na první pohled nesmyslným textovým sekvencím. Více se této problematice budu věnovat v kapitole 3.3.2.

## 2.5 Základní dimenze analýzy textů

V této kapitole představím základní úrovně, na kterých jsou texty v sociálněvědním výzkumu zkoumány. Nepůjde o detailní přehled, protože jednotlivé analytické metody a techniky budou předmětem dalších kapitol. Cílem této kapitoly je ukázat na tři strukturální úrovně textu (úroveň slov, vět a nadvětých celků), které mají implikace pro zprostředkování významu a v nemalé míře tak předurčují spektrum možností jejich analýzy. Následující části, v nichž se budu věnovat obsahové, diskurzní a narativní analýze, budou všechny tři úrovně textů stále analyticky dostupné. I ve vyprávěních používáme slova, věty a nadvěté struktury; i diskurz sestává ze slov, vět a větších struktur nemluvě o obsahu textů. Proto tato kapitola ty ostatní předchází.

### 2.5.1 Analýza významu v jednotkách kratších než věta

Z hlediska významu (smyslu) je elementární jednotkou textu slovo. Když rozložíme slovo na hlásky nebo slabiky, ztratí se význam. Například slovo *škola*: slabiky *ško* a *la*, případně hlásky *š*, *k*, *o*, *l*, *a* nenesou samy o sobě v textu relevantní význam. (Ponechme stranou zvláštní případy, ve kterých slabiky nebo i hlásky význam nesou, například v textu popěvku „la-la-la“ nebo ve výčtu, kde jednotlivé položky jsou řazeny jako a), b), c) atd.) Jistý abstraktní význam mají v tomto slovu složka *škol* a složka *a*. První z těchto složek má obecný význam, který spojuje slova jako *škola*, *školní*, *školník*, *školství* aj., druhá z těchto složek nás informuje o pádu, čísle a rodu tohoto slova, jde tedy o význam gramatický. V každodenním dorozumívání ale teprve celé slovo *škola* nese význam (zpravidla zapuštěné do nějaké rozvinutější věty). Jaký význam



nese, zjistíme nejspíše ve výkladovém slovníku: 1. vyučovací ústav, vyučování; 2. skupina spisovatelů, umělců atp. jdoucí ve stopách svého učitele [Váša, Trávníček 1937: 1459]. Opět nebudeme brát zřetel na zvláštní případy, kdy slovo nese jiný význam, než který je uveden ve slovnících, například jako vlastní jméno osoby nebo místa.

Slova lze seskupovat do různých kategorií jako třeba slovní druhy (podstatná jména, slovesa apod.) nebo dávat do různých vztahů souřadnosti, příbuznosti, odvozenosti; rozlišovat můžeme také slova domácí a přejatá, spisovná a nespisovná, i když to není tak jednoduchá záležitost, jak by se mohlo na první pohled zdát. Mnohé z těchto lingvistických kategorizací jsou pro textovou analýzu užitečné, protože vypovídají o autorech textů, resp. o cílech, jichž mělo být prostřednictvím textů dosaženo. Pokud text chápeme jako sdělení odkazující na okolní svět, tak z hlediska sociálněvědních textových analýz můžeme říci, že výskyt určitého slova v textu buď odkazuje na existenci nějaké skutečnosti mimo daný text, nebo se podílí na organizaci odkazu na takový fenomén. První typ slov nazýváme plnovýznamová (autosémantika) a druhý typ neplnovýznamová slova (synsémantika). První označují určité předměty, osoby, vlastnosti, děje, stavy a jde převážně o podstatná a přídavná jména, zájmena, číslovky, slovesa a příslovce. Nejčastěji v textech sledujeme právě plnovýznamová slova, protože předpokládáme, že jejich výskyt a vztahy mezi nimi nám pro sociálněvědní účely toho o textu, podmínkách jeho vzniku a užití (čtení) vypovídají víc než slova neplnovýznamová. To jsou zejména předložky, spojky a částice. Někdy je však žádoucí je rovněž sledovat, jak samostatně, tak v jejich spojení se slovy plnovýznamovými. Například spojky *vždyt*, *přece* signalizují argumentační styl [příkladem výzkumu zaměřeného na anglické neplnovýznamové slovo *just* je Weltman 2003].

#### **Jak se vyskytují slova v textu**

Slova se v textu nevyskytují ani náhodně ani rovnoměrně. To první může být sice triviální informace, protože kdyby byla slova v textu náhodně rozložena bez ohledu na gramatiku a jejich význam, nebyl by text srozumitelný; co se týče nerovnoměrnosti výskytu slov, tak ta nemusí být ze čtenářského pohledu evidentní. Při statistickém zpracování textu se však ukazuje velice zřetelně. Platí zde totiž Zipfův zákon [Zipf 1949], podle kterého je základem každého dlouhého textu několik málo velmi častých slov, nejčastěji gramatická slova jako předložky, spojky, pomocná slovesa atd. nezávisle na jazyku textu (míněny jsou přirozené jazyky; pro texty v umělých jazycích, např. počítačové programy, tento zákon platí jen částečně). Tímto rozložením slov se texty liší od typického rozložení jevů ve společnosti, které má zvonovitý tvar

tzv. normálního rozložení, případně trochu vychýlený (šikmý, zkosený) jedním či druhým směrem, kdy nejčastější jevy jsou jevy průměrné nebo nacházející se na křivce blízko průměru (Zipf se ovšem pokusil dokázat, že jeho pravidlo, založené na principu nejmenšího úsilí, platí i pro mnohé společenské jevy, jako třeba rozdělení práce). U textů tomu tak nebývá, slova s průměrnými vlastnostmi (např. počet hlásek) nejsou slova nejčastější (tím je v češtině spojka *a*) [srov. Baayen 2001; Čermák 2007a].

Je třeba upozornit, že výše uvedené se týká rozsáhlých textů (obsahujících tisíce a více slov). V sociálních vědách však často narážíme na texty krátké, například několikáslovní odpovědi na otevřené otázky v dotazníku, mediální texty, rozhovory, pro které běžné frekvenční distribuce neplatí [Baayen 2001: 2–6].

#### **2.5.1.1 Klíčová slova, kódy, lemmata**

Z pohledu analýzy textu je důležitý koncept klíčového slova; klíčové slovo, jak už termín napovídá, je slovo, které něco otevírá. Slovo tak nestojí jen za sebe, ale zastupuje nějakou významovou oblast nebo téma. Klíčové slovo *škola* může zastupovat téma školství, třeba ještě spolu s dalšími slovy. Vyskytne-li se v textu toto slovo, usuzujeme, že je v něm obsažena významová oblast školství. Čím častěji se v textu slovo vyskytuje, tím silněji figuruje daná významová oblast v textu. Klíčové slovo nám tak dává klíč k poznání obsahu textu. Někdy se namísto klíčového slova používá termín kód.

Kód může být klíčovým slovem, ale častěji je kód teoreticky zkonstruovaný a zastupuje více klíčových slov. Například kód „školství“, zastupující významovou oblast spojenou se školní výukou a vzděláváním, může obsahovat klíčová slova *škola*, *učitel*, *učitelka*, *třída*, *gymnázium*, *učebnice*, *zkouška* apod. Každý výskyt klíčového slova je následně uniformně kódován jako „školství“. Vzhledem k ohebnosti slov nejsou při definování kódů používána přirozeně se vyskytující klíčová slova, ale tzv. lemmata, což jsou základní tvary slov reprezentující všechny jejich formy; například lemma *škola* reprezentuje toto slovo v různých tvarech *školy*, *školou*, *školu*, *škol*, *školami* atd. V některých případech je vhodné definovat kód pomocí slovních základů (*-škol-*), které pokrývají i slova odvozená, například *škola*, *školní*, *předškolní*, *školský*, *školák* atd.

Obsahová analýza často postupuje tím způsobem, že podle zaměření výzkumné otázky vytváří kategorie, do kterých systematickým způsobem zahrnuje všechna v daném jazyce nebo korpusu studovaných textů se vyskytující slova nebo tvary slov; například kategorie osoby, místa, činnosti, času; aktivity – pasivity; budoucnosti – přítomnosti – minulosti apod. Aplikace takového kategorizovaného slovníku může rozlišit úseky textu, které pojednávají o určitém tématu, nebo srovnávat například do jaké míry jsou diskusní

příspěvky čtenářů k určité zprávě zveřejněné v internetovém médiu pojaté aktivně nebo pasivně, orientované na přítomnost, budoucnost nebo minulost a jaké další sociálně a textově významné charakteristiky tyto jednotlivé druhy příspěvků mají.

#### 2.5.1.2 *Kolokace, spoluvýskyty, mapy a sítě*

Frekvence slov v textu je základní informací o textu na úrovni slova, ale nikoli jedinou. Další důležitou informací jsou kolokace neboli řetězce (dvojice, trojice, čtveřice atp.) slov, které se vyskytují v textu častěji než náhodným spoluvýskytem. Některé kolokace jsou fráze (ustálená slovní spojení). Při analýze textu nás mohou zajímat určitá ustálená slovní spojení či víceslovné termíny, například *lidská práva, právní stát, znečištění životního prostředí, světový řád, mírové hnutí*; fráze typu *šikmá plocha, nedá se nic dělat, lidská svoboda*; kolokace určitého klíčového slova, tzn. s jakými slovy společně se v textu vyskytuje, například *národ: náš národ, historický národ, ohrožení národa, k národu, pro národ*. Kromě takových ad hoc kolokací můžeme kolokace třídit na abstraktní, gramatické, prostorové, časové, slovesné apod. [Čermák, Šulc 2006].

Vedle kolokací, tedy slov vyskytujících se v textu bezprostředně za sebou, můžeme pozorovat i spoluvýskyty slov v námi definované kontextové jednotce. Teoreticky vychází tento přístup z Firthovy distribuční hypotézy, která říká, že slova vyskytující se v podobných kontextech mají i příbuzný význam [Firth 1957]. Ačkoliv tato hypotéza je primárně lingvistická a má pomoci při zjišťování sémantické podobnosti slov, lze ji využít i pro sociálněvědní zkoumání textů. Například nás může zajímat, která z klíčových slov v odborářské mediální komunikaci (*práce, spravedlnost, právo, mzda, zaměstnanci, odbory*) se vyskytují spolu relativně častěji než ostatní a indikují tak existenci nějakého společného tematického (významového) pole, například pole odborářských nároků [viz část 4.7].

Na bázi měření spoluvýskytů lemmat, klíčových slov nebo kódů lze vytvářet prostorová zobrazení vztahů mezi nimi, které mají formu map. I když jistou informaci nám podá rovněž jednoduché zobrazení polohy vybraného slova v textu, kterou znázorníme body na úsečce (jestli se v textu vyskytuje rovnoměrně nebo je jeho výskyt v některých částech textu častý a v jiných vzácný či se nevyskytuje vůbec), mnohem informačně bohatší je zobrazení ve dvourozměrném (případně trojrozměrném) prostoru, ve formě jakési mapy. Získáváme tak informaci o vnitřní struktuře textu, jestli je tematicky nebo diskurzivně polarizován, zda má ústřední linii tvořenou opakujícími se slovy nebo naopak se slova průběžně obměňují. Jedná-li se o texty vytvořené v rámci kontrolovaného experimentu různými jedinci, lze na základě ana-

lyzy spoluvýskytů slov vytvářet například mentální či paměťové mapy, individuální sémantické regiony apod. [Carley, Palmquist 1992].

Mapám je podobné síťové zobrazení struktury slov, které rovněž vychází ze sledování spoluvýskytů klíčových slov. Na rozdíl od map se analýza sítí soustředí na popis vztahů mezi jednotlivými prvky či uzly sítě, reprezentovanými nejčastěji slovy nebo kódy. Tyto vztahy mohou být posuzovány podle síly, směru, příznaku. Průkopnickou práci v používání modelu sítě při analýze textových dat vykonala Kathleen Carley [Carley, Palmquist 1992; Carley 1997; Diesner, Carley 2005]. Jinou aplikací síťové analýzy spoluvýskytů jsou tzv. citační mapy, které ukazují na vztahy mezi autory odborných textů podle toho, jak na sebe vzájemně v textech odkazují [např. White, McCain 1998].

#### 2.5.1.3 *Další vztahy či atributy textu na úrovni slov – KWIC*

Podle Firthovy hypotézy uvedené výše, kterou Firth lakonicky vyjádřil výrokem „you shall know a word by the company it keeps“ [1957: 179], mají slova vyskytující se v podobných kontextech podobný význam. Například slova *lékař* a *doktor* se vyskytují v podobných kontextech: *lékař/doktor diagnostikoval, předepsal, operoval; šel jsem na vyšetření k lékaři/doktorovi; potřebuji lékaře/doktora*. Můžeme sice nalézt výraz *pilot diagnostikoval* (závadu na motoru), ale nikoli už *šel jsem na vyšetření k pilotovi*. To znamená, že slova *lékař* a *doktor* mají podobný význam. Postup však můžeme i obrátit a ptát se, zda se dvě slova, o nichž se domníváme, že mají podobný význam, vyskytují též v podobných kontextech. K tomu je vhodná metoda keyword-in-context (KWIC), ve které sledujeme textové okolí všech výskytů zvolených klíčových slov v textu (např. u slov *homosexuál* a *gay* mohou odlišné kontexty ukázat odlišnosti v jejich významech nebo významových konotacích).

Na konec této kapitoly je nutné uvést, že největší analytickou sílu mají výše uvedené přístupy, když jsou užívány komparativním způsobem. Samotná frekvence klíčového slova v daném textu nemá prakticky žádnou informační hodnotu. Pokud ji srovnáme s četnostmi ostatních slov, už to něco o daném textu vypovídá. Pokud navíc srovnáme četnost určitého slova nejen vůči ostatním slovům v textu, ale i vůči jiným textům, dozvíme se toho o textu ještě více. Jistým ideálním stavem zde může být existence referenčního textu, resp. korpusu, který představuje jakýsi teoreticky podepřený standard a vůči kterému můžeme ostatní texty poměřovat. Příkladem referenčního textu může být významný manifest nějaké politické strany, vůči němuž můžeme poměřovat texty, které strana produkovala předtím a potom, a sledovat změny, jež se ve frekvencích slov, kolokacích, spoluvýskytech slov, jejich kontextech a dalších vztazích na úrovni slov objevují [srov. Laver, Benoit, Garry 2003]. Příkladem referenčního korpusu je synchronní repre-

zentativní korpus současné psané češtiny označený SYN2010, který obsahuje publicistické, odborné a beletristické texty čítající v souhrnu 100 milionů textových slov (<http://ucnk.ff.cuni.cz/syn2010.php>).

#### 2.5.1.4 *Hodnocení*

Sociálněvědní výzkumy soustřeďující se při analýze textů výhradně na podvětné jednotky, tj. slova, jejich frekvence a jejich asociace, jsou dlouhodobě kritizovány za redukcionismus. Výtka směřuje ze „čtenářských“ pozic – člověk nikdy nečte text jako oddělená slova nebo slovní spojení, ale vždy nachází větší a komplexnější významové celky, minimálně na úrovni vět. Teprve věta nebo řada vět dává aktuálně, dopředně i zpětně slovům v textu jejich konkrétní význam, například ironií, metaforickým použitím, karikaturou.

Tato kritika pro slovní redukcionismus má sice částečné oprávnění, ale rozhodně nediskvalifikuje textovou úroveň slov pro analýzu. I když lze tvrdit, že význam existuje na úrovni věty a nadvětných struktur, tak to ještě neznamená, že slova a jejich asociace žádný sociálně relevantní význam nést nemohou. Jak jsem ukázal výše, čtenář s textem kooperuje už na úrovni slovníku. Když někdo v orálně-historickém vyprávění definuje období po listopadu 1989 jako období *po revoluci*, jiný *zas po převratu* a ještě další *po listopadu*, dává té události potenciálně odlišný význam. Jaký význam, to už nelze z úrovně slov zjistit, museli bychom do analýzy zahrnout i interpretační aktivitu čtenáře, ale potenciální významová odlišnost se dá sledovat už zde.

#### 2.5.2 *Analýza významu na úrovni věty*

Zásadní odlišností úrovně slov a úrovně věty spočívá v tom, že slova zpravidla nevytváříme, ale používáme ta, která jsou v daném jazyce dostupná, zatímco věty vytváříme zpravidla vždy nové nebo jako nové. I když slova nesou subjektivně srozumitelný význam, tento význam je určen komunitou uživatelů jazyka, nikoli jednotlivcem; význam slova (nikoli volba) se nevztahuje primárně k mluvčímu nebo posluchači, ale k jazyku (k řeči). Oproti tomu význam věty se váže primárně k mluvčímu/pisateli nebo posluchači/čtenáři – je to *jeho* věta. Věta je projevem naší individuální vůle i schopnosti komunikovat v jazyce. Tato principiální odlišnost slova a věty otevírá nové horizonty pro výzkum textů.

Lingvistika vede polemiky o tom, co přesně věta je, zda píšeme či mluvíme ve větách nebo ve výpovědích, ale to není pro sociálněvědní textové analýzy až tolik podstatné. Nám postačí vědět, že věta je uzavřená gramatická struktura složená ze slov (a interpunkčních znamének), že zpravidla obsahuje přísudek a další větné členy a je úplným sdělným celkem.

#### 2.5.2.1 *Pragmatika jako výzkum užívání řeči*

Pro sociálněvědní textové analýzy jsou syntaktické a částečně i sémantické lingvistické analýzy věty vcelku nezajímavé. Buď si z nich berou jen to triviální, jako to, že věta spojuje nějaký *děj* (proces či jednání) s nějakými *aktanty* (činnými nebo trpnými osobami či předměty) a nějakými *okolnostmi* (časovými, místními apod.), nebo je prakticky ignorují a zabývají se tím, za jakých okolností má věta nebo výrok smysl a stává se tak součástí komunikace. Větší význam má pro sociologický kontext pragmatika, tj. ta odnož jazykovědního zkoumání, která se zabývá používáním řeči a snaží se mapovat vztahy mezi řečí a světem, ve kterém je používána, tj. jak mluvena či psána, tak poslouchána a čtena. Pro účely tohoto přehledu považuji za vhodné seznámit čtenáře se základními koncepty pragmatiky.

Je zajímavé, že k rozvoji pragmatiky přispěli zásadní měrou filozofové (Wittgenstein, Austin, Searle nebo Grice) a poukazuje to na to, že pole pragmatiky je současně v řeči i v jednání a ve světě mimo ni. V pragmatice počátku převažoval kódový model komunikace, tj. názor, že řečová komunikace je proces, ve kterém jeden člověk komunikuje (předává) nějaké sdělení druhému člověku tím, že sdělení kóduje do společného jazyka (kódu). Toto sdělení je úzce vázáno na věty jako sdělné celky. Kódový model dobře odpovídá naší zkušenosti, kterou sami jako uživatelé jazyka máme: když chceme někomu něco sdělit, nějakou myšlenku nebo pocit, hledáme pro to vhodná slova, z nichž utvoříme věty a ty vyslovíme (nebo napíšeme). Převod myšlenky do vět můžeme označit jako kódování. Zvláště intenzivní je tato zkušenost jazykového kódování v případě, že máme něco sdělit v řeči, kterou dobře neovládáme. Když je sdělení zakódováno do patřičných vět, které jsou vysloveny nebo napsány, přichází moment posluchače/čtenáře, jenž věty dekoduje, aby se dostal k našemu sdělení. Aby mohla být komunikace úspěšná, musejí obě strany ovládat stejný komunikační kód, stejnou řeč, to je evidentní. Předpokladem kódového modelu komunikace je existence nějakých myšlenek, zkušeností, pocitů, jednoduše nějakého mentálního materiálu, který se může stát předmětem sdělení, a hlavní otázka spočívá v tom, jakými kompetencemi mluvčí a posluchači musí disponovat, aby při jeho kódování do řeči a dekodování z řeči správně zacházeli s patřičnými kódy.

Ačkoliv je tato představa komunikace intuitivně srozumitelná a odpovídá naší zkušenosti užívání řeči, nedokáže vysvětlit některé komunikační situace. Běžně se setkáváme s tím, že výrokům nerozumíme pouze na základě jejich doslovného znění, ale k porozumění používáme i jiné dostupné informace. Například slyšíme-li konstatování „Ten je nadobro mrtvej.“ a mluvčí ukazuje na ležícího mrtvého psa, nebo slyšíme-li tutéž větu a mluvčí mávne rukou směrem ke klimbajícimu opilému mladíkovi, pochopíme sdělení mluv-

čího vždy jinak. Přesto nám k tomu žádný jazykový kód obsažený ve větě nedává oporu; co věta znamená, musíme rozhodnout na základě skutečností přicházejících z neязыkové oblasti (z kontextu). Konkrétní okolnosti, ve kterých byla věta řečena, způsobí, že v prvním případě budeme rozumět výroku doslovně, zatímco v druhém případě ho budeme považovat za hyperbolu. Jiným příkladem může být konverzační žádost „Zapomeň na to.“, kdy nechápeme výrok doslovně (ani nelze si něco přestat pamatovat na požádání), ale usoudíme, že mluvčí chce, abychom se k dané záležitosti v budoucnu nevraceli. To znamená, že sdělení porozumíme nejen na základě toho, co je v něm zakódováno, ale i na základě našeho úsudku o záměrech, jež vedly mluvčího k vyslovení jeho žádosti. Podobným případem jsou ironické či sarkastické výroky typu „Ten to teda vyhrál!“, které, abychom druhému správně porozuměli, nesmíme brát doslova. Konečně posledním příkladem nedostatečné explanatorní síly kódového modelu jsou tzv. performativní výroky [Austin 2000/1962]; jejich smyslem není informovat, ale něco provést. Například výrok oddávajícího „Oddávám vás.“ nepřináší sdělení jenom o tom, že dotyčný právě oddává („Tímto vám sděluji, že vás oddávám.“), ale samotným výrokem to dělá.

Takové a další příklady vedly k tomu, že kódový model přestal být vyhovující. Nikoli proto, že by byl nesprávný, ale byl schopen vysvětlit pouze malou část řečových komunikací; tu část, kde k tomu, aby byla zdařilá (felicity condition), stačil pouze společný jazyk. Velká část komunikace je však komplexnější a pro tu je potřeba najít jiné vysvětlení. Tím je inferenční (usuzovací) model, který tvrdí, že řečová komunikace je úspěšná tehdy, když posluchač rozpozná komunikační záměr mluvčího, tedy to, co mu chce mluvčí sdělit. K poznání komunikačního záměru nestačí, aby mluvčí a posluchač sdíleli společný komunikační kód (jazyk), ale musí sdílet i určité usuzovací strategie. Čtenář by se mohl v tuto chvíli zarazit a ještě než se pustí do čtení o inferenčním modelu si položit otázku: nepostačoval by v případě analýzy textů kódový model? Vždyť uváděné komplexní formy komunikace, kde si lidé rozumí na základě jiných skutečností, než které jsou obsahem sdělení, se týkají především komunikace mluvené, zejména konverzace, která probíhá v simultánní přítomnosti. Textové analýzy ovšem nejsou zpravidla analýzou živé konverzace a týkají se textů jako záznamů řečových aktů již proběhnutých. Proč tedy nezůstat u kódového modelu komunikace?

Odpověď je jednoduchá: málokdy rozumíme textům pouze na základě jazykového kódu, tj. porozumění významu slov a gramatických pravidel. Jak jsme si už ukázali v kapitole o čtení textů, čtenář, ať už je to textový analytik nebo ne, nepoužívá pouze slovník a gramatiku, aby textu porozuměl, ale rovněž svou encyklopedickou kompetenci (znalost kulturního kontextu textu).

při čtení narativních textů navíc čtenář za účelem porozumění textu konstruuje různé možné světy [viz část 3.3.2], což není dekódovací, ale inferenční strategie. V případě přepisů rozhovorů, které tvoří nemalou část textových dat pro sociálněvědní analýzu, je nutné brát v úvahu i to, že aktéři, jejichž řečové jednání je zaznamenáno, při komunikaci používali inferenční strategie porozumění druhému a textová analýza by měla být schopna je při interpretaci zohlednit. Proto při analýze textů nemůžeme zůstat u kódového modelu.

Inferenční model je založen na předpokladu, že komunikace je interakce, ve které se mluvčí snaží, aby posluchač rozpoznal jeho komunikační záměr (tj. co mu chce sdělit), a to nejlépe tak, že zvolí výrazy, jež to v dané situaci/kontextu usnadní. Posluchač zase usiluje o pochopení komunikačního záměru mluvčího na základě výrazů, které mluvčí zvolil, a kontextu jeho promluvy. Inferenční model tak nepracuje pouze s řečovou komunikací, ale bere v úvahu všechny další způsoby, jimiž lze vyjádřit komunikační záměr, zejména gesta a zvukové signály při komunikaci tváří v tvář a různé formy paratextových signálů při textové komunikaci (např. typy písma). Kompetentní komunikační partner umí tyto signály jednak strategicky vytvářet, jednak diagnosticky „číst“. Například když jsem v předchozí větě dal slovo „číst“ do uvozovek, dal jsem tím čtenáři najevo, že slovu „číst“ má rozumět jinak než je běžné, že to slovo označuje nějaké ne-tak-úplně-normální čtení. Převodeme-li tento můj řečový akt na obecnou úroveň, uvozením slova číst, aniž by šlo o přímou řeč, jsem mírně narušil čtenářovo běžné očekávání pravidel textové komunikace; toto narušení očekávání mělo pravděpodobně ten důsledek (implikaci), že se čtenář zamyslel, co jsem tím „diagnostickým „čtením““ chtěl říci (jaký byl můj komunikační záměr).

Uvedený příklad nám dále ukázal logiku fungování komunikace podle inferenčního modelu. Lidé sdílejí určité předpoklady o tom, jak by měla normální komunikace probíhat. Jedním ze základních očekávání je to, že mluvčí/pisatel má pro svou promluvu/psaní nějaký komunikační důvod. Obvykle nepředpokládáme, že někdo mluví bez záměru něco sdělit, i když občas se to stává (např. některé ceremoniální promluvy, promluvy při nácviu výslovnosti slov). Od posluchače/čtenáře zase očekáváme, že je schopen z promluvy v daném kontextu odvodit smysl sdělení. Dále předpokládáme, že lidé používají slova v jejich doslovném významu, pokud nedají nějak najevo, že tomu tak není (např. intonací nebo uvozovkami, viz výše).

P. Grice [1989/1975: kap. 2] k těmto obecným komunikačním očekávaním přidal ještě specifická očekávání, která lidé vůči sobě mají, pokud spolu rozmlouvají: jednak že to, co říkají, je pro rozhovor relevantní; dále že se snaží mluvit uspořádaně, maximálně jasně, ani příliš stručně, ani příliš rozvláčně a pokud možno pravdivě. Že tato konverzační očekávání nemusí být

vždy naplněna, si dokážeme snadno představit, nicméně poskytují interpretační oporu pro usuzování, co nám chce náš konverzační partner sdělit [ibid.]. Když dojde ze strany mluvčího k zjevnému porušení některého z těchto očekávání, máme důvod se domnívat, že nám tím dává najevo to, co nechce nebo nelze (z nějakého důvodu) říci doslovně, stručně, jasně, uspořádaně a pravdivě. Názorným příkladem je rozhovor rodiče (R) s adolescentním potomkem (A), který se chystá večer odejít ven:

R: „Kam jdeš?“

A: „Nikam.“

R: „S kým jdeš?“

A: „S nikým.“

R: „Kdy se vrátíš?“

A: „Nevím.“

I když by se mohlo podle kódového modelu zdát, že rodič dostal odpovědi nesmyslné (odcházející říká, že neodchází), nepravdivé (odchází na schůzku s někým) nebo neinformativní (neví, kdy se vrátí), podle inferenčního modelu tomu tak nutně nemusí být: rodič mohl usoudit například to, že slovem „nikam“, chtěl mluvčí sdělit, že půjde na nějaké obvyklé místo, výrazem „s nikým“, že půjde s kamarády, a slovem „nevím“, že přijde možná pozdě. Jiný rodič by mohl usoudit, že příčinou negativních odpovědí je neochota prozradit skutečný úmysl potomka, protože by mohl vyvolat diskuzi či hádku; další inferenční variantou by mohlo být to, že potomek touto formou demonstruje rodiči svou dospělost a tedy i právo nesdělovat důvody svého jednání („Co vám je do toho?“). Jak přesný bude tento odhad komunikačního záměru, bude záviset na tom, jak dobrou inferenční strategii rodič použije. Na druhé straně, zda tento rozhovor, tato komunikace, proběhne zdárně, záleží i na tom, jestli adolescent-mluvčí poskytne svým celkovým chováním rodiči-posluchači dostatečnou oporu pro to, aby svůj inferenční odhad přijal jako věrohodný.

Inferenční model dobře vysvětluje způsob, jaké komunikační strategie mluvčí a posluchači používají na úrovni vět. Nicméně nezabývá se už tím, jaké motivy vedou mluvčí k tomu, aby porušili komunikační očekávání. Existují nějaké pádné důvody pro to, abychom nemluvíli doslovně, jasně, dostatečně stručně atd.? Odpověď na tuto otázku formuloval na obecné úrovni J. Austin [2000/1962]: lidé často svou řečí (ale myslel tím větami) chtějí něco udělat. Známý je příklad náčelníků jistých jihoamerických indiánů, kteří, pokud si chtěli udržet svou autoritu a moc, museli často a dlouze mluvit ke svým soukmenovcům, i když ti jim naslouchali jen na půl ucha, pokud vů-

bec; řeč náčelníků byla ryze ceremoniální a spočívala v opakovaném opěvování tradičního způsobu života [Clastres 1974: 135]. Mohli bychom říci, že náčelníci prostřednictvím své řeči dělali politiku. Ale důvodů, proč nemluvit v souladu s komunikačním očekáváním, je celá řada. K nejčastějším důvodům patří zdvořilostní ohledy, tj. snaha udržet svou tvář a nepoškodit cizí [politeness theory; Brown, Levinson 1987; Goffman 1955], statusové rozdíly (používání latiny lékaři, cizích slov odborníky) nebo tzv. performativní výpovědi, kterými vykonáváme určité skutky [Austin 2000/1962]. Už jsem zmiňoval výrok „Oddávám vás.“; Austin ukázal, že takových performativů je velké množství, například všechny typy právních verdiktů (obviňovat, vyměřovat, zprošťovat obžaloby, stanovovat apod.), rozhodnutí (jmenovat někoho někým, hlasovat, degradovat, apod.), závazných prohlášení (slíbit, zavázat se, přísahat, sázet se apod.).

Při důkladném pohledu na různé typy vět, které pronášíme, však zjistíme, že jen málokterá výpověď je čistě komunikační, bez performativního aspektu. Eco [2010: 69] uvádí příklad toho, že se performativům prakticky nelze vyhnout: Osoba u kavárenského stolu vytahuje cigaretu z krabičky a musí – pokud nechce vypadat nezdvěřile – nabídnout také svému společníkovi s otázkou „Kouříte?“ Dotyčný, pokud odpoví prostě „Ne.“ se může dopustit aktu nezdvěřilosti, zatímco když zdvořile odpoví „Děkuji, nechci.“, neodpovídá na položenou otázku; když řekne „Ne, děkuji.“, tak polovina této odpovědi je opět zdvořilostní, nikoli jen jazykově komunikační akt. Dokonce i taková na první pohled ryze komunikační promluva „Dnes jsem obědval doma.“ může být chápána performativně, pokud ji vysloví třeba host před slavnostním obědem, kam byl pozván, nebo papoušek („Umí mluvit!“).

### 2.5.2.2 Gramatika příběhů

Ještě než přejdeme k analýze zaměřující se v textech na nadvětné celky, uveďme si jeden specifický druh větné analýzy vyvinutý R. Franzosim – gramatiku příběhu (story grammar) [Franzosi 1990, 1994, 1997, 2010]. Návaznost na naratologické studie V. Proppa, A. Greimase nebo T. Todorova, kteří v duchu strukturalistického přístupu hledali invariantní „gramatiku“ vyprávění, je pouze slabá [více o tom v části 5.4.1]. Franzosiho metoda se opírá o jednoduchou lingvistickou poučku, že úplná věta z gramatického pohledu obsahuje činnou složku reprezentovanou přísudkem, aktanční složku reprezentovanou podmětem anebo předmětem a okolnostní složku (příslovná určení místa, času). Proto *gramatika* příběhu. Druhým pilířem metody je vlastnost každého příběhu spojovat aktéry, jednání, místo a čas. Proto *gramatika příběhu*. Když se tyto dvě vlastnosti spojí, můžeme předpokládat, že každá věta, která je součástí vyprávění dějové linie příběhu, bude obsaho-

vat tyto stavební prvky. Na tomto základě je vytvořena metoda analýzy vyprávění, která rozkládá příběhy do vět a věty na tzv. sémantické triplety. Sémantický triplet sestává z aktéra, který jedná, vůči nějakému objektu (někomu nebo něčemu); tvoří tak kostru narativních vět, která je základem kódování. Tímto způsobem je možné systematicky analyzovat velké množství příběhů (např. novinových zpráv o určitých událostech nebo biografických vyprávění) co do typů aktérů, jednání, objektů a okolností. Více rozvedeme tuto metodu v části 5.4 věnované „strojové“ narativní analýze.

### 2.5.3 Analýza významu v nadvětném rozsahu

Zhruba od 70. let minulého století se rozvíjí interdisciplinární bádání, které si bere za svůj předmět lingvistické celky větší než věta – texty. Ukázalo se totiž, že texty mají vlastnosti, které nemohou být vysvětleny vlastnostmi věty [Eco, 2010: 24]; pro přehled viz [Beaugrande, Dressler 1981] a nemohou tak být popsány standardními lingvistickými koncepty. I když nadvětnými strukturami (texty) se již od antiky tradičně zabývá rétorika a poetika, později pak stylistika a literární věda, nově se bádání zaměřilo zejména na to, jak jsou texty (mluvené i psané) organizovány a jak jsou používány a interpretovány ve společenském kontextu. V důsledku terminologických překryvů, víceznačností a nesrovnalostí se výzkum textů na úrovni nadvětných struktur označuje různými názvy, například lingvistika textu, gramatika textu, textová analýza, analýza diskurzu, diskurzivní analýza, nehledě na tradiční názvy rétorika, stylistika, literární věda, hermeneutika nebo speciální přístupy jako konverzační analýza nebo narativní analýza. Protože se problematice analýzy diskurzu budeme věnovat později jako jednomu z hlavních přístupů textové analýzy, není v tuto chvíli názvoslovné vyjasnění klíčové; co je však důležité, je věcně si tento široký směr bádání uspořádat. Začneme tím, že v rámci výše uvedeného interdisciplinárního bádání volně rozlišíme a) analýzu promluv (diskurzů), tj. užívání řeči v různých sociálních kontextech a b) analýzu diskurzu jako typu výpovědi, která má (zpravidla) nadvětnou strukturu. Na tomto místě se budeme věnovat jen tomu druhému, protože analýza promluv prostupuje všechny dimenze analýzy textů (od slov až k rozsáhlým textům). Bude nás zde tedy zajímat, jak lze studovat význam vytvářený nadvětnými celky. Připomínám, že jako pro analýzu vět platí některé poznatky získané na úrovni slov, protože mluvčí vytvářejí věty ze slov a posluchači interpretují význam věty na základě konfigurace slov v ní obsažených, tak i pro nadvětnou úroveň platí určité poznatky získané na úrovni vět, obzvláště pragmatika komunikačního aktu.

Zajímá nás, jak jsou věty uspořádány do sekvencí a jak jsou soubory vět konstruovány a organizovány smysluplným způsobem. Konkrétně se ptáme: jestliže slova ve větě jsou uspořádána gramatickými a syntaktickými pravidly (která ovšem nezaručují její smysluplnost, tu přiděluje větě posluchač/čtenář), tak co uspořádává text? Existuje vůbec nějaká textová gramatika, která by byla nezávislá na významu textu?

Lingvistika textu/analýza diskurzu definuje několik standardů textovosti, tzn. vlastností, které musí určitý soubor slov a vět vykazovat, aby bylo možné ho považovat za (čitelný) text [Beaugrande, Dressler 1981; Brown, Yule 1983; van Dijk 1977].

Běžné je spojení vět do větších diskurzivních celků prostřednictvím textové koheze, tzn. slovních a gramatických prvků, které dokážou „lepit“ věty k sobě, překonat tak omezení dané gramatickou strukturou a tím vytvářet rozsáhlejší text [srov. Daneš 1985: kap 7]. Například slovo *například* z této věty vytváří spojení s větou předcházející, protože na ni odkazuje jako na obecné tvrzení, jehož konkrétní příklad se chystá poskytnout. Ale už ne s touto větou, která je gramaticky přilepena k předchozí spojkou *ale*. A další spojka *a* napojuje gramaticky tuto poslední větu. Koheze lze také dosáhnout zopakováním slova nebo synonyma z předcházejících vět jako v této větě slovo *koheze*. Jak je z výše uvedeného zřejmé, pro textovou kohezi není kooperační činnost čtenáře nebo posluchače nutně potřeba [pro více pragmaticky zaměřený pohled na kohezi srov. Bublitz 2011].

Další textové nástroje uspořádávání sekvencí vět jsou více závislé na významu, a proto i na čtenáři a posluchači. Jedná se o textovou koherenci, která pracuje na bázi propojování významů v rámci textu, a tím ho činí pro posluchače/čtenáře smysluplný jako celek [Bublitz 2011; van Dijk 1980]. Přesun pozornosti na posluchače/čtenáře však neznamená, že text není smysluplný již pro jeho producenta – mluvčího nebo pisatele. Je to jako v případě věty, význam textu je autorem konstruován v okamžiku produkce (je produkován jako smysluplný) i v okamžiku recepce (inferenční aktivitou jeho posluchače/čtenáře).

Mezi drobné, ale efektivní spojníky vět do koherentního diskurzu, zejména v konverzaci, patří tzv. diskurzivní markery – slova nebo slovní spojení, která odkazují na porozumění významu předchozí věty v kontextu celé promluvy. Například: *aha, opravdu?, hmm, neříkej, no ne, fakt?, kecáš!, no vidíš to!* apod. Spojení může vypadat takto: „Představ si, co se mi včera stalo. Jel jsem tramvají a nastoupil nějaký důchodce a šinul si to ke mně.“/„Hmm.“/„Nejdřív jsem ho nepoznal, ale pak jo. Víš, kdo to byl?“/„Kdo?“/„Vodvárko, školník, ze základky!“/„Fakt? A poznal tě?“

Jiným způsobem, jak se dosahuje koherence v promluvách, je referenční identita (osob, předmětů, dějů, míst, časů). Text získává pro čtenáře/poslu-

chače na koherenci, pokud je v něm identita osob, míst, předmětů, na které se v textu odkazuje, identická napříč větami. Vezměme si příklad: „Kamila byla včera v Brně na úřadě. Musela si tam vyřídit doklady. Ale měli zavřeno. Takže odjela domů bez ničeho.“ Čtenář předpokládá: a) identitu osob – „Kamila“ z první věty je tatáž osoba, jako ve větách následujících; b) identitu míst – „úřad“ v první větě je „tam“ v druhé a „(oni)“ ve třetí větě; identitu času – „včera“ z první věty je čas „měli“ ve větě třetí; identitu objektů – „doklady“ z druhé věty se vztahují k „bez ničeho“, tj. bez dokladů, ve větě čtvrté. Povšimněme si, že referenční identita funguje jen jednosměrně, obrácení pořadí vět nebo jejich přeskupení by porušilo význam textu. To poukazuje na linearitu textu, který se odvíjí pouze jedním směrem, zvláště pak v mluveném textu (diskurzu). V psaných textech je možné přeskakovat, vracet se, opakovat čtení, to znamená číst text jako hypertext.

Na sémantické úrovni zajišťuje koherenci textu téma. Problematika tématu (angl. topic) a tematizace je komplikovaná [Brown, Yule 1983: kap. 3; Eco 2010: kap. 5; Chafe 2008: kap. 35] a v tomto přehledu se ji budu snažit rekonstruovat maximálně srozumitelně, což může vést k potlačení některých, z mého pohledu ne zcela podstatných aspektů pro sociálněvědní bádání. Předně je třeba uvést, že téma je metatextový nástroj produkce významu, to znamená, že je v textu vytvářeno mluvčím, pisatelem a následně přisouzeno posluchačem/čtenářem. Nejprve vymezíme téma vůči komentáři v rámci jedné věty. Ve větě „Kamila byla včera v Brně na úřadě.“ je Kamila jako osoba téma, to znamená určitá známá skutečnost, o které pojednáváme, a část věty „byla včera v Brně na úřadě“ je komentář, tzn. nová, dosud kontextově neurčitá skutečnost. Toto rozlišení je analogické rozlišení téma – réma (známá – neznámá skutečnost ve větě). V případě delších textů přesahujících úroveň jedné věty je stanovení tématu a případně jeho odlišení od komentáře méně jasné. Nicméně při analýze textů se bez konceptu tématu často neobejdeme.

Když na to půjdeme oklikou, můžeme říci, že téma vymezuje hranice textu; tam, kde text začíná „být o něčem“ a tam, kde přestává „být o tom“, se objevuje a ztrácí téma textu. Protože konstrukce tématu textu je žádoucí jak ze strany mluvčího/pisatele, aby posluchač/čtenář textu porozuměl, tak i ze strany posluchače/čtenáře, aby mu text dával smysl, tzn. věděl, o čem text je, bývá téma v textu často zvýrazněno textově, například může být vyjádřeno první větou („V následujícím úvodu/odstavci/několika větách vysvětlím, co je to diskurz.“) nebo poslední větou („Nyní už by mělo být posluchači/čtenáři jasné, co to diskurz je.“); v případě psaného textu lze téma vymezit i paratextově, například titulkem („Definice diskurzu“) nebo tučným písmem („Jedním z klíčových pojmů současných sociálních věd je **koncept diskurzu.**“) nebo

jednoduše odsazením odstavce. Další strategií je přítomnost klíčových slov, která se v textu častokrát opakují a tím dávají posluchači/čtenáři najevo, o čem text je. V souvislosti s rozdělením na psané a mluvené texty lze říci, že psané texty využívají jak textových, tak paratextových strategií, mluvené texty jsou odkázány převážně na textové strategie, nicméně mohou využívat i gestikulace, intonace a další neverbální prostředky komunikace; například odmlčení se, úklona zdůrazňují konec promluvy nebo její jedné části.

Jelikož téma je z velké části pragmatický akt posluchačského/čtenářského usuzování (příčemž důležitou roli hraje hypotéza komunikačního záměru mluvčího/pisatele), různí čtenáři mohou podle své encyklopedické kompetence najít ve stejném textu i několik témat současně – anebo žádné. Jakmile si posluchač/čtenář učiní hypotézu o tématu textu, interpretuje následující text v rámci této hypotézy až do doby, kdy se v textu objeví téma nové. V tomto bychom mohli téma přirovnat ke Goffmanově pracovním konsensu na definici situace: „Účastníci přispívají k jednomu společnému výkladu situace, což neznamená ani tak skutečnou shodu o tom, co skutečně existuje, ale spíše skutečnou shodu o tom, že dočasně budou respektována ta a ta tvrzení týkající se toho a toho“ [Goffman 1999: 17]. Tato podivuhodná shoda mezi teorií sociální interakce a teorií textu je dokladem toho, jak silně je čtení/poslouchání textů kooperativní aktivita. To je ještě patrnější, když textem je konverzace. Účastníci konverzace vytvářejí společnou definici situace ve smyslu jejího tématu (tedy o čem konverzace je). Jak je zřejmé z Goffmanova tvrzení, není důležité, zda mezi nimi existuje skutečná shoda o tom, o čem v danou chvíli konverzuji – jeden může myslet, že o počasí, druhý že si vyměňují zdvořilostní fráze – ale že budou schopni dále generovat jim smysluplný text.

Zajímavou otázkou je téma, resp. jeho udržení a změna v konverzaci [např. Drew, Holt 1998] nebo ve vyprávění [např. Eco 2010]; podle Eca je charakteristickým momentem narace dějovost (příběhovost), která se vynořuje na textovém povrchu jako dějová zápletka neboli narativní komplikace a její následné vyřešení. Ve vztahu k nadvětným strukturám textu [makrostrukturám, srov. van Dijk 1980] je však vhodné poznamenat, že narativní potenciál má i jednoduchá věta, například „Pomalou jsme sjížděli dolů z kopce.“ Také platí, že makrostruktura (např. téma) může být mluvčím/pisatelem zexplicitněna jako v následující větě, uvozující biografické vyprávění: „Já začnu asi od konce té kauzy babička.“

Další textovou makrostrukturou je argumentace, která typicky obsahuje nejprve předpoklady, pak argument a končí závěry. Vědecké texty vykazují neobyčejně pevné makrostrukturní vlastnosti, tj. způsoby nadvětného uspořádání [Čmejková, Daneš, Světlá 1999].

#### 2.5.4 Analýza významu na kontextové úrovni

V oddílu 2.3 jsem popsal vztah textu, kotextu a kontextu. *Kotext* je textové okolí určitého úseku textu (slova, věty, promluvy), *kontext* je interpretačně relevantní prostředí textu patřící do jiného sémiotického systému (určitý text, např. promluvu, interpretujeme s ohledem na kontext, pokud vezmeme v úvahu i neязыkové prvky jeho prostředí, např. gesto, které promluvu doprovází).

Proč je kontext při studiu textů důležitý? Asi není třeba dokazovat, že kontext má vliv na porozumění textu; když se setkáme s textem „Byl pozdní večer, první máj...“ budeme ho interpretovat jinak, když víme, že je to první verš ze slavné básně K. H. Máchy, než když bychom to nevěděli. V prvním případě řekneme, že jde o první verš atd., v druhém případě to budeme považovat za popisné časové určení. Nebo když na otázku: „Kde je nádraží?“, dostaneme odpověď: „Jděte rovně a pak vlevo.“ a mluvčí pohybem nenaznačí, jakým směrem konkrétně „rovně“ je. Nebo konečně, když uslyšíme „Ukončete nástup a výstup, dveře se zavírají.“ mimo nástupiště nebo nějaký dopravní prostředek, a když stejnou větu uslyšíme při nastupování do vozu metra.

Ve všech případech jsme rozuměli, co tvrzení znamenají „obecně“, ale podle kontextu, ve kterém se vyskytly, jsme jim porozuměli specificky. To znamená, že kontext je zde chápán jako určitá situovanost textu v prostředí, jež poskytuje posluchačům nebo čtenářům vodítko k jejich interpretaci, k nalezení jejich smyslu. Zkoumání této situovanosti bychom mohli chápat jako ekologii textu, tzn. zkoumání vztahu textu a prostředí, kde prostředím rozumíme svět kolem textu včetně mluvčího/pisatele a posluchače/čtenáře. Vztah text – kontext je důležitou, i když ne jedinou složkou této ekologie textu. Hlavní kategorie, které vliv kontextu na význam textu popisují, jsou: odkaz, předpoklad, implikatura, situace, úsudek a archiv [Brown, Yule 1983].

*Odkaz* (reference) je takový vztah textu a prostředí, ve kterém mluvčí/pisatel prostřednictvím textu (slova, věty, promluvy) odkazuje na svět mimo text (na jeho prostředí, na jeho kontext) [Brown, Yule 1983: 28]. Důležité je uvědomit si, že tím, kdo odkazuje, není text samotný, ale mluvčí/pisatel. Uvedme si příklad: Bude-li věta „MÁM HLAD“ napsaná na kusu kartonu položeném na lavičce v parku, plní svou funkci lingvistickou – dokážeme ji interpretovat, ale nerozumíme sdělení, protože nechápeme její referenci; nevíme, *kdo* oznamuje, že má hlad. Když druhý den půjdeme kolem stejného místa a uvidíme tam skupinu bezdomovců, kolem kterých běhá voříšek se včerejší cedulkou pověšenou na krku, můžeme se tomu v duchu i zasmát, protože rázem pochopíme vztah textu k prostředí – ti lidé dali cedulku svému

psovi na krk, aby tím (humornou formou) sdělili, že má hlad nebo že je stále při chuti, nebo chtěli jen udělat vtip.

Jestliže jsme se voříškovi s cedulkou zasmáli, potvrdili jsme tím kontextové *předpoklady* (presupozice), se kterými text počítal, tzn. a) že psi mívají hlad, b) že psi neumějí psát, c) že dotyčný pes nejeví známky hladovění a d) že psovi dali cedulku na krk lidé, kteří psát umí. Když jsem napsal, že s předpoklady počítal text, tak by někdo mohl namítnout, že s nimi počítal jeho pisatel, nikoli text samotný. Lze ovšem namítnout, že pisatel byl žebrák, který den předtím cedulku používal k žebrání, a i když pravděpodobně by s předpoklady souhlasil, určitě je neměl při psaní textu jako kontext na mysli. Teprve až kolemjdoucí, který uvidí cedulku na psovi a zasměje se tomu, aktualizuje uvedené předpoklady. Tedy ačkoliv jsou předpoklady mimo text samotný, text s nimi počítá, dokáže je na sebe vázat. Kdyby na cedulce bylo napsáno „Alík“, žádný z uvedených předpokladů by čtenáři aktualizován pravděpodobně nebyl. Je zřejmé, že koncept předpokladu je komplexní, nicméně pro vztah textu a prostředí klíčový.

*Implikatura* je koncept, který vytvořil již dříve zmiňovaný P. Grice v souvislosti s pravidly konverzace [Brown, Yule 1983: 31–33; Grice 1989: 24 a násl.]. Pokud má posluchač dojem, že mluvčí porušil některé z konverzačních pravidel, třeba maximu „mluv pokud možno pravdivě“, má to důsledky pro interpretaci toho, co text doslovně říká. I když náš příklad s cedulkou není nejlepší příklad konverzace, můžeme ho přesto využít, abychom si ukázali, jak koncept implikatury popisuje vztah textu a prostředí. Implikaturu totiž můžeme chápat jako (záměrné i nezáměrné) porušení v dané situaci očekávaných předpokladů, které má interpretační důsledky. Kdybychom podle očekávaných předpokladů („mluv pravdivě“) interpretovali cedulku, tak bychom se textu nezasmáli, protože bychom usoudili, že pes má hlad, na čemž není nic k smíchu. Poněvadž ale víme, že zprávu nám nesděljuje pes, a protože si nelze představit, že cedulku napsal, a tudíž nemůžeme hledat komunikační záměr u něj, ale u toho, kdo mu jí na krk pověsil, nebudeme interpretovat text v doslovném významu. Shrnu-li to, porušení očekávaného předpokladu o vztahu textu a jeho prostředí, konkrétně že pravdivě komunikuje ten, kdo text prezentuje jako svůj (to by byl případ, kdy by cedulku měl žebrák, ale v případě psa to neplatí), vedlo k tomu (mělo implikaturu v tom), že čtenář začal hledat alternativní interpretaci, tj. a) vhodného pisatele textu, jehož b) předpokládaný komunikační záměr by dával textu v dané situaci větší smysl.

Po celou dobu, kdy na příkladu cedulky vysvětluji vztahy textu a prostředí, je čtenář tohoto textu informován o *situaci*, ve které k popisovanému komunikačnímu jednání dochází, jen velmi skoupě. Lavička v parku, skupina bezdomovců, voříšek s cedulkou s nápisem. Nicméně předpokládám, že



většinu čtenářů tento popis dostačuje, aby si danou situaci představili (alespoň pro účely porozumění tomuto textu). To znamená, že situační kontext je zde dvojitý: jednak spojený se situací, která je textem popisována (pes s cedulkou v parku), a jednak spojený se situací samotného čtenáře textu (čtouceho knihu vysvětlující situační kontext). V prvním případě se jedná o situační kontext v rámci posluchačem/čtenářem konstruovaného *možného světa* [Eco 2010: kap. 8], zatímco ve druhém případě se jedná o situační kontext aktuálního světa posluchače/čtenáře. Jiným příkladem by mohla být situace hladového člověka, který si čte povídku o bohaté hostině. Aby povídce porozuměl, musí si umět zkonstruovat hypotetický (možný) svět, v němž bohatá hostina může probíhat; ale zároveň se může rozhodnout, že povídku přestane číst, protože mu nedovolí zapomenout na vlastní hlad. K ekologii textu bychom řadili pouze vztah text – situace aktuálního světa, protože posluchačská/čtenářská konstrukce možného světa umožňující pochopit význam toho, na co text odkazuje (jak se to má se psem s cedulkou), je otázkou textové reference, o které byla řeč výše. V případě situačního kontextu se jedná o to, v jaké situaci je text produkován a konzumován, například ve výše uvedeném textu „Ukončete nástup a výstup, dveře se zavírají.“ Jestliže posluchač nenalezne ve svém okolí nic, do čeho by bylo možné vstoupit a vystoupit, nebo dveře, které by se mohly zavřít, bude se snažit na základě svých aktuálních situačních okolností nalézt pro text jako sdělení jinou interpretaci, než je ta doslovná; například metaforicky jako sociální kritiku (omezené sociální mobility) v rámci vystoupení punkové skupiny na jejím koncertě. Jak vyplývá z uvedeného příkladu, kontext sociální situace můžeme definičně vymezit souřadnicemi místa, času, mluvčího/pisatele, posluchače/čtenáře, jejich nelingvistického jednání a předchozích událostí, s tím, že a) se jedná jen o skutečnosti vnímané jako relevantní a b) ne všechny souřadnice musí být vždy ve hře.

Do kontextu, tj. interpretačně relevantního prostředí textu, můžeme možná překvapivě řadit i *usuzovací aktivitu* (inferenci) [Brown, Yule 1983: 33–35]. I ta je totiž součástí prostředí, nikoli textu, a bezprostředně ovlivňuje význam textu. O inferenční činnosti byla řeč výše [2.5.2], na tomto místě jen zdůrazním roli encyklopedické kompetence, která s inferencí bezprostředně souvisí. Obvykle je encyklopedická kompetence neboli obecné sociokulturní vědění považováno za nezbytné pro interpretaci textu. Slovy de Beaugranda „otázka, jak lidé vědí, o co v textu jde, je zvláštní případ otázky, jak lidé vědí, o co jde ve světě vůbec“ [cit. dle Brown, Yule 1983: 233; srov. též Beaugrande 1996: 26]. Takto pojatá inferenční aktivita je omezena na to, o co jde v textu; nicméně jak je zřejmé z výše probírané otázky kontextu, inference musí probíhat také na úrovni toho, jak text souvisí s tím, co se děje v dané situaci, již

je text součástí. Je rozdíl mezi tím, když začnu číst text, o kterém vím (a je jedno jak), že je zákonem dané země, a text, o kterém vím, že je sbírkou anekdot. Inferenční aktivita proto nesouvisí jen s tím, o co jde v textu, ale i o co jde v dané situaci, v níž s textem interaguji. Schopnost rozlišovat mezi texty, které je třeba interpretovat doslovně a které s nadsázkou, které brát vážně a kterým se smát, může být pro pochopení významu textu klíčová.

Tím se dostáváme k poslednímu vztahu textu a kontextu, kterým je to, co bychom s určitým přiblížením mohli nazvat *archiv*. Doposud jsme zaujívali perspektivu, v níž byl význam textu odhadován z jeho vlastností a z vlastností prostředí nebo přesněji řečeno, z interakce textu a prostředí, tzn. jeho kontextu. Text byl chápán jako faktický, zatímco prostředí – veškerý myslitelný svět mimo text – jako interpretační zdroj. Text „MÁM HLAD“ na jedné straně a cedulka na provázku, pes, bezdomovci, kolemjdoucí, park na straně druhé; text je aktér a prostředí jsou kulisy, ve kterých aktér hraje. To jsem nazval ekologií textu, nutnou situovaností každého textu v konkrétní konfiguraci světa kolem. Nyní však změníme perspektivu a hledejme, jestli je něco, co je všem nebo alespoň různým druhům situovanosti společné. Lze odhalit pravidla nebo pravidelnosti, které se budou vztahovat na interpretaci textu „MÁM HLAD“ napříč situacemi a rozmanitostí mluvčích/pisatelů, posluchačů/čtenářů? Je možné přejít od popisu a klasifikace situovanosti k popisu a klasifikaci situačního [Goffman 1964: 3; srov. též Kaderka 2013 a literaturu tam uvedenou]? O analýzu situačních pravidel používání textů (promluv) se pokusil M. Foucault ve své teorii řádu diskurzu [Foucault 1994/1971: 7–39; viz též kap. 4.2]. Tvrdí, že v každé situaci si uvědomujeme, že nemůžeme mluvit o čemkoliv a bez ohledu na okolnosti. Tím, co určuje zákazy, vylučování a celkově rozmístění (disperzi) produkce textů do jednotlivých situací, je specifický diskurz. Například diskurz lidství, který vyděluje lidi z oblasti ostatních zvířat, způsobuje, že někdo dá psovi na krk cedulku „MÁM HLAD“ a někdo se tomu zasměje. Zkoumání takového diskurzu není podle Foucaulta poměřování pravdivosti textů, není hledání obsahu (zjevného nebo skrytého) textu, ale zjišťování „hodnoty“ textu spočívající v jeho kapacitě někde být a někde ne, být v oběhu a být směřován za jiné statky (např. úctu, status, loajalitu); více o tom budeme mluvit ve čtvrté kapitole. Nyní ještě zbývá objasnit, proč je tento kontextový vztah zde označen termínem *archiv*. Archiv je pravidlo určující to, co může a smí být řečeno, které řídí objevování se promluv jako jedinečných událostí a uspořádává je do rozmanitých figur. Archiv je v tomto pojetí (archi)funkční kontext textu, který umožňuje jeho existenci a garantuje jeho smysl jako jedinečné výpovědní události [Foucault 2002/1969: 197–199].

## 2.6 Souhrn

V této úvodní kapitole jsem definoval text a jeho vztah k jiným textům, ke světu, ke čtenáři. V druhé části jsem popsal základní úrovně analýzy textů, které lze rozdělit na slovní, větné a nadvětné. Konečně na závěr jsem ukázal na roli kontextu. V následujících kapitolách se budeme k různým částem této úvodní kapitoly ještě vracet podle toho, jak budou jednotlivé probírané metody pojmově ukotveny. Pokud se tedy čtenář cítí dostatečně konceptuálně vybaven k tomu, aby přistoupil k výkladu obsahové, diskurzivní a narativní analýzy, kapitola splnila svůj účel.

## Kapitola 3. Obsahová analýza

Obsahová analýza je ze zde probíraných přístupů v sociálních vědách nejstarší a nejrozšířenější. Je popsána v řadě dostupných příruček a učebnic [Franzosi 2008; Krippendorff 2004; Neuendorf 2002; Weber 1990] a její standardní výklad by byl nošením dříví do lesa. V této kapitole se proto budeme zabývat jen základními předpoklady obsahové analýzy a specifickými aplikacemi rozlišující čtenářskou a strojovou perspektivu analýzy textů, zejména kvalitativní obsahovou analýzou založenou na postupu zakotvené teorie (Grounded Theory) na jedné straně a korpusovými přístupy na straně druhé. Budu se držet hlavně postupů, s nimiž mám vlastní zkušenosti, které považuji za inovativní a které jsou v současnosti použitelné v sociologickém výzkumu. Samozřejmě to neznamená, že neexistují jiné zajímavé a slibné aplikace obsahové analýzy.

Klasická Berelsonova definice zní, že obsahová analýza je „výzkumná technika k objektivnímu, systematickému a kvantitativnímu popisu zjevného obsahu komunikace“ [Berelson 1952: 18]. I když v současnosti se mnoho výzkumů k této definici v její úplnosti nehlásí, stále platí, že obsahová analýza by měla být objektivní, systematická a vztahovat se k obsahu komunikace nezávisle na její formě. Ponechme nyní stranou problematický požadavek objektivnosti a věnujme se pojmu obsahu, i když v rámci definice není nejexponovanějším konceptem a jako by se rozuměl samo sebou; z hlediska dnešní aplikace je jeho smysl zcela klíčový. Berelsonovská představa, že lidská komunikace je jednání, během něhož je pomocí komunikačního média (hlas, obraz, písmo apod.) přenášeno určité sdělení (obsah komunikace), které lze podrobit objektivnímu zkoumání, je totiž dnes neobhajitelná. Jak je patrné z předchozí kapitoly, sdělení nelze jednoduše oddělit nejen od komunikačního média, ale ani od účastníků komunikace. Z toho důvodu by bylo vhodnější mluvit o analýze komunikace bez toho, že by se analýza soustředila na její obsah. To je na druhou stranu až příliš široký pojem, který zahrnuje velkou a hlavně velmi různorodou část lidských aktivit. Sperber s Wilsonovou, vycházející z pojmu relevance, na základě analogie s pojmem pohybu dokonce pochybují, jestli je možné vytvořit teorii komunikace, která by podle

jednotného principu vysvětlovala tak odlišné komunikační jevy, jako je okázalé nadýchnutí se čerstvého vzduchu a právní text [Sperber, Wilson 1995: 1–4]. Proto se setrvává u původního termínu obsahová analýza jako zavedené nálepky; někteří autoři, aby se vyhnuli problému s obsahem, používají termín textová analýza, zejména tam, kde je analyzovaným prostředkem komunikace text [Popping 2000; Roberts 1997].

Pouhé vymazání pojmu obsahu z názvu nebo definice ovšem nestačí k tomu, abychom se s problémem obsahu komunikace vypořádali. Jak v každodenním, tak v odborném životě totiž intuitivně předpokládáme, že komunikace může mít nějaký obsah. Například když se ptáme, co píší v novinách nebo o čem je kniha, kterou právě někdo čte; nebo když se díváme na stránku nějaké knihy nazvané „Obsah“, abychom zjistili, jestli je pro nás zajímavá; nebo když hodnotíme nějaký text jako bezobsažný, ve kterém se „nic neříká“. Obsah je tedy něco, s čím běžně počítáme a jsme schopni ho identifikovat, resp. identifikovat jeho absenci. Je však také možné, že nás tato laická intuice klame a že představa obsahu je formou komunikačního fetišismu, tzn. mylné víry, že jádrem úspěšné komunikace je nějaký obsah (význam, myšlenka, představa, tvrzení apod.), který si komunikující předávají. V následující části se budeme otázkou obsahu komunikace zabývat podrobněji se zvláštním zřetelem na komunikaci textovou.

### 3.1 Co je obsah textu?

Existence obsahů knih a delších textů obecně, tzn. předpokládaná možnost, že text může být užitečně doplněn paratextem označeným „Obsah“, který explicitně stanovuje, co je „uvnitř“ textu, nás nutí nenechat problém obsahu textu stranou jako nedůležitý. Dokládá totiž, že něco jako obsah textu prakticky existuje, jinak by tento druh paratextu pozbýval smyslu. Současně však nesmíme jednoduše zaměňovat paratextuální „obsah“ s hypotetickým vlastním obsahem textu, protože – jak jsem ukázal v úvodní kapitole – role čtenáře je při interpretaci textových dat klíčová.

Otázku „Co je obsah textu?“ nevyřešíme jen pouhým přemýšlením o textech; je třeba si pomoci obecnější teorií komunikace a následně ji upravit na konkrétní případ textu. V první kapitole jsem se zabýval rozdílem mezi kódovým a inferenčním modelem komunikace [Sperber, Wilson 1995: 2].<sup>5</sup>

<sup>5</sup> Terminologie je zde neustálená; srovnej [Craig 1999]; Jiráková a Köpplová rozlišují přenosový a konstitutivní model komunikace [Jiráková, Köpplová 2009: 26 a násl.], které zhruba odpovídají tomu, co je zde označováno jako kódový a inferenční model.

Pro připomenutí, kódový model předpokládá, že lidská komunikace probíhá tak, že jedna osoba převede svou myšlenku do kódu a vytvoří tak sdělení („Kolik je hodin?“), které je prostřednictvím nějakého média (zvuku, obrazu) přeneseno k jiné osobě, jež sdělení dekóduje a zjistí sdělovanou myšlenku. Komunikace souhrnně probíhá takto: myšlenka A – převod do kódu – přenos médiem – převod z kódu – myšlenka A. Kódový model nelze považovat za univerzální model lidské komunikace, protože velká část komunikace takto neprobíhá a je mnohem komplexnější („Kolik je hodin?“ – „Tak ty mě nenecháš spát?!“); na druhou stranu vysvětluje důležitý rys komunikace, a to využití kódu a média. Druhý – inferenční – model komunikace předpokládá, že explicitní část komunikace probíhá podle kódového modelu, ale zároveň existuje stejně tak důležitá implicitní část komunikace (tj. to, co není přímo řečeno, ale z řečeného to v dané situaci vyplývá), která kódem určená není a závisí na usuzování (inferenci) komunikujících (viz výše uvedený příklad). Celkově tak můžeme komunikaci chápat jako *jednání* s komunikačním záměrem, kterým v dané situaci *vybíráme* (činíme zjevné) určité kódující i nekódující prvky *prostředí* (slova, zvuky, obrazy, pohyby apod.), o kterých *předpokládáme*, že z nich náš komunikační partner buď bude schopen *odvodit* to, co mu chceme sdělit (myšlenku, představu, pocit, výzvu apod.), nebo naopak, že s jejich prostřednictvím zjistíme, co chce sdělit komunikační partner nám [srovnej Luhmann 2006: 159 a násl.].

Lidská komunikace je často multimodální, tzn. jsou při ní využívány různé způsoby komunikace a nikoli jen jazyk [Kress 2010]. Sperber s Wilsonovou [1995: 25] uvádějí hezký příklad multimodálnosti komunikace, kdy na Petrovu otázku „Jak se dnes cítíš?“ odpovídá Marie beze slov vytažením krabičky aspirinu z kabelky. Jak je v této situaci zajištěno, že odpovědi bude možné správně porozumět? Pokud vezmeme v úvahu jen ostenzivní komunikaci, tzn. když jedna osoba dává svým chováním druhé osobě očividně najevo, že ji chce o něčem informovat, tak podle Sperbera a Wilsonové platí princip relevance. Ten formulují takto: každý akt ostenzivní komunikace současně komunikuje předpoklad své vlastní optimální relevance [ibid.: 158]. Jinak řečeno, komunikující své sdělení podává v takové podobě, o které očekává, že bude pro příjemce komunikace dostatečným stimulem, aby se významem sdělení zabýval, tj. shledal ho v dané situaci relevantním. V uvedeném příkladu Marie vytažením krabičky aspirinu předpokládá, že Petr pochopí relevanci tohoto gesta pro danou komunikaci a začne ho interpretovat jako odpověď na svou otázku.

Jak nám výše uvedené úvahy pomohou odpovědět na otázku o existenci a povaze obsahu textu, který předpokládá obsahová analýza? Následovně:

1. Protože každý text je součástí komunikačního jednání, je na jedné straně fixovaným výběrem prvků prostředí (zejména znaků), který učinil komunikující aktér se záměrem komunikovat určité sdělení, a na druhé straně předmětem usuzování jiného komunikujícího aktéra (čtenáře, např. analytika) se záměrem zjistit komunikované sdělení („Komunikace je koordinovaná selektivita.“ [Luhmann 2006: 176]). Obsah komunikace je proto vždy efektem (účinkem) individuální projekce určitého sdělení činná souborem voleb (znaků). Asi jako když ve výtahu zmáčkneme tlačítko se znakem „3“, abychom tím informovali stroj o našem požadavku vyjet do třetího poschodí. Obsah se proto nachází mimo komunikaci, není předmětem komunikace, ale je komunikací generován. Vytváření textu, stejně jako čtení textu tak obsah generuje, resp. generuje jeho absenci.
2. Text je vytvářen i čten s ohledem na jeho manifestní i implicitní (konotovanou) složku [Eco 2010]. Manifestní složka je předmětem kódování a dekódování, implicitní složka je předmětem usuzování (Proč je napsáno zrovna tohle a ne něco jiného? Co se tím chce říci?). Obsah textu je proto účinkem obou těchto složek. Navíc, i když samotný text je nemateriální a nesituovaný, vyskytuje se vždy v materiálním a situačním kontextu (a zpravidla je doprovázen určitým kotextem). Proto i obsah je kontextem, v němž je text produkován nebo konzumován, spoluurčován, zejména jeho implicitní složka. Obsah textu, který je nadepsán *Můj tajný deník* (kotext) a nalezneme ho v truhle po babičce (kontext), bude čtením vytvářen v tomto kotextuálním a kontextuálním rámci. Z toho vidíme, že obsah textu je komplexním účinkem psaní, resp. čtení textu nějakým aktérem.

Doposud jsme se pohybovali v oblasti teorie lidské komunikace a nedbali jsme sociologické relevance pojmu obsah textu. Je ze sociálněvědního pohledu důležité, je-li obsah textu inherentní vlastností textu nebo jeho účinkem? Vždyť důležité snad je, zda obsah textu (komunikace) existuje a je přístupný analýze, nikoli jeho původ. Ba ne, uvědomění si původu obsahu, jeho pragmatický (činnostní) původ, je pro sociálněvědní analýzu klíčový, protože znamená: kolik je různých zacházení s textem, tolik je různých obsahů textu. Obsahová analýza je vhodný název nikoli z toho důvodu, že by dokázala odhalit „objektivní“ obsah textu, ale že dokáže systematicky vytvářet obsah textu (zejména z jeho manifestní složky, ale nikoli výhradně). Krippendorff navrhuje definici, která tyto podmínky splňuje, ale navíc zdůrazňuje důležitost

kontextu užití textů. „Obsahová analýza je výzkumná technika pro vytváření replikovatelných a validních úsudků (inferences) z textů (nebo jiných smysluplných věcí) na kontexty jejich užití“ [Krippendorff 2004: 18].

Všimněme si, že Krippendorfova definice obsahové analýzy přenáší explanans (to, co chceme analýzou vysvětlit) z textů na jejich užití. Reflektuje tím pragmatický obrat v teorii komunikace, který přenesl pozornost od studia domnělých vnitřních charakteristik komunikace k vnějším způsobům používání komunikace – teorii mluvních aktů (speech act). Nebudu zde tuto problematiku hlouběji rozebírat a odkážu čtenáře na první kapitolu, kde se o tom píše. Přenos důrazu na užití textů (co se texty „dělá“) otázku obsahu prakticky nemění. Protože nic jiného než interpretace se s texty dělat nedá [Wilson 2012]; pokud z novin udělám vlašťovku a hodím jí z okna, nepoužívám text, ale novinový papír. Jakmile si noviny začnu číst, již generuji jejich obsah a na otázku „Co v nich píšou?“ jsem připraven odpovědět. Když píše „Slibuji...“, výběrem slova *slibuji* se snažím komunikovat svůj závazek, protože v tom slově je, jak předpokládám, můj závazek pro daného komunikačního partnera obsažen. Dovolím si proto upravit Krippendorfovou definici tak, aby byla jasnější: Obsahová analýza je výzkumná technika pro vytváření replikovatelných a validních úsudků z textů (nebo jiných smysluplných věcí) na kontexty vytváření jejich obsahů.

Pro ještě větší jasnost definice bychom ještě mohli nahradit termín kontexty slovem okolnosti. Lze totiž říci, že vytváření obsahů má určitý kontext, ale protože tento pojem je zatížen mnoha diskuzemi o jeho významu, jednodušší ho bude nahradit termínem okolnosti, který nevyžaduje takovou interpretaci. Okolnosti vytváření obsahů mohou být fyzické i symbolické.

### 3.2 Kódování

Jestliže jsme si na základě rozboru pojmu obsahu textu definovali obsahovou analýzu jako metodu analýzy textů, jejíž pomocí můžeme vztahovat texty k okolnostem vytváření jejich obsahů, nyní přistupme k otázce, jakým způsobem analyzovat texty, abychom odhalili možnosti vytváření jejich obsahů. Vlastně se tím dostáváme k základní otázce textové analýzy, tj. otázce významu. Nebude nás však zajímat to, jak je v textu význam obsažen, protože touto otázkou jsme se zabývali v předchozí části, ale to, jak transformovat analyzovaný text, jehož účinkem jsou nějaké obsahy, do analytického textu, který reprezentuje (re-konstruuje) obsahy analyzovaného textu. Jednoduše řečeno, jak textově postihnout obsahy, které analyzovaný text nebo texty vytvářejí?

Nástrojem, který se k této transformaci v obsahové analýze používá, je kódování. Jeho klíčová role je uznávána v kvalitativních verzích obsahové analýzy [Seidel, Kelle 1995], méně už v kvantitativních přístupech. Příznačná je Krippendorfova kvazisynonymizace kódování se zaznamenáváním [kapitola Recording/Coding v Krippendorf 2004], sugerující, že kódování je v obsahové analýze jen jakési dokumentování, jehož interpretační aspekt je vzhledem analyzovanému textu minimální anebo přísně kontrolovaný: jde přece o to, „zaznamenat“ přirozené obsahy v pojmech jazyka dat („data language“) [ibid. 125]. „Zaznamenávání (recording) probíhá, když pozorovatelé, čtenáři nebo analytici interpretují to, co vidí, čtou nebo nacházejí, a následně vyjadřují své zkušenosti ve formalizovaných pojmech analýzy; kódování (coding) je termín, který obsahoví analytici používají, když proces zaznamenávání probíhá v souladu s na pozorovateli nezávislými (observer-independent) pravidly“ [ibid. 126]. Celý problém zaznamenávání-kódování se pak přenáší na otázku kvality (lidských) kodérů, tzn. jak je vycvičit, aby kodovali validně a reliabilně.

Termín „kódování“ implikuje nějaký sémiotický systém kódů, do kterého zkoumaný text překládáme. V kvantitativní obsahové analýze to může platit tehdy, když máme systém kódů ve formě vzájemně propojených škál (přítomnost – absence; přítomnost: pozitivní – neutrální – negativní; přítomnost pozitivní: minulá – současná – budoucí; atd.), ale většinou jde o pouhou transformaci dat pro účely statistického zpracování přiřazováním předem definovaných kódů k určitým úsekům textu bez předpokladu koherentního systému kódů. Kódování v obsahové analýze si proto nesmíme spojovat s kódovacím modelem komunikace, protože je mnohem jednodušší.

V principu existují dvě funkce kódů [Seidel, Kelle 1995]. *Faktická* funkce, kdy kód označuje nějakou skutečnost v textu jednoznačným způsobem; ukazuje, že něco je případem něčeho obecnějšího. Například studujeme-li líčení vzdělávací dráhy v biografických rozhovorech a hledáme všechny případy, kdy bylo vzdělání ovlivněno nestandardními procedurami (politický tlak, protekce...), kód s faktickou funkcí zastupuje to místo v textu, kde se tyto důvody uvádějí. Například v orálně-historickém rozhovoru nalezneme tuto pasáž:

*A tak si popovídali a ředitel říká tomu mistrovi mému a co ty honíš? A on říká, ale mám žáka, který se sem přihlásil na přijímací zkoušky a tak by mě zajímalo, jak dopadl. A on říká, tudle to jsou určitě přihlášky, tak je probereme. Tak probrali sloupec přijatých přihlášek, kde jsem nebyl, tak říká, no tak to je nepřijaté. Tak probrali štos nepřijatých přihlášek, samozřejmě mě tam našel, vytáhl nepřijatou, přehodil do sloupce přijatých, asi to ještě neměli, sekretářka ještě neměla seznamy a já byl přijat na průmyslovou školu. (ID093).*

Uvedený úsek lze označit kódem „nestandardní procedury“ a dále lze už pracovat se samostatným kódem – kód reprezentuje jev v textu. Kódy s primárně faktickou funkcí lze dále třídit a shlukovat do kategorií, například akteři muži nebo ženy, druh školy, pomoc při přijetí nebo naopak bránění v přijetí apod.

Druhým typem funkce kódu je *referenční* funkce. Kód v tomto případě označuje určitou část textu relevantní vzhledem ke zkoumanému jevu nebo přinejmenším hypoteticky relevantní. Například ve výše uvedeném úryvku vypravěč mluví o průmyslové škole a jakou roli v přijetí na ni hrály známosti. V případě, že bychom zkoumali vzdělávací dráhy vypravěčů, tento úsek vyprávění bychom označili hned několika kódy s referenční funkcí, například průmyslová škola, nestandardní přijetí, známosti, ředitel, mistr, protekce, přiznání, úspěch. Referenční funkce znamená, že kód daný úryvek nezastupuje, pouze na něj odkazuje, asi jako číslo v telefonním seznamu odkazuje na danou osobu, a chceme-li zjistit, kdo je zač, musíme jí zatelefonovat. Proto u kódů s primárně referenční funkcí musíme pracovat stále s původním úryvkem, nikoli kódem, který na něj odkazuje.

V kvantitativně orientované (hypoteticko-deduktivní) obsahové analýze je cílem dosáhnout maximální faktické funkce kódování tak, abychom mohli operovat s kódy bez obavy, že nemáme jasně definovaný sémantický vztah mezi kódem a textem, který zastupuje. Kódy si teoreticky ponechávají svou referenční funkci, nicméně prakticky už nikdo nekontroluje, jestli jednotlivé instance kódovaných úryvků odpovídají přiřazenému kódu. Daleko častěji se ověřuje reliabilita kódování, tzn. zda různí kodéři kódují stejně. Jak podotýká Seidel s Kellem, faktické kódování předpokládá jasný konceptuální rámec, ve kterém se kódy pohybují [Seidel, Kelle 1995: 54]. Kód „nestandardní procedura“ musí mít jednoznačnou definici, v tomto případě komplementární, tj. vymezení, co jsou standardní procedury přijetí na školu. Pro výzkum by bylo pohromou, kdyby někteří kodéři považovali politická kritéria přijetí za standardní a jiní za nestandardní.

Při kvalitativní obsahové analýze je kladem důraz na referenční funkci kódů, a to především z toho důvodu, že kódy vznikají ad hoc „za pochodu“ a tudíž výzkumník nemůže garantovat jejich kvalitu. Kódy zde představují heuristickou pomůcku ke konstruování kategorií. Z logických důvodů zde mnohem více hrozí záměna „slabších“ referenčních kódů za „silnější“ faktické. Seidel s Kellem [Seidel, Kelle 1995] uvádějí dvě rizika plynoucí z této záměny, která připadají v úvahu zejména tam, kde je ke kódování využíván počítačový software. První riziko je zvěcnění kódu, tím i ztráta opory v datech; to, co jsme si provizorně označili za (hypoteticky) relevantní vzhledem k určité kategorii nebo jevu, začneme považovat za případ toho jevu, aniž by-

chom tento vztah zpřesnili. Ve výše uvedeném případě bychom za nežádoucí zvěčnění mohli považovat to, kdybychom s kódem „známosti“ začali pracovat jako s fakticky kódujícím namísto pouhé referenční funkce. Vždyť jak můžeme s určitostí tvrdit, že přijetí bylo ovlivněno známostí? Vždyť žádnou definici „známosti“ nemáme v dané chvíli k dispozici. Pokud bychom chtěli funkčně kódovat kódem „známosti“, nejdříve bychom ho museli definovat ve vztahu k podobným jevům, například vztahům nadřizenosti a podřizenosti apod. Následně bychom museli všechna data takto nově definovaným kódem překódovat. Tento postup je legitimní, ale dosti náročný. Druhým rizikem záměny referenční funkce kódu za faktickou je ztráta kontextu. V případě referenčního kódování označujeme oblasti textu za relevantní vzhledem ke kontextu aktivovanému naším čtením, který bývá značně rozsáhlý. V některých případech je hypotetická relevance vytvořena jen pozicí výpovědi v řadě jiných výpovědí (Proč zrovna zde mluví o přijetí na školu?). Pokud kód, například úspěch, faktualizujeme a tím dekontextualizujeme, ztrácíme tak podstatnou část jeho relevance. V důsledku toho nebudeme schopni dospět k validní interpretaci kódu a ke konstrukci smysluplných, na datech založených kategorií.

Referenční kódování samo o sobě není schopno generovat obsah (kromě specifických případů, o nichž bude řeč níže); je jako rejstřík v knize, který nás může informovat, o čem všem se v ní píše, ale už z něj nezjistíme, o čem kniha je – jaké je její téma a její zápletka nebo důvod jejího napsání. Proto referenční kódování musí vyústit v konstrukci kategorií nebo, jak bylo zmíněno výše, ve vytvoření systematického a koherentního kódovacího rámce s faktickou funkcí, kterým se text znovu kóduje. Běžně se ovšem přistupuje k tvorbě kategorií, při které jsou referenční kódy používány jako pomůcka. Nejpropracovanějším způsobem tvorby kategorií z referenčních kódů je postup zakotvené teorie, jemuž se budeme věnovat v další části této kapitoly.

Shrnu-li dosavadní úvahy o obsahu a kódování v obsahové analýze, ukazuje se, že před výzkumníkem stojí následující výzvy:

1. analýza by měla vytvořit takový obsah textu, který odpovídá jak textu, tak okolnostem, v nichž byl text vytvořen anebo v nichž je obvykle čten;
2. analýza by měla postupovat prostřednictvím kódování, které v různé míře kombinuje referenční a faktickou funkci; referenční funkce indexuje text, zatímco faktická vytváří jeho obsah.

Podle odpovědí na tyto výzvy můžeme definovat různé přístupy k obsahové analýze:

Tabulka 2 Přístupy k obsahové analýze

předmět vysvětlení	převažující funkce kódování	
	faktická	referenční
okolnosti tvorby textu	deskriptivní přístup (explorační/ konfirmační)	hermeneutický přístup
okolnosti čtení textu		sémiotický přístup

Rozdělení přístupů podle toho, jestli se zaměřují na vysvětlení okolností tvorby textu, nebo čtení textu, není obvyklé, ale podle mého názoru je opodstatněné. Při využití faktického kódování se plně neukazuje, protože tento typ kódování není schopen odlišit proces konstruování textu od jeho čtení (vypravěčem líčené přijetí na školu je kódováno jako nestandardní procedura bez ohledu na to, jestli je jako taková konstruována vypravěčem nebo čtenářem). Když však využijeme referenční funkci kódování, musíme se vždy ptát, zda nás zajímá způsob, jakým byl obsah textu vytvářen mluvčím, nebo způsob, jak obsahu textu rozumí čtenář. V prvním případě jsem označil přístup za hermeneutický, protože se snaží vyložit text z pozice jeho autora, tzn. jaký obsah chtěl autor textem vytvořit. Když je analýza zaměřena na čtení textu, pak referenční kódování vede k sémiotickému přístupu, tzn. jak čtenář kooperuje s textem, aby vytvořil jeho obsah.

Následující část kapitoly je rozdělena na přístupy předpokládající lidského čtenáře a na ty, které si vystačí se čtenářem strojovým.

### 3.3 Lidský čtenář

Lidské čtení textu je komplexní sémiotická aktivita. Jak obrazně tvrdí Eco [2010], „text přenechává interpretační iniciativu na čtenáři, třebaže si obvykle přeje být interpretován s co největší jednoznačností. Chce, aby mu někdo pomáhal fungovat“ [Eco 2010: 68]. Čtenář interpretuje text v několika úrovních současně, od základní lexematické úrovně (jaká slova jsou v textu obsažena) přes úroveň encyklopedickou (co tato slova znamenají v kontextu kultury, jíž je čtenář členem) až po úroveň konstrukce možných světů, tzn. hypotetických konfigurací aktérů, objektů a situací, ve kterých interpretovaný obsah textu dává smysl (např. když čteme o teroristickém

útoku v Severním Irsku, abychom textu porozuměli, představujeme si možnou organizaci aktérů, objektů i situací, za kterých se útok odehrál; jsme ochotni připustit, že policie zasáhla střelbou z pušek, ale těžko bychom přijali, že útočníky zahlnala luky a šípy; v takové případě bychom hledali jinou interpretaci celého textu, např. jako aprílové vydání novin). Když čte text stroj, k interpretaci textu nedochází, stroj pouze redukuje komplexitu textu podle určitého, předem zadaného algoritmu. Například při sledování spoluvýskytů klíčových slov převádí stroj text na numerickou matici četností spoluvýskytů klíčových slov, z níž může výzkumník obsah textu vytvářet. Pokud budou stroje disponovat kulturními znalostmi (encyklopedickou kompetencí), bude rovněž možné, aby text interpretovaly i na úrovni encyklopedické; například bude-li text obsahovat výraz „tak jsme se kapku pohádali“, bude stroj schopen nejen rozpoznat, že slovo *kapka* nese v textu význam „trochu“ a *pohádali* znamená „verbálně komunikovali vzájemné neshody“, ale zároveň dokáže hledat v nejbližším okolí i příčinu nebo důsledek hádky (rozdílné názory na výchovu dětí). K využívání encyklopedických znalostí (databázového zpracování již existujících textů) již dochází při strojovém generování obsahů textů, například při automatickém detekování nevyžádané elektronické pošty.

Sémiotická komplexita procesu čtení není to, co v principu lidské a strojové čtení odlišuje. Podle mého názoru je podstatný rozdíl v tom, že lidský čtenář vztahuje čtením aktualizované významy a konstruovaný obsah textu k vlastní identitě a to ovlivňuje následné čtení i interpretaci textu. Pokud MNE text nezaujme, čtu ho dál jen povrchně; pokud (dojdu k závěru, že) MNE uráží, přestanu ho číst; pokud (zjistím, že) MI lichotí, čtu text s potěšením třeba i opakovaně. Lidé čtou text, protože ho chtějí číst jako součást svého života a očekávají, že text jim osobně něco přinese (zábavu, poučení, emoce apod.). Proto může malá část textu, která rezonuje s identitou čtenáře, ovlivnit čtení a interpretaci celého textu. Strojové čtení tuto vlastnost nemá, protože stroj nečte text pro sebe, ale pro lidského operátora. Stroj se sice čtením může proměňovat (získávat nové informace do databáze a přehodnocovat již uložené informace) podobně jako člověk, ale nevztahuje text ke své identitě jako klíčovému aspektu interpretace textu. To je jeho silnou i slabou stránkou při využití strojového čtení pro analýzu textu v sociálních vědách.

Pro bližší seznámení s obsahovou analýzou jsem vybral tři přístupy – dva lidské (obsahovou analýzu v rámci zakotvené teorie a analýzu sémiotickou) a jeden strojový (sledování spoluvýskytů klíčových slov).

### 3.3.1 Přístup zakotvené teorie

Grounded Theory (zakotvená teorie) je současně metodou výzkumu i výsledkem výzkumu; sestává ze souboru pravidel, jak vést induktivně založený výzkum, jehož cílem je vytvoření teorie, nikoli pouze empirické generalizace nebo tvrzení o přijetí či odmítnutí teorie stávající. Zakotvená teorie jako metoda se zaměřuje zejména na analytickou část výzkumu, to znamená na analytickou a interpretační práci s daty, ale významným způsobem ovlivňuje i přípravu výzkumu a sběr dat. Hlavním pilířem metody a vodítkem pro výzkumníky je neustálé vzájemné srovnávání (dat, terénu, interpretací, konceptů) a interaktivita (výzkumníka s terénem, s jinými lidmi, koncepty, s jinými výzkumníky).

Metodologii zakotvené teorie vytvořili v 60. letech minulého století dva sociologové, Barney Glaser a Anselm Strauss. Oba už předtím spolupracovali na výzkumech sociální organizace umírání v nemocnici, což je pro výzkum velmi náročné téma a má přesah i do praxe – jak lépe zorganizovat péči o pacienty v terminálním stádiu nemoci. To je i v současnosti velmi diskutovaná otázka. Pro takový náročný výzkum potřebovali vyvinout metodologii, která by odpovídala povaze kvalitativního výzkumu, tzn. aby byla současně systematická, rigorózní a flexibilní. Svou představu takovéto metodologie ztělesnili v knize *Discovery of Grounded Theory* [Glaser, Strauss 2006/1967]. Grounded Theory měla kvalitativní výzkum vymanit z tehdejší podřadné pozice, kterou měl vzhledem k výzkumu kvantitativně orientovanému. Její autoři k tomu měli všechny předpoklady. Glaser studoval na Kolumbijské univerzitě, kde v té době působil například P. F. Lazarsfeld a kde se zaměřovali na rozvoj sofistikovaných kvantitativních metod. Ze studia si odnesl přesvědčení, že dobrá sociologická metodologie musí být dobře promyšlená, jasně formulovaná a systematicky zpracovaná. Zejména ho trápila otázka, jak testovat pracovní hypotézy vzniklé z kódování. A. Strauss studoval na Chicagské univerzitě, která byla vždy baštou kvalitativního terénního výzkumu (ostatně ne nadarmo se v dějinách sociologie učí o chicagské škole sociologického výzkumu). Strauss byl ovlivněn zejména pragmatismem a symbolickým interakcionismem, ze kterého si vzal zaměření na výzkum v terénu, nutnost pevného empirického ukotvení každé teorie, důraz na proměnlivost sociálního dění a důležitost symbolického rozměru lidského jednání. Když jejich kniha *Discovery of Grounded Theory* vyšla, umožnila kvalitativně orientovaným výzkumníkům vést výzkum takovým způsobem, že nemohli být obviňováni z nevědeckosti, nesystematičnosti a jejich závěry nebylo možné odmítnout mávnutím ruky z důvodu, že pocházejí z metodologicky nezvládnutého výzkumu. Ironicky se časem stala jakousi legitimizační

záštitou pro celou řadu výzkumů, které postupy zakotvené teorie nepoužívaly, ale jejich autoři se domnívali, že jim tato metoda dodá punc vědeckosti [srov. Becker 1993].

Později došlo k rozdělení zakotvené teorie do dvou verzí, jedné zastávané Glaserem a druhé Straussem. Hlavní rozdíl mezi nimi spočíval v důrazu, který kladl Glaser na induktivní přístup [Heath, Cowley 2004; Walker, Myrick 2006]. Výzkumník by měl přicházet do terénu bez předem nastudovaných vědomostí, věnovat se sbírání všech dostupných dat a ty pak podrobně – řádek po řádku – analyzovat s vědomím základních sociálních procesů. Cílem je vytvoření konceptů, které povstanou z analýzy dat, a jejich vztahy pak umožní zkonstruovat zakotvenou teorii [Glaser 1978, 1998, 2001]. Strauss spolu s Corbinovou se spíše přikláněli k méně ortodoxní verzi, kdy výzkumník má vstupovat do terénu a posléze i analyzovat data se znalostí prostředí a již existující odborné literatury o tématu, které chce studovat [Corbin, Strauss 1990; Strauss, Corbin 1997; Strauss, Corbinová 1999]. Navzdory odlišnostem základní princip zakotvené teorie, tzn. vyvíjení teorie z dat a její těsná vazba na empirická data, zůstává stále platný.

### 3.3.1.1 Kódování

Kódování je jedním ze dvou klíčových analytických kroků v rámci budování zakotvené teorie. Druhým klíčovým analytickým krokem je vytváření kategorií a vztahů mezi nimi. Připomeňme si, jak probíhá analýza textů v běžné kvalitativní obsahové analýze. Výzkumník je postaven před množství přirozeně strukturovaných dat (dokumenty, přepisy rozhovorů, zápisy pozorování apod.). Cílem analýzy, aby to byla opravdu analýza a nikoli jen popis, je transformace přirozeně strukturovaných dat do interpretace jevu v odborném kontextu. Například jsme svědky pokřikování fotbalových fanoušků pražské Sparty na utkání s ostravským Baníkem: „Vy ste špína naší ligy – vy ste přece polský Židi!“ Rozumíme tomu, chápeme, že cílem je zesměšnění, a to že je činěno pomocí slov *špína*, *polský* a *Židi*. Dále nad tím nepřemýšlíme, jen si řekneme, že to dokazuje otevřenou xenofobii a rasismus fanoušků, a někteří z nás si pomyslí, že by se s tím (xenofobií a rasismem) mělo něco dělat. Pokud ovšem chceme chování fanoušků analyzovat, nestačí pokřiku porozumět v jeho přirozeném kontextu a k tomu přidat hodnotící soud. Analytik musí nejprve znesamozřejmit samozřejmý jev pokřikování na stadionu a vysvětlit ho z odborného pohledu v kontextu sociálněvědních poznatků a teorií. Jinými slovy, musí poskytnout jiné vysvětlení, než že pokřik má zesměšnit druhý klub a činí to pro danou situaci hanlivými výroky, a toto vysvětlení musí odborně argumentovat v kontextu dosavadního disciplinárního vědění (např. o sociálním konfliktu).

V obsahové analýze je nástrojem k transformaci přirozeně se vyskytujícími dat do odborného kontextu právě kódování. Pomocí kódování, resp. jeho prostřednictvím, hledáme v textech typy, příběhy, témata, vztahy, základní kategorie atd. Kódování není nikdy cílem samo o sobě, ale je prostředkem k vytvoření kategorií či teorie určitého jevu.

V postupu budování zakotvené teorie rozvinutém Straussem a Corbinovou existují tři hlavní typy kódování: otevřené, axiální a selektivní [Strauss, Corbinová 1999; pro srovnání s Glaserovým přístupem viz Kelle 2005]. Hranice mezi typy kódování jsou však při samotném výzkumu nezřetelné, protože v průběhu analýzy kódujeme tak trochu všemi způsoby najednou: jsme stále otevřeni novým kódům, pokud by nám přišly vhodné, stále přemýšlíme, jak bychom kódy mezi sebou pospojovali do kategorií, a současně máme na mysli nějakou klíčovou kategorii, ke které se celá naše interpretace vztahuje.

#### 3.3.1.1a Otevřené kódování

Otevřené kódování je v postupu zakotvené teorie nejčastěji tím analytickým krokem, kterým začínáme. Máme před sebou nasbíraná první data a výzkumnou otázku, jež nás při sběru dat vedla. Otevřené kódování znamená, že označujeme ty kusy dat, například části textu, o nichž se domníváme, že by mohly být pro zkoumaný jev nějakým způsobem relevantní. Vezměme si výše uvedený pokřik: „Vy ste špína naší ligy – vy ste přece polský Židi!“ Označit můžeme jednotlivá slova: *vy* – adresát pokřiku, *špína* – hanlivé označení adresáta, *naší ligy* – skupinová identita, *prece* – rétorická pomůcka, *polský* – etnické označení adresáta, *Židi* – „rasové“ označení adresáta. Všechna slova jsou zde relevantní, i když slovesa jsme samostatně nekódovali, protože jejich funkce je zde gramatická. Označit můžeme také celou větu jako zesměšňující a ponižující pokřik. Prvním krokem je tedy označení relevantních úseků textu kódy (pojmy), které reprezentují významy. Umění otevřeného kódování spočívá ve schopnosti zachytit co nejvíce konkrétních informací co nejmenším počtem kódů. Všimněte si, že z deseti slov jsme získali šest různých kódů. Když bychom takto kódovali biografický rozhovor, který má tisíce slov, dostali bychom stovky, ne-li tisíce kódů. A co pak při analýze deseti biografických rozhovorů? Proto se musíme vyvarovat obou krajností: mít kódy příliš konkrétní, které výstižně reprezentují jevy, ale které se v datech prakticky neopakují a mezi kterými se jen těžko hledají obecnější vztahy, nebo mít kódy příliš obecné, které jevy reprezentují jen volně, ale zase se z nich dobře vytvářejí kategorie. V našem případě by příliš konkrétní kód pro slovo *špína* bylo „označení adresáta jako špína“ a naopak příliš obecný by byl kód „označení adresáta“, který by nediferencovaně sdružoval slova *špína*, *polský*, *Židi*.



Poté, co si data okódujeme, začneme z kódů vytvářet kategorie. Kategorie se vytvářejí seskupováním kódů. V našem případě bychom vytvořili kategorii „označení adresáta“, která by mohla mít podkategorie hanlivé, pochvalné a neutrální. Dále bychom mohli vytvořit kategorii „skupinová identita“, která by byla totožná se stejně nazvaným kódem. Smyslem vytváření kategorií je jednak snížení počtu kódů do zvládnutelného počtu, jednak vytvoření takové analytické úrovně, na níž bychom mohli sledovat dimenze pojmů a určit/navrhnout teoretické vztahy mezi nimi.

Co jsou to dimenze kategorií? Například označení adresáta může mít dimenze pozitivní až negativní, ale může mít též dimenze stigmatizující – nestigmatizující (*buzeranti/debilové*), přímé – nepřímé (*debilové – Pražáci*), etnické, rasové, genderové, zdravotní atd.

Vztahy mezi kategoriemi a jejich dimenzemi je potřeba definovat. Například skupinová identita a označení adresáta. V našem výroku je skupinová identita kategorií zahrnující kategorii označení adresáta. Pokřik řadí adresáta, tj. tým Baníku a jeho fanoušky, do „naší ligy“, ale v jejím rámci mu dává nízký status, který byl vyjeven („přece“; buď: vždyť to z vás na první pohled čiší, nebo: navzdory tomu, že tak na první pohled nevypadáte). V jiném pokřiku by adresát mohl být považován za nepřítele, někoho cizího, dvě identity by stály proti sobě, místo aby byly v hierarchickém vztahu.

Strauss s Corbinovou nabízejí několik technik, jak konstruovat kvalitní kódy a kategorie. Nejprve upozorňují, že překážkou může být naše předporozumění zkoumaným jevům, které způsobí jakousi analytickou slepotu, neschopnost rozpoznat v datech relevantní informace. Když fanoušci společně skandují „Smrt rozhodčímu!“, jak takový pokřik kódovat? Jako výhrůžku? Jako rituální pokřik, jehož význam je ve vyjádření nesouhlasu s rozhodovacím zápasem? Běžný člověk totiž chápe výhrůžku smrtí jako velmi vážnou, prakticky nejvážnější. Ale jsou situace, v nichž se až tak vážně nebere, například když někdo zapomene něco pro druhého udělat, ten mu odpoví: „Já tě zabiju!“. Proto, abychom se stali analyticky citlivější, můžeme zkusit hledat protiklady, abychom zjistili, na jaké škále se pohybujeme. Co se týče pokřiků, existují pokřiky, že někdo „je bůh“, „je král“ apod. Jejich nerealističnost nám umožňuje vzít v úvahu i to, že výhrůžka smrtí na stadionu při utkání je chápána jinak, než když by takto pokřikovali fanoušci na rozhodčího před jeho domem. Jinou technikou získání analytické citlivosti je porovnávání dvou nebo více jevů. Například můžeme srovnat pokřikování u různých druhů sportů, u různých skupin, válečné pokřiky atd. Toto srovnání může ukázat, co je na pokřiku fanoušků obecné a co specifické.

### 3.3.1.1b Axiální kódování

Pokud máme dostatek kategorií, můžeme začít s axiálním kódováním. Jestliže smyslem otevřeného kódování je rozložit přirozenou strukturu dat na soubor kategorií, které tento smysl zachycují, ale nejsou už přímo spojeny s konkrétními daty, cílem axiálního kódování je složit kategorie tak, aby vytvořily strukturu odborného vysvětlení jevu. Můžeme to přirovnat ke zkoumání nějakého přístroje, například mobilního telefonu. Otevřené kódování je jako rozzebírání (je otevřené a zároveň otevírací), kdy jednotlivé součástky rozdělujeme do kategorií podle různých kritérií. Axiální kódování je naproti tomu aktivita syntetická, kdy se snažíme porozumět vztahům mezi kategoriemi.

Protože svět se neodehrává náhodně (i když někteří filozofové tvrdí opak) a ani sociální dění není libovolné, navrhuji Strauss s Corbinovou obecný model pro hledání vztahů mezi kategoriemi, který odpovídá sociálnímu dění. Tento model nazývají paradigmatický model (paradigm model) neboli vzorový model. Model vypadá takto [Strauss, Corbinová 1999: 72]:

Tabulka 3 Paradigmatický model v zakotvené teorii

příčinné podmínky	kontext		následky
	jev	strategické jednání a interakce	
intervenující podmínky			

Účelem modelu není analýzu sešněrovat, ale nabídnout obecný model pro systematické vysvětlení sociálního dění, který umožňuje kategorie a jevy vzájemně mezi sebou vztahovat složitými způsoby. Abych axiální typ kódování vysvětlil, použiji opět stejný příklad pokřiku fotbalových fanoušků. Bez paradigmatického modelu bychom mohli poskládat kategorie do vzájemných vztahů (např. označení ke skupinové identitě), ale vše by bylo statické. Víme totiž, že sociální dění je neustálý proces, který vnímáme jako proud zkušeností. Model nám pomáhá tuto procesualnost dění zachytit tím, že nás nutí hledat ty aspekty jevu, které jsou před ním, po něm a po „stranách“. Zde nabídnutá ilustrace bude ovšem jen špatnou simulací opravdového axiálního kódování, protože nemáme k dispozici potřebná data.

Zkoumáme pokřiky fanoušků. Naším jevem je pokřikování fanoušků Sparty Praha na hráče a fanoušky Baníku Ostrava. Nuže, hledejme *příčinné podmínky* – co vede fanoušky k tomu, aby na stadionu pokřikovali, že ti druzí jsou špína ligy. V rozhovorech s fanoušky bychom zjistili, co si myslí o fanoušcích jiných týmů, co si myslí o ligových hráčích, proč chodí fandit, jaké jsou druhy fanoušků a ke kterým se řadí. Z toho bychom mohli pomocí otevřeného kódování získat poznatky o identitách jiných aktérů a mohli by-

chom srovnávat s tím, co se křičí. Zda je příčinou zrovna takového pokřiku přesvědčení, že hráči a fanoušci Baníku jsou špina ligy, nebo jestli je cílem snížit sebevědomí protivníka a zvýšit vlastní sebevědomí. Pak bychom hledali, proč zrovna ponižování souvisí se zvyšováním sebevědomí, a to bychom museli nahlédnout i do odborné literatury. To nás může vést zpátky do terénu zjistit, jak pokřiky fanoušci připravují, jestli je nacvičují či zda existuje jejich stabilní repertoár apod.

Kontext pokřikování je evidentní, je jím sportovní utkání, ale musíme z dat zjistit, jak fanoušci utkání kategorizují, jako jaký typ události ho chápou. Někteří fanoušci za hlavní bod utkání nepovažují samotný souboj hráčů na hřišti, ale bitku s fanoušky druhého týmu. Sportovní utkání je pro ně zámkou pro kolektivní rvačku. V tomto kontextu je pokřik podobný pokřikování bojovníků před bitvou, které jim má dodat odvalu zesměšněním nepřitele a morální posílení poškozením jeho tváře [srov. Goffman 1967: 5–46]. Ale jsou možné i jiné kontexty.

Na hřišti nelze křičet cokoli. Například když by někdo zakřičel, že „Zítří je neděle!“ nebo „Baník jsou strojuvůdci!“, asi by se z toho pokřik nestal. *Intervenující podmínkou* je například fanouškovský diskurz. Jen některá slova mohou sloužit jako ponižující označení [Bargiela-Chiappini 2003; Brown, Levinson 1987]. Dalšími intervenujícími podmínkami jsou etnické a rasové stereotypy, které pokřik používá. Kdyby takto pokřikovali v Německu nebo Holandsku, málokdo by tomu rozuměl, protože si nespojí Ostravu s blízkostí Polska a Polsko s naším sousedem. I to, že pokřik nelze vymyslet libovolně, že aktéři mají k jeho tvorbě určité vymezené podmínky, je třeba při interpretaci vzít v úvahu.

Důležitým prvkem v modelu je *strategické jednání a interakce*. Teoreticky zakotvená teorie vychází z pragmatismu a ze symbolického interakcionismu, které tvrdí, že významy se vytvářejí až v interakcích jedinců, kteří do nich vstupují s určitými záměry [Corbin, Strauss 1990]. Je proto nanejvýš důležité při interpretaci dat vzít v úvahu interakční kontext, v němž se jev vyskytuje, včetně příčinných a intervenujících podmínek a kontextu. V našem případě musíme zjišťovat, za jakým účelem jedni fanoušci pokřikují na druhé rasistické a xenofobní urážky. Jak na to ti druzí reagují? Není to hra na přebíjeňnou, kdo toho druhého více poniží a tím i našve? Jsou pokřiky svého druhu konverzací? Na druhé straně některé prvky pokřiku mohou být nezáměrné. Primárním cílem možná není urážet Poláky a asi ani Židy. Tato označení jsou použita pro jejich nízký symbolický status vůči „naší lize“ a v rámci fanouškovské subkultury.

Konečně posledním prvkem modelu, který bychom při axiálním kódování měli brát v úvahu, jsou *následky*, jaké jev má. Následky opět musíme

hledat v sebraných datech a nikoli o nich spekulovat, jak to zde činím já. Následky jevu můžeme rozdělit na zamýšlené a nezamýšlené ze strany účastníků jednání. Zamýšleným následkem pokřiku může být posílení identity a sebevědomí fanoušků, kteří křičí. Nezamýšlené následky může pokřikování mít na ty fanoušky v obecnstvu, kteří s obsahem pokřiku nesouhlasí a mohou intervenovat u managementu klubu, aby takovým pokřikům zabránil.

To je ve zkratce paradigmatický model. Práce s modelem znamená systematické vytváření otázek a hledání odpovědí v datech. Například hledáme nezamýšlené důsledky jednání jednoho aktéra, které mohou vytvářet kontext pro jednání jiného aktéra. Při tomto úsilí nepracujeme jen na úrovni kategorií, ale sestupujeme i na úroveň kódů a primárních dat. Zjišťujeme, jestli jsme nezapomněli kódovat něco, co se ukázalo pro pochopení pokřiku jako formy bojovné konverzace důležité. K tomu nám dobře poslouží software pro analýzu kvalitativních dat, jako například Atlas.ti [Konopásek 1997, 2005, 2008, 2011], NVIVO [Bringer, Johnston, Brackenridge 2006] nebo Qualrus. V množství papírů může analytik ztratit hlavu, ale pokud má vše v počítači, je orientace ve velkém množství dat jednodušší.

Ještě než přistoupíme k selektivnímu kódování, chtěl bych zdůraznit, že ke kódování nelze přistupovat mechanicky. Paradigmatický model slouží jako pomůcka pro systematické porovnávání a uspořádávání kategorií, ale není povinným stadiem výzkumu nebo nějakou garancí, že něco objevíme. Když vzpomenu příklad s mobilním telefonem, podobně jako na domácí rozebírání telefonů není žádný přesný návod, na kódování také ne. Ale pokud dodržujeme určitá pravidla (např. že do telefonu nemlátíme kladivem), časem se to naučíme, podobně jako kódování. Paradigmatický model je něco jako obecný model telefonu, tzn. říká nám, že když jsme ho rozebrali, měli bychom mezi díly najít nějakou anténku, nějakou baterii, SIM kartu, nějaké čudlíky, kterými se spouští apod. Podobně každý sociální jev má nějaké příčiny a následky, kontexty a interakce. Axiální kódování je uspořádávání kategorií do vztahu jak vůči sobě navzájem, tak i k celku.

### 3.3.1.1c *Selektivní (výběrové) kódování*

Třetím typem kódování je selektivní kódování. Je to práce s kategoriemi, jejímž cílem je vybudovat zakotvenou teorii nějakého jevu. Je to závěrečná integrace kategorií do jednoho celku. Opět, jako v případě předcházejících typů kódování, nejde o závazný postup nebo o kuchařku; k dispozici jsou určitá pravidla, pokud se jimi řídíme, můžeme kódování úspěšně provádět.

Prvním krokem je vytvoření příběhu, resp. nalezení pointy, která určitou kategorii zvýznamní a učiní z ní kategorii centrální. Pointu příběhu nejsnadněji nalezneme tak, že vyprávíme někomu jinému, na co jsme přišli.

Abychom udrželi posluchačovu pozornost, přirozeně nás to svádí zvýraznit určitý aspekt jevu. Není ovšem výjimkou, že při snaze vyprávět zjištění někomu jinému zjistíme, že jsme vlastně nic zvláštního neobjevili, že žádnou pointu nejsme schopni vymyslet. Nebo se nakonec posluchač zeptá: „No a? A to má být objev?“ V takovém případě je třeba se znovu vrátit ke kategoriím, případně kódům, a hledat mezi nimi dosud nezjištěné vztahy, které by nás k pointě navedly. Je ale potřeba říci, že jev sám žádnou pointu nemá, tu tam vkládáme až my jako analytici a interpreti. Někdy se ale stane, že potenciálních point je více; v tom případě je nutné si zvolit jen jednu. Proto se tento typ kódování nazývá výběrovým.

Vezměme si náš fanouškovský příklad. Zkoumáme pokřiky fanoušků a zjistíme o nich mnoho informací typu kdo, kdy, kde, jak a proč. Stojíme teď před otázkou, co se stane našim objevem. Můžeme zkusit tvrdit, že centrální kategorií je interakční povaha skandování, jakési paralelní utkání k utkání hráčů. Že existují určitá nepsaná pravidla skandování. Čím agresivnější a tvrdší je utkání sportovní, tím agresivnější je i skandování fanoušků. Můžeme zkusit vybrat centrální kategorii, která by souvisela s nějakou již existující teorií, například teorií konfliktu. Pak bychom mohli tvrdit, že jde o formu nerealistického konfliktu, který se jako konflikt tváří, ale ve skutečnosti jde o posílení skupinové solidarity. Ať už bychom vybrali první, nebo druhou interpretaci, museli bychom podle ní uspořádat ostatní kategorie a podřídit je logice této interpretace. Někdy se stane, že centrální kategorie je celkově dobře podpořená, ale někde existují slabá místa. Tehdy musíme zpátky ke kódům a ke kategoriím nebo dokonce přímo k datům a pokusit se buď smysl centrální kategorie upravit, nebo najít v datech potřebnou podporu.

V této fázi výzkumu je důležité, abychom měli načteno dostatek literatury, která nám poskytne zdroje pro hledání centrální kategorie. Může to být buď přímo, jako v případě nerealistického konfliktu, který byl původně konceptualizován Georgem Simmelem a později ho rozpracoval Lewis Coser [Coser 1956], nebo nepřímo, jako v případě interakční povahy skandování.

V praxi se často setkáváme s otázkou, kdy můžeme přistoupit k selektivnímu kódování. Je evidentní, že i jeden rozhovor s fanouškem můžeme kódovat, vytvořit kategorie, paradigmatický model a začít s budováním teorie. Nicméně byla by to jen slabě zakotvená teorie, protože by byla s empirickou realitou propojená jen jedním rozhovorem. Opačný problém může nastat, když máme nasbíráno velké množství dat a vytvořenou řadu kategorií, ale chybí nám odvaha začít stavět teorii. Pro rozhodování o tom, kdy lze začít s teorií, existuje pomůcka, která se nazývá teoretická saturace. Teoretická saturace nastává v okamžiku, kdy už další nová data nepřinášejí významné

nové poznatky ani neobohacují existující kategorie, jejich dimenze a vzájemné vztahy [Glaser, Strauss 2006: 62]. V tomto okamžiku je vhodné přejít k hledání centrální kategorie, pokud jsme na ni už dříve nepřišli.

Samotné vytvoření zakotvené teorie poté spočívá v argumentovaných výrocích o vztazích mezi kategoriemi, které jsou v souladu s dominantní rolí centrální kategorie. Teorii pak můžeme formulovat narativně, ve formě příběhu, nebo systematicky, podle nějakého modelu (např. rozpracovat a přizpůsobit paradigmatický model).

### 3.3.1.2 *Shrnutí kódování v zakotvené teorii*

Kódování postupem zakotvené teorie je variantou obsahové analýzy prováděné od začátku do konce lidským čtenářem. Je charakterizováno snahou o maximální indukčnost, tzn. kódy a posléze kategorie by měly být v co nejvyšší míře odvozeny z analyzovaných textů. To neznamená, že by byl analytik osvobozen od všech forem předporozumění, hermeneutický kruh [Gadamer 2004: 265 a násl.] zde platí. Nejprve operuje s referenčním kódováním, které umožní vytvoření kategorií a následně i paradigmatického modelu. Kategorie, s nimiž pracuje axiální a selektivní kódování, jsou tak i nadále ukotvené v kódech s referenční funkcí. Z tohoto důvodu je v tomto typu obsahové analýzy často používáno „in vivo“ kódování, kdy k označení kódu nebo kategorie použijeme přímou citaci analyzovaného textu.

Silnou stránkou tohoto typu obsahové analýzy je právě induktivní postup, který slibuje o data opřenou konstrukci teoreticky nepředpokládaných vztahů; odhaluje obsah, jenž nejvíce odpovídá tomu, který generují účastníci sociálního dění zacházející s texty (vytvářejí je, distribuují i je čtou). Jinými slovy, postupem zakotvené teorie můžeme dosáhnout značného průniku relevantní aktérů i analytiků ve vztahu k obsahu analyzovaných textů.

Nevýhodou tohoto postupu je obtížnost. Nezkušený analytik velmi často dojde buď k banálním zjištěním, kdy pouze tematizuje obsahy běžně tematizované laickými čtenáři textů, nebo naopak k sofistickým, nicméně spekulativním výsledkům, které jsou významově adekvátní, ale těžko dokazatelné. Udržet se na interpretační úrovni blízké aktérům a současně dojít k odborně relevantním zjištěním je nesnadné a vyžaduje zkušenost a hermeneutické nadání.

### 3.3.2 *Sémioticky orientovaná obsahová analýza*

V úvodu této kapitoly jsem zdůraznil, že čtení je sémiotická aktivita. Samozřejmě to platí i pro vytváření textů. V rámci obsahové analýzy se proto

můžeme zaměřit na studium textů jako *výsledků* sémiotické aktivity jejich autora nebo na obsahy textů, které jsou generovány čtenáři. Jinou možností je zaměřit se na analýzu tvorby textů nebo čtení textů jako na aktivity generující významy a obsahy textů. Vezměme známý příklad – gymnaziální otázku – co chtěl básník říci verši:

*My přijdem blíž, my přijdem blíž,  
my světů dožijeme,  
my bijem o mříž, ducha lvi,  
a my ji rozbijeme!*

Obvyklá „správná odpověď“ je ta, že Neruda tím chtěl vyjádřit jednak vůli překonat malost poměrů a jednak víru v lidský pokrok; tato odpověď je vlastně výsledek obsahové analýzy básně. Analýzou skutečného studentského čtení dané básně bychom pravděpodobně zjistili, že jednotliví čtenáři vytvářejí obsah básně různě, nicméně vlastním čtením vytvořený obsah, resp. jeho fragmentárnost či absenci, nepovažují (z ryze pragmatických důvodů) za legitimní obsah, který by mohl být prezentován jako „správná odpověď“ na výše uvedenou otázku. Školní lavice nejsou jediné místo, kde se vyžaduje, aby vytvářený obsah textu odpovídal především očekávání ostatních; v mnoha různých situacích vytváříme čtením takové obsahy, které jsou v dané situaci legitimní. A právě zkoumáním toho, jak se během čtení textů vytvářejí obsahy, se zabývá sémioticky orientovaná analýza [pro klasické texty viz Eco 2010; Hall 2005/1973]. V následující části představím jednu konkrétní aplikaci takové analýzy, při níž se zjišťovalo, jak příslušníci tří generací kooperují s úryvky narativních textů a vytvářejí jejich obsahy; jedná se o výzkum, který jsem prováděl s Marií Dlouhou a dalšími studenty [Hájek, Dlouhá 2011, 2014].

#### **Výzkum čtenářské kooperace s úryvky biografických vyprávění**

Ve výzkumu jsme se zaměřili na dosud jen málo prozkoumaný způsob, jakým jsou biografická vyprávění pamětníků, která můžeme chápat jako určité dokumenty kolektivní paměti, přijímány členy komunity paměti. Přijetí nějakého dokumentu jako dokumentu paměti, tzn. jako (legitimní) součásti kolektivní paměti, je komplexní proces a probíhá v několika úrovních, rozlišitelných pouze analyticky. První úroveň je situace, v níž se člověk s dokumentem setkává, například muzeum, dokumentární pořad v televizi, hraný film v kině, kniha, billboard apod. Tuto úroveň charakterizuje očekávání, že dotýčný se „naladí“ na určitý žánr či sociální kontext, který situace implikuje. Příjemce se může rozhodnout, že daný žánr nepřijme, že jej bude popírat

nebo k němu přistoupí subverzivně (např. ironicky). Příkladem může být neochota nebo právě ochota jít do Muzea totality nebo začít číst knihu *Paměti milicionáře*. Druhou úrovní je lokální kontext, ve kterém je dokument prezentován a který orientuje potenciálního příjemce směrem do určité minulosti, například předmluva k vyprávěním pamětníků, komentář k muzejnímu exponátu, datum na fotografii apod. Příjemce dokumentu paměti zde musí opět aktivně participovat v tom smyslu, že přijímá (nebo nepřijímá) určitá pravidla hry, například v historickém románu kombinaci historických a fiktivních momentů, v komentáři k muzejnímu exponátu důvěru k faktičnosti sdělovaných informací nebo u biografického vyprávění předpoklad autenticity sdělení (na rozdíl od fiktivních autobiografií) [Hall 2005/1973]. Třetí úroveň předpokládané kooperace při recepci dokumentu jako dokumentu paměti je jeho obsah, například text básně, obsah filmu, podoba budovy, znění písně. Aby si dokument udržel status dokumentu paměti, je zapotřebí, aby recipient pochopil jeho obsah ve vztahu k sdílené minulosti určitého společenství paměti. Například ve filmu uvedeném jako historický film musí porozumět, v čem je film historický, tzn. události v něm zpodobňované musí úspěšně přiřadit k vlastním představám o dané době. Případný nesoulad může vést k odmítnutí dokumentu jako dokumentu paměti (např. ve prospech „propagandistického dokumentu“).

Při aktuální interakci s dokumentem působí všechny tři úrovně současně, tzn. v určité situaci je specifickým způsobem prezentován objekt zvláštní povahy, například v knihovně je kniha biografických rozhovorů s pamětníky socialismu obsahující i průvodní studii. Ne nutně však všechny úrovně působí stejným směrem: táž kniha tvořící součást uměleckého objektu vytvořeného z množství knih různého stáří, který se nachází v předsálí knihovny, nebude značnou částí návštěvníků knihovny považována za „dokument paměti“, tzn. za legitimní součásti jejich „kolektivní paměti“, i když v jiném kontextu (na polici označené „Životopisy“) by nepochybně za dokument paměti považována být mohla a byla.

Z uvedených tří analyticky pojímaných úrovní procesu recepcce dokumentů paměti jsme se v našem výzkumu soustředili na úroveň třetí, na to, jak je posuzován obsah určitého dokumentu jako potenciálního dokumentu paměti. Jinými slovy, chtěli jsme zjistit, jak členové komunity paměti interagují s předkládaným textovým dokumentem tak, aby ho mohli začlenit do intersubjektivně sdílené reprezentace minulosti a dát mu tak status dokumentu paměti. Domnívám se totiž, že kolektivní paměť nesestává pouze z intersubjektivně sdíleného souboru reprezentací minulosti, ale je současně specifickou schopností jednat, kompetencí s dokumenty paměti zacházet. To, co činí z člověka člena určité komunity paměti, není výlučně jeho znalost sdí-

lené minulosti, ale – a možná přednostně – umění náležitě konfrontovat sebe a ve společnosti existující dokumenty se vztahem k minulosti. Prakticky vzato, člověk je nucen aktivovat „kolektivní paměť“ vždy, když se setká s nějakým dokumentem, pro který je nárokován statut dokumentu paměti.

### 3.3.2.1 *Interpretační kooperace ve vyprávění*

Vzhledem k narativnímu charakteru textů použitých ve výzkumu (jeho účastníkům jsme předložili úryvky z přepisů autobiografických vyprávění pamětníků socialismu) jsme při analýze metodologicky vycházeli ze dvou zdrojů: z Ricoeurovy teorie vyprávění, zejména z jeho konceptu trojí mimésis rozpracovaného v prvním díle *Času a vyprávění* [Ricoeur 2000: 88–134; srov. Hamar 2008: 28–40], a z Ecovy analýzy čtenářské kooperace v narativních textech [Eco 2010]. Ricoeur charakterizuje každé vyprávění jako proces trojí mimésis [viz kap. 5.4.2]. Tato konceptualizace vyprávění je ze sociologického hlediska přínosná, protože do vyprávění integruje jak interpretaci světa vypravěčem, která vyprávění předchází, tak následnou interpretaci vyprávění čtenářem. V našem výzkumu jsme respondentům nabídli úryvek autobiografického vyprávění jako dokument paměti a ten byl jimi jako takový přijat, pokud v něm našli sdílený svět (socialistické) minulosti, na který vyprávění odkazovalo, a současně ho dokázali vnímat jako aktérskou biografickou výpověď, které je možné porozumět s ohledem na nynější situaci.

Při analýze čtenářské aktivity respondentů, tj. při vytváření obsahu textu, jsme použili sémiotický přístup, jak ho rozpracoval U. Eco v knize zabývající se čtenářskou kooperací v narativních textech [Eco 2010], tzn. tím, co by Ricoeur označil jako třetí mimésis. Eco ve své práci vychází z Peirceova pragmatismu, v němž „odhalit objekt [...] znamená odhalit *modus operandi*, čili vyprodukovat jej (nebo vyprodukovat jeho praktické použití)“ [Eco 2010: 59; kurziva původní]. Na tomto základě pak staví svou teorii interpretační kooperace v narativních textech, jejichž porozumění čtenářem je úkol veskrze praktický, spočívající v aktualizaci jejich obsahu. Proto jsme se v našem výzkumu zaměřili na způsoby, jakými čtenáři s úryvky biografických narativů kooperují při jejich interpretaci. Poněvadž jednou z klíčových čtenářských strategií porozumění vyprávění je konstrukce hypotézy narátora [Bortolussi, Dixon 2003: 60–96; Eco 2010: 79; Mullins, Dixon 2007], zajímalo nás, jak budou recipienti biografických příběhů konstruovat identitu vypravěčů a jak bude identita domnělých vypravěčů spojena s interpretací textu. Druhou věcí, kterou jsme sledovali a pro kterou Ecova sémiotika poskytuje analytické nástroje, byla skutečnost, že čtenářská interpretace textu (aktualizace jeho obsahu) se nezakládá jen na skutečnostech v textu fakticky přítomných (nezávisle na intencích autora), ale současně vypovídá i o samot-

ném světě čtenáře [Eco 2010: 82; Ricoeur 2000: 113]. Touto cestou jsme byli schopni postihnout způsoby, jakými se realizují specifické kompetence příjemce dokumentů paměti vedoucí k jejich přijetí a performované integraci do kolektivní paměti.

### 3.3.2.2 *Experiment*

Provést výzkum formou experimentu jsme se rozhodli vzhledem k položeným výzkumným otázkám i ke způsobu analýzy. Jelikož jsme se snažili prozkoumat v podstatě obyčejné čtení, tj. samozřejmě vnímanou a rutinní interakci – předpokládanou kooperaci – s dokumenty paměti, nabízelo se navázat na etnometodologickou tradici zviditelnění taken-for-granted řádu prostřednictvím jeho narušení [Bonková 2002; Garfinkel 1967: 35; McHoul 1982]. Logika experimentu se odvíjela od předpokladu, že člen komunity paměti (v tomto případě české společnosti) při setkání s domnělým dokumentem paměti dokument jako takový buď odmítne, nebo přijme. Pokud ho přijme (což jsme vzhledem k situaci předpokládali), bude se ho snažit zapracovat do kontextu legitimního obrazu minulosti. Abychom tento proces zviditelnili, předložili jsme účastníkům experimentu zvláštní dokumenty – krátké fragmenty autobiografických vyprávění pamětníků socialismu – které od čtenářů vyžadovaly značnou kooperaci (dle našeho soudu až na hranici únosnosti), aby mohly být integrovány do všeobecně přijímaného (nebo přijatelného) historického kontextu.

Konkrétním úkolem respondentů-čtenářů pak bylo na základě četby přiřadit vypravěčům politickou identitu (podrobnosti uvádím níže). Následně byly s respondenty provedeny rozhovory o důvodech jejich volby. Provedený experiment sice nebyl narušující v garfinkelovském pojetí, protože radikálním způsobem neporušoval tichá očekávání (background expectation) účastníků a nevyvolával v nich nejistotu, zmatek či dokonce zlobu vůči experimentátorům [Garfinkel 1967: 37–49]. Respondenti věděli, že se účastní experimentu a byli připraveni na nekaždodenní, nerutinní činnosti. Narušení běžného řádu spočívalo v převrácení okolností při setkání s dokumenty paměti; v přirozených situacích je znám autor či zdroj dokumentu, zatímco je nejasná autenticita dokumentu, která je následně vůči autorovi či zdroji poměřována (zda mluví/píše pravdu, „jak to bylo“, či nikoli). V našem experimentu jsme respondenty naopak postavili před deklarovaně autentické dokumenty s nejasným autorstvím (disident, komunistický funkcionář nebo obyčejný člověk?). Druhé převrácení přirozeného pořádku spočívalo v předložení neobvykle krátkých úryvků; respondenti, zvyklí rozumět výpovědím na základě kontextů, ve kterých se vyskytují, tzn. opírat se o indexikální charakter přirozeně se vyskytujících výrazů [Garfinkel 1967: 4–7], tak byli nu-

cení naopak potřebný kontext dotvářet, aby se úryvky staly srozumitelnými vzhledem k domnělé identitě vypravěče.

Celkem se experimentu zúčastnilo čtrnáct osob. Účastníci pocházeli ze čtyř rodin a v jejich rámci reprezentovali tři generace – prarodiče, rodiče a děti. Nejstarší generace prožila většinu svého života v období socialismu, střední generace si socialismus pamatuje ze svého mládí, případně středního věku, a nejmladší generace, dnešní dvacetiletí, znají socialismus jen zprostředkovaně. Nejmladší generaci tvořili studenti (jeden muž a tři ženy) magisterského stupně studia sociologie na Fakultě sociálních věd Karlovy univerzity, střední generaci jejich rodiče (jeden muž a čtyři ženy) a nejstarší jejich prarodiče (dva muži a tři ženy; v jednom případě byl příbuzenský vztah nepřímý). Z hlediska sociální homogenity lze považovat náš vzorek nejvíce homogenní v generaci studentů a nejméně v generaci prarodičů. Stojí za zmínku, že žádný ze starších respondentů nebyl v době socialismu významně politicky angažován, ať už pro nebo proti režimu. Relativně malý počet respondentů (14) nepovažujeme za slabinu výzkumného designu vzhledem k tomu, že jsme pro analýzu získali téměř 300 konverzací o jednotlivých úryvcích, tj. 100 pro každou generaci, a v rámci generace 30 konverzací pro každou ze tří identitních kategorií vypravěčů. Celkem jsme měli k dispozici 120 normostran přepisů. Při analýze rozhovorů jsme vzali v potaz i to, že nejmladší generaci zastupovali studenti sociologie; z analýzy nám však nevyplynulo, že by jejich odborné znalosti a zkušenosti hrály při kooperaci s úryvky signifikantní roli.

Technicky proběhl experiment následovně. Každému účastníkovi jsme předložili list s 21 krátkými úryvky vyprávění vybraných ze sbírek přepisů biografických rozhovorů vytvořených v rámci dvou projektů řešených v Centru orální historie (COH) Ústavu pro soudobé dějiny (ÚSD) AV ČR. Konkrétně šlo o sbírku rozhovorů *Politické elity a disent v období tzv. normalizace. Životopisná interviews.*, ve které byli narátoři buď bývalí vysocí komunističtí funkcionáři, nebo příslušníci disentu, a sbírku *Průzkum české společnosti v období tzv. normalizace. Biografická vyprávění příslušníků dělnických profesí a inteligence.*, ve které vyprávějí, jak už z názvu vyplývá, příslušníci dělnických profesí a inteligence. Charakteristiky studovaných sbírek rozhovorů jsou dostupné na webových stránkách COH ÚSD AV ČR<sup>6</sup>; značný počet rozhovorů byl publikován v editované podobě knižně [Vaněk, Urbášek 2005; Vaněk 2009].

Vybrané úryvky sestávaly z 35 až 45 slov dlouhých úseků textu obsahujících zájmeno *nás*. Důvodem záměrné přítomnosti zájmena *nás* byla potřeba asociovat předkládané texty s nějakým aspektem kolektivní identity

6 <http://www.coh.usd.cas.cz/cs/katalog-a-sbirky-rozhovoru>

vypravěče – komunistického funkcionáře, disidenta nebo tzv. obyčejného člověka. Naším cílem tudíž bylo vybrat takové fragmenty vyprávění, v nichž by byla patrná identitní práce vypravěčů [Becher 2008; Snow, Anderson 1987]. Předpokládali jsme, že v takových úryvcích bude narace obsahovat určitou diskurzní konstrukci skupiny, k níž se vypravěč řadí a která umožní respondentům snáze odhadnout jeho politickou identitu [Galasinski 2010]. Úryvky byly náhodně vybrány z korpusu rozhovorů prostřednictvím sady náhodných čísel, vždy osm z každé skupiny vypravěčů. Z celkově vybraných 24 úryvků jsme dva vyřadili, protože se vztahovaly ke druhé světové válce a nikoli k období socialismu, a v jednom případě jsme úryvek vyřadili z toho důvodu, že text byl celkově obtížně srozumitelný.

Jak jsem již uvedl, účastníkům experimentu byl předložen list s 21 úryvky. Respondenty jsme vyzvali, aby označili, do jaké ze tří výše uvedených politicky definovaných skupin osob by zařadili narátora, který vyslovil daný úryvek v rámci svého životního vyprávění. Respondentům přitom nebylo známo rovnoměrné zastoupení narátorů pro každou skupinu (nechtěli jsme, aby jejich určování ovlivňovala určitá kalkulace a chtěli jsme omezit vzájemné srovnávání jednotlivých úryvků) ani nebyli nuceni se k některé z identit přiklonit. Naopak, umožnili jsme jim v případě pocíťované nemožnosti určení neoznačovat identitu žádnou, resp. označit identit několik. Kromě samotného označování domnělých vypravěčů jsme respondenty současně požádali o podtržení slov či oblastí textu, které je vedly ke zvolení narátorovy identity. Poté, co respondenti určili identity vypravěčů a podtrhli informativní slova a slovní spojení, jsme s nimi provedli rozhovor, ve kterém jsme je požádali o vysvětlení důvodů svých voleb, tzn. proč zvolili danou identitu a proč podtrhli ta která slova. Rozhovory vedl vždy student se svým příbuzným (rodičem, prarodičem), resp. kolegou/gyní bez přítomnosti cizích osob; ve dvou případech jsem vedl rozhovor se studenty já, příslušník střední generace. Rozhovory jsme nahrávali a jejich přepisy následně podrobili analýze.

### 3.3.2.3 *Metoda analýzy*

Technicky se analýza opírala o konceptuální nástroje vytvořené U. Ecem [Eco 2010]. Eco stratifikuje textovou kooperaci čtenáře s narativem do několika úrovní [Eco 2010: 85–105]. První kooperace se odehrává na úrovni slovníku, tzn. řetězce slovních výrazů neboli lexematického povrchu textu [Eco 2010: 91]. Čtenář kooperuje s textem tím, že na slova tvořící text aplikuje systém pravidel (kódů), který mu umožňuje text převést do první úrovně obsahu. Čte-li v textu „tak jsme se kapku pohádali“, je schopen rozpoznat, že slovo *kapka* nese v textu význam „trochu“ a *pohádali* znamená „verbálně

komunikovali vzájemné neshody.“ Na rozdíl od tohoto příkladu slovní výrazy *čínský budoár* nebo *sofistikované systémy* nemusejí být všem čtenářům na této základní úrovni jednoznačně srozumitelné a jejich význam se budou snažit odhadnout z kontextu. V našem výzkumu nás tato úroveň slovních výrazů (slovníku) zajímala především v souvislosti se slovy a slovními spojeními, která čtenáři označili podtržením jako klíčová pro identifikaci vypravěče. Přirozeně, námi sledovaný „podtrhaný“ slovník tvoří jen výseč ze slovníku, který čtenářům umožnil pochopit význam textů. Jednalo se poněvíc o slova, jež měla z pohledu čtenářů vztah k minulosti, resp. k tomu, co z minulosti mělo potenciální význam pro určení identity vypravěče. Pro názornost můžeme uvést několik slovních výrazů, na něž čtenáři upozorňovali: *vězení, nás tajemníky, co studujeme, protěžovaný* atd. Od úrovně slovníku jsme očekávali, že nás přivede k odpovědi na dvě otázky: jaký má slovně výrazová složka vyprávění podíl na přijetí dokumentu čtenářem jako dokumentu paměti a do jaké míry je tato slovně výrazová složka vyprávění sdílená napříč generacemi i jednotlivými respondenty?

Úroveň slovníku umožní čtenáři udělat „první pokus o amalgám“ [Eco 2010: 96], tzn. první mlhavou představu o tom, o co se v přečteném textu jedná. Druhou úroveň kooperace čtenář realizuje na úrovni encyklopedie. Eco hovoří o encyklopedii, nebo přesněji o encyklopedické kompetenci, jako o schopnosti čtenáře rozumět slovům v textu vzhledem ke kulturní tradici, k níž se základní slovník textu vztahuje [Eco 2010: 96]. Například objevili se v textu výraz *byli jsme ve vězení*, čtenářova encyklopedická kompetence mu poskytne nutné vědění k tomu, aby mohl dále pokračovat ve čtení (co je to vězení pro účely čteného textu, např. nemožnost být s rodinou) a současně v konstrukci toho, o čem daný text pojednává. V sociologickém myšlení je konceptu encyklopedie blízký Schützův pojem „balíku příručního vědění“ (stock of knowledge at hand), nicméně Schütz svůj pojem chápe obecněji, jako interpretační schéma pro současné i minulé zkušenosti každodenního života [Schütz 1970: 74]. Eco encyklopedii pojímá úzce jako schopnost propojit jazyk s kulturou, podobně jako knižní encyklopedie propojuje heslo s jeho kulturním významem. V našem výzkumu jsme zkoumali úroveň čtenářské encyklopedie z toho důvodu, že odkazuje k něčemu, co bývá intuitivně spojováno s kolektivní pamětí, tj. s určitou zásobou vědění o minulosti. Zjišťovali jsme, jaké obsahy aktualizují u čtenářů slova základního slovníku označená v první úrovni kooperace s textem.

V rámci prováděné analýzy jsme se ještě pokusili o zkoumání třetí – nejkomplexnější – úrovně čtenářské kooperace, kterou je konstrukce „možných světů“. V sémiotice a literární vědě je pojem možného světa velmi diskutovaný, viz [Doležel 2003; Eco 2010], a není mojí ambicí zde do této dis-

kuze vstupovat. Termín jsme přejali a zacházeli s ním v jednom důležitém ohledu odlišně od běžného užití – vypravěči autobiografií si na rozdíl od autorů fiktivních příběhů nemohou sestavovat svět svého životního příběhu libovolně, protože by se mohli dostat do konfliktu se čtenářem (posluchačem), jenž by mohl namítnout, že „takhle to přece vůbec nebylo!“ Analogicky čtenáři biografií nemohou nekriticky přijímat text a na jeho základě libovolně konstruovat hypotézy struktur možných světů, protože si uvědomují, že referentem je minulost skutečná, ne fikce (jako v případě historických románů nebo retrofilmů). Nicméně i přesto považuji pojem možného světa za vhodný k analýze interpretační kooperace v biografických textech, a to zejména z toho důvodu, že umožňuje zachytit čtenářovu (nikoli vypravěčovu) cestu k významu, který pro něj úryvek vyprávění nabývá. V autobiografickém vyprávění čtenář na základě čteného textu a své encyklopedie činí propozice o tom, jak se bude děj vyvíjet, anticipuje následné stavy, které se mu po přečtení dalšího textu buď potvrdí, nebo ne [Eco 2010: 139]. Naši čtenáři se však museli potýkat s tím, že text jim poskytoval pouze omezený materiál pro vytvoření hypotézy o tom, co se děje, a svým rozsahem už jim zpravidla neumožnil tuto hypotézu ani dílčím způsobem ověřit. Respondenti-čtenáři tak za pomoci své encyklopedie konstruovali situační kontext či kontexty, v nichž by předložený úryvek pro ně nabýval smysl. Tyto kontexty lze chápat sémioticky ve vztahu k textu jako možné světy, tj. situace považované respondenty za reálně možné. Současně musíme připustit, že v některých případech si byli respondenti natolik jisti kontextem a svou interpretací úryvku, že je vztahovali jen k určitému subjektivně reálnému historickému světu, který znali z vlastní nebo zprostředkované zkušenosti. To v naší analýze znamenalo limitní případ možného světa. A právě interpretační pohyb čtenáře mezi „možným“ a „skutečným“ světem byla hlavní kooperativní aktivita, kterou jsme sledovali.

#### 3.3.2.4 Kooperace na úrovni slovníku

V Ecově pojetí čtenářská kooperace na úrovni slovníku představuje především čtenářské, zpravidla rutinní a neuvědomované rozpoznání elementárních sémantických vlastností výrazů, například že výraz *žena* znamená lidskou bytost ženského pohlaví [Eco 2010: 96]. V našem případě se však většinou nejednalo o takto bazální úroveň kooperace, ale o sledování, která slova či slovní spojení čtenáři-respondenti považovali za významná pro určení identity mluvčího jako komunistického funkcionáře, disidenta nebo obyčejného člověka. Konkrétně jsme pozorovali, jak se mezi generacemi respondentů liší zvýrazněná slova a slovní spojení a zejména zdůvodnění jejich role a významu při konstrukci hypotézy celkového významu úryvku za úče-

lem nalezení identity vypravěče. Vycházeli jsme nejen ze slov podtržených při čtení, ale především z následných rozhovorů-komentářů, v nichž respondenti význam podtržených slov rozváděli.

Z hlediska generační struktury respondentů musím překvapivě konstatovat, že rozdíly mezi generacemi nejsou výrazné. Jediná znatelná odlišnost se týká nejmladší generace, která se častěji zaměřovala na hledání historicky relevantních významů jednotlivých slov než na skupinově specifické výrazivo nebo potenciální vysvětlující kontexty (viz dále). Přestože tedy čtenáři prožili v období socialismu různou dobu a jejich osobní zkušenosti jsou různé, na úrovni jednotlivých slov a slovních spojení není jejich kooperace s úryvky významně odlišná.

Popíšme si proto nalezené sdílené charakteristiky čtenářské spolupráce s úryvky narativů. I když hlavním úkolem respondentů bylo rozpoznání identity vypravěče, nejednalo se rozhodně o jedinou činnost v rámci čtenářské kooperace, kterou účastníci prováděli. Identifikaci provázela nebo předcházela snaha zakomponovat a dát určitý význam i událostem a postavám, jež byly v úryvku explicitně nebo implicitně přítomny, tedy snaha navrhnout hypotézu původního širšího kontextu výpovědi, ze které byl úryvek za účelem experimentu vytržen. Hledání souvislostí mezi obsahem textu a představovaným kontextem nebylo náhodné, protože všechny prvky byly vytvářeny či zpětně upravovány tak, aby byly konzistentní s hypotetickou identitou vypravěče, kterého ať už provizorně, nebo definitivně označili za autora úryvku. Čtenáři považovali za nutné, aby úryvek byl v souladu s navrhovaným širším kontextem biografického narativu postaveným na základě jejich interpretačních voleb. Obraz tak implikoval jistou celistvost, avšak ne vždy byli respondenti schopni přesně označit, která slova je vedla ke konkrétnímu interpretačnímu rozhodnutí o tom, jakou ze tří kategorií vypravěč reprezentuje. Například jeden respondent označil vypravěče za obvyčejného člověka a současně podtrhl slovo *protěžovaný*<sup>7</sup>:

**T[AZATEL]:** Proč protěžovaný vede k obvyčejnému člověku? Že byly ty kuchařky protěžovaný...

**R[ESPONDENT]:** Nó, nó, nó. To spíš odkazuje na to, že by to byla komunistka.

**T:** Hmm.

**R:** Ale prostě já jsem to podtrhnul... Automaticky asi.

(úryvek 1, student B; slova převzatá z úryvku jsou zde a v dalších přepsaná základním písmem)

<sup>7</sup> V prepisech zůstává zachován charakter mluvené řeči, tzn. hovorové formy slov; zde např. *protěžovaný* místo spisovného *protěžovaný*.

většinou ale respondenti bez obtíží dokládali vazbu mezi podtrženými slovy a identitou vypravěče, jak ukazuje následující příklad:

**R:** No, tady to druhý... škola jim to nedovolila, poslali nás ostříhat, vždycky přišel ředitel...

**T:** Takže tys podtrhla vždycky přišel ředitel?

**R:** Jo. A říkal ty, ty, ty běžte se ostříhat... no, protože to si přesně pamatuju. [smích]

**T:** Jo.

**R:** Takže to byli ale taky obvyčejný lidi.

(úryvek 2, rodič A)

Stojí však za povšimnutí, že vazba mezi jednotlivými slovy a smyslem výroku zde není bezprostřední. Výraz „poslali nás ostříhat“ odkazuje k určité situaci a teprve ta situace odkazuje k identitě mluvčího. O primární roli slovníku můžeme mluvit jen tehdy, kdy k identitě mluvčího, a tím i k pochopení textu, vedou už sama jednotlivá slova jako v následujících dvou případech:

**R:** No, tohle si myslím, že byli – vůbec nechápu smysl.

**T:** [smích]

**R:** [...] vůbec jo, ale jako když mluví o zámečnicích, elektrikářích, posunovačích, takže to byl asi... asi normální, obvyčejnej člověk.

(úryvek 7, student A)

**R:** No tady jsem dal taky komunistu, protože slovo barák je pro ně taky typický, tohle [KV – pozn. aut.] bude určitě krajskej výbor.

(úryvek 9, rodič Cb)

Ukazuje se, že i když už i na úrovni slovníku jsou čtenáři schopni dosahovat interpretačně úspěšné semiózy orientované na identitu vypravěče a někde dokonce výhradně jen na této úrovni (viz úryvek 7, student A), zpravidla se hledání identity mluvčího a významu výroku opíralo o větší celky než jednotlivá slova nebo slovní spojení.

Druhé zjištění se týká toho, jak se čtenáři vyrovnávali se samým úkolem, který v rámci experimentu řešili, tj. jakým způsobem určit identitu jednotlivých narátorů. Respondenti v tomto případě postupovali různými způsoby, které se objevovaly nezávisle na tom, o jakého šlo respondenta a o jaký úryvek se jednalo. Variabilita způsobů textových heuristik nás vedla ke zjištění, že nelze říci, že by existoval sdílený či předepsaný způsob, kterým by si čtenář utvářel představu o tom, do které kategorie si vyprá-



vějícího zařadit; jde zde o určitou individuální kompetenci. Jedním ze způsobů, kterými čtenáři konstruovali podobu vypravěče, bylo postupné vytváření významů na základě určitých slov či jednoduchých slovních spojení. Například v jednom z úryvků vedlo k identifikaci vypravěče jako komunistického funkcionáře slovní spojení *nás jako tajemníky*. V jiném úryvku výrazy *99 % společnosti a nás volilo* směřovaly čtenáře k tomu, že interpretovali situaci jako běžný výsledek komunistických voleb, a proto mluvčího identifikovali jako komunistického funkcionáře, jak dokládá následující úryvek z rozhovoru:

R: *Tak 99 % společnosti, k tomu mě vedlo, že vlastně 99 % bylo pro komunistickou stranu, to znamená, že to spojuju. Které nás volilo, to znamená 99 % obyvatelstva volilo teda komunistickou stranu a tím, že je tam to slovo nás, tak je jakoby toho komunistu, toho mluvčího, že nás podporuje.*

(úryvek 8, student D)

Jiným způsobem hledání významu výpovědi orientované na odhalení mluvčího bylo něco, co bychom mohli označit jako skupinový styl mluvy, který byl podle respondentů charakteristický pro určité skupiny obyvatel v minulosti. Identifikaci komunistického funkcionáře či disidenta podle zvláštních výrazů, resp. stylu, dokládají následující úryvky z rozhovorů:

T: *No, tady, velice sofistikovaný systém, je to jak bolševická mluva, no.*

R: *Hm.*

T: *Gorbačov zřejmě leccemus nepřeje, z toho taky čiší bolševik, že jo.*

R: *Hm.*

T: *No, tak proto jsem dal komunistu.*

(úryvek 6, rodič Cb)

R: *No komunistu určitě ne. Protože ten by prostě nebyl v té době zavřenej. No... vlastně jenom proto, že tam není ten důvod... naznačený. Ale říkám, možná teda, když teď nad tím uvažuju, že to spíš byl ten disident. Z důvodu toho, že je tam ta kultivovaná mluva... v té větě. A že obvyčejnej vězeň vlastně... mladej... by byl patrně... nějaký dělník. Takže by mluvil jiným způsobem.*

(úryvek 10, rodič B)

V mnoha případech úroveň slovníku (výrazu) nebyla pro určení identity vypravěče podstatná nebo se alespoň neobjevovala ve vysvětleních identitní volby, kterou respondenti prováděli. Vysvětlení se opíralo o komplexnější úrovně semiózy, tzn. o encyklopedii a možné světy. Příkladem může být ná-

sledující vyjádření respondentky, které se vztahuje ke stejnému výroku jako v předchozím případě:

R: *Tady bych řekla, že je to nějaký obyčejný člověk, který, který se neorientoval v té době, který prostě nevěděl, o jaké rozhodnutí jde, šlo zřejmě o nějaké rozhodnutí, ve kterém vězeňský systém byl mírně vylepšen, oslaben nebo ehm. Snažili se hlavně mladé vězně už proti sobě nepopuzovat.*

(úryvek 10, prarodič Da)

Na úrovni slovníku bych ještě rád poukázal na dvojí kooperační strategii, která vedla čtenáře-respondenty ke konstrukci identity vypravěče. Jednou strategií bylo vytváření celkového smyslu úryvku z vyskytujících se slov a slovních spojení, k jeho zařazení do širšího historického kontextu a vytvoření si představy o povaze situace, o níž text vypovídá. Jednalo se zejména o slova, jež se vztahovala k situaci, v níž se úryvek odehrával, buď místně (v *kuchyni, tábor Mariánská, ÚV KSC*) nebo časově (*16. ledna 1989, v 55., 56. roce*), nebo z hlediska identity vypravěče (*nás tajemníky, kuchařky*). Smysl výroku pak byl odhadován na základě těchto indicií. Druhým způsobem čtenářského uchopování významu fragmentů výpovědi bylo více či méně reflektované spojování slov nebo slovních spojení s určitými identitami vypravěčů nezávisle na místě, době či historické identitě. Například slovní spojení *nějací soudruzi* bylo chápáno jako způsob vyjadřování toho, kdo se se „soudruhy“ neidentifikuje; význam slova (ve smyslu slovníkového hesla) byl až druhotný, primární roli hrála konotace neurčitěho osobního zájmena ve spojení se stranickou příslušností. Dalším příkladem tohoto druhu kooperace může být čtení výrazu *studujeme*, který byl v tomto gramatickém tvaru vykládán jako součást výraziva obyčejných lidí, nebo „rozhodně [to] není nějakéj komunistickéj funkcionář, kterej by používal ustálenou frazeologii, ustálené výrazivo, jo?“ (úryvek 13, prarodič Db). Tyto strategie tedy ukazují na různé metody přiřazování identit na základě výrazových charakteristik mluvčích, které se uplatňují pravděpodobně nejen při recepci dokumentů paměti, ale i při recepci jiných dokumentů či v interakcích tváří v tvář.

### 3.3.2.5 Kooperace na úrovni encyklopedie

Druhá úroveň narativní kooperace s textem je vázána na encyklopedickou kompetenci čtenářů. V našem výzkumu byla encyklopedická kompetence chápána především jako vědomosti, které čtenář aktivuje za účelem porozumění textu v průběhu čtení a které jsou spojeny s minulostí společnosti, již se text týká. Základním předpokladem, jakousi vstupní hypotézou, bylo, že mezi generacemi se bude tato kompetence lišit, že v textu přítomné vý-

razy budou aktualizovat jiné či jinak rozsáhlé soubory vědomostí podle generační příslušnosti čtenářů-respondentů. Předpokladem tedy byla jistá podmíněnost encyklopedické kompetence vlastní zkušeností čtenáře, tedy to, zda v době socialismu žil část svého života či zda jeho podobu zná pouze zprostředkovaně (z mediálních sdělení, rodiny, literatury, známých apod.).

Tento předpoklad se však nepotvrdil, alespoň ne tak, jak jsme očekávali. Encyklopedické vědomosti o minulosti, které se čtením textů aktivovaly a na které respondenti při vysvětlování svých voleb odkazovali, obsahovaly zpravidla jen sdílené stereotypní podoby disidentů, obyčejných lidí, komunistických funkcionářů a života za socialismu. Jak se pokusím v interpretaci ukázat, není to z důvodu, že by všichni respondenti věděli o socialismu „to samé“ nebo „stejně“, ale proto, že pro přijetí dokumentu jako dokumentu paměti a určení identity vypravěče nebylo potřeba aktualizovat rozsáhlou encyklopedii; povětšinou dostačovaly běžné rejstříky znalostí, pravděpodobně často aktualizované jinými dokumenty v každodenním životě respondentů. Na druhou stranu navzdory tomu, že charakteristiky hlavních skupin aktérů byly sdílené stereotypně napříč generacemi, respondenti ve své interpretaci pro ně vytvářeli rozmanité kontexty, což ve svém důsledku vedlo k neshodám při konečném určování identity vypravěčů. V tomto smyslu jsou názorným příkladem interpretace následujícího úryvku:

**ÚRYVEK 21:** „Což mě dneska spíš baví, tak aby proběhlo takhle. To byl takovej souběh událostí, kdy jsme nikdo z nás netoužili po tom, aby se něco semlelo. Já k tomu disentu, když už jsem se k tomu tak nějak dvakrát dotknul, já jsem to víceméně bral jako okrajovou záležitost.“ (komunistický funkcionář)

Respondenti indentifikovali vypravěče jako zástupce všech tří identit a ve všech případech dokázali uvedený výrok podpořit odkazem na encyklopedické vědění o dané identitě či době.

Příklad, kdy byl narátor určený jako disident:

**R:** *A teď tady je teda souběh událostí, pro mě klíčový slovo, a ještě bych teda tady možná podtrhla netoužili po tom, aby se něco semlelo. A tady jsem na vahách, protože... Souběh událostí použije někdo, kdo už, kdo vůbec o těch věcech přemejšlí.*

**T:** *Mhm.*

**R:** *A kdo teda možná v těch věcech tak, jak se udály, viděl souběh víc událostí než jenom jednu věc. To znamená, to trošku ukazuje na někoho jako informovanějšího.*

**T:** *Mhm.*

**R:** *No, ale... Je tu netoužili po tom, netoužili, nikdo z nás netoužil po tom, aby se něco semlelo. Takže já bych to tipovala na nějakého disidenta a mohl teda mluvit o něčem, kdy to ještě třeba nebylo připravený a kdy ještě to načasování, [...] prostě mělo bejt jiný, takže třeba nechtěli, aby nějaká, aby už v tu dobu bylo něco vyprovokovanýho.*

(úryvek 21, rodič C)

Narátor určený jako obyčejný člověk:

**R:** *No, to si taky myslím, že to byl takovej normální člověk, kterému nešlo o nějaký změny, aby se bůhvíco stalo a o tom disentu zas tak moc nepřemýšlel. To zas nebylo tak rozšířený mezi lidma, že někdo je v disentu a že někdo něco vymýšlí a že někdo by chtěl něco změnit, nebo tak.*

(úryvek 21, prarodič C)

Narátor určený jako komunistický funkcionář:

**R:** *No tohle říkal nějakej papaláš... komunistickéj. Disent považoval za okrajovou záležitost. No. Voni to podcenili. Pro ně to jako...*

**T:** *Hmm... Takže voni to podcenili.*

**R:** *Já si myslím, že voni to trošku podcenili. No.*

(úryvek 21, prarodič B)

Extrahujeme-li z komentářů explicitní encyklopedické vědění, na jehož základě respondenti zdůvodňují své volby, dostaneme následující informace: disidenti byli o politické situaci informovanější než obyčejní lidé a nehrnuli se po hlavě do konfrontace s režimem; obyčejní lidé příliš o politice nepřemýšleli ani se politicky neangažovali, ať už na straně režimu nebo protirežimně; komunističtí funkcionáři (mylně) považovali dissent za nedůležitou politickou sílu. Lze říci, že encyklopedické kompetence jednotlivých respondentů nejsou v rozporu navzdory tomu, že vedou k odlišnému určení mluvčího. Odvážil bych se i tvrdit, že všechny výše uvedené znalosti by mohly být součástí jedné a téže encyklopedie a odlišné určení vypravěčů má původ v odlišných rejstřících, které text diferencovaně aktivuje.

Z generačního hlediska stojí za zmínku jeden rozdíl v encyklopedické kompetenci, který se vytvořil mezi generací studentů a respondenty z generací, jež socialismus zažily. Znalosti aktérů ve studentských encyklopediích se podobaly typickým charakteristikám postav obývajících minulost (disident, komunistka), zatímco znalosti starších generací odkazovaly převážně

na konkrétní či generalizovanou minulou zkušenost. Takto vysvětluje své určení jedna studentka:

*R: A v posledním... příspěvku ten mluvčí zmiňuje disent, ale já jsem se nakonec přiklonila k tomu, že to je spíš obyčejný člověk než aktivní disident, jako člověk, který vyjadřuje odpor k režimu.*

(úryvek 21, student A)

Další student uvádí:

*R: A poslední... teda... disident... Tak tam mi jako přišlo, že je to disident, který se k tomu dostal tak nějak jako slepý k houslím. Že se stýkal s těmi lidmi. S disidenty. A nějak tak... netoužil po tom, aby se něco semlelo. A tak jako mluví o všech, jako by netoužili... po tom aby se něco semlelo.*

(úryvek 21, student B)

Konečně třetí příklad studentského typologizování:

*R: Tak to jsem teda moc nevěděla, přišlo mi to takový jako, že jsem nevěděla, ale tak... Tak pak mi to přišlo disidentský z hlediska toho, že to chování je takový non-konformní a kdy, kdy jakoby se nechce podřídit tomu, co mu jako nějaká instituce nebo organizace říká, tak tomu, že se vzpírali vlasit nosy, nosit vlasy, takovým způsobem jako všichni ostatní, no, to mi tak jako přišlo, že by to mohlo souviset s tou povahou toho disidenta, no.*

(úryvek 2, student D)

Studenti, zdá se, popisují spíše možný svět historických aktérů, než aby se snažili přiblížit k tomu, co se odehrálo. Ale jejich „typologický“ přístup neznamená, že by se častěji mylili v odhadování identity vypravěčů a toho, o co v úryvku běží. Zatímco encyklopedická kompetence pamětníků umožňuje nalézání a čtenářské aktualizování mnoha kontextů výkladu úryvku a tím vnáší do určování identity vypravěče neurčitost, encyklopedická kompetence studentů založená na typech aktérů a situací umožňuje navrhovat jednoduchá a zdánlivě povrchní řešení, jež nicméně často „fungují“. Všechny generace tak disponují ve vztahu k textu rozdílným věděními o komunistech či disidentech, jak do rozsahu, tak do povahy, ale při kooperaci s krátkým biografickým textem jsou – jak již bylo uvedeno výše – schopni využít vždy jen jeho relativně malou část. Možná je to způsobeno i „historickým“ charakterem encyklopedické kompetence pamětníků, tzn. jsou schopni nalézat možné kontexty a vysvětlení z pohledu popisované doby,

zatímco studenti aplikují spíše taková vysvětlení, která cirkulují v současném diskurzu (uměleckém, dokumentárním, odborném) o socialistické minulosti.

### 3.3.2.6 Konstrukce možných světů

Třetí, nejkompexnější interpretační, či chcete-li čtenářskou kooperací úrovní, je konstrukce možných světů. V Ecově pojetí se možné světy vážou na čtenářskou anticipaci toho, co se v textu bude dít, jsou to propoziční postoje vůči běhu událostí, předvídané stavy jevů a věcí [Eco 2010: 139]. V kontextu našeho výzkumu jsme uvažovali možný svět jako čtenářskou konstrukci, která prostřednictvím vědění aktualizovaného textem (encyklopedie) nabízí možnosti toho, co se mohlo, ale nemuselo odehrát, resp. co se odehrát nemohlo. Možným světem se zde tedy nerozumí svět rekonstruovaný na základě vlastní přímé zkušenosti, na základě individuální paměti a vzpomínek na minulost, ale je s jejich pomocí a pomocí dalšího vědění *návrhem* věrohodných možností dění. Zjednodušeně řečeno, jednalo se o ty úryvkem aktualizované oblasti minulosti, kde čtenář nemohl daný úryvek vykládat výhradně podle vlastní zkušenosti či pozitivní vědomosti, ale pracoval i s interpretacemi, které byl ochoten ještě v daném kontextu připustit, tj. které se mu jevily jako možné.

Hned z prvních analýz bylo patrné, že charakter možných světů, jež respondenti v rozhovorech konstruovali, podléhal v zásadním ohledu experimentálnímu zadání, tj. tomu, aby čtenáři během čtení v úryvcích pátrali po identitních značkách, které by je mohly vést k určení vypravěčovy identity. To stimulovalo a zároveň i limitovalo vytvářené možné světy. Jejich nejčastější struktury se proto vztahovaly k vlastnostem připisovaným jednotlivým kategoriím mluvčích či situací, v nichž se mohli nacházet. Příkladem návrhu možného světa socialismu, ve kterém se vylučuje status kuchařky a komunistické funkcionářky, je následující komentář:

*R: Tak vzhledem k tomu, že bylo řečeno, že pokud se jedná o komunistu, tak to byl... ehm... člověk ve vysoké pozici v komunistické straně, tak se dá předpokládat, že kuchařka asi nebyla ve vysoké pozici a... ehm... vyvolává to prostě dojem běžného člověka za minulého režimu. Když o sobě mluví jako o tom, že byly jsme dvě kuchařky a mluví o vedoucí*

(úryvek 1, student A)

Pro srovnání můžu uvést komentář příslušníka generace rodičů, ve kterém k identifikaci vypravěče postačuje jeho encyklopedická kompetence:

R: No, identifikovala jsem to jako na běžného člena partaje a identifikuju to z toho důvodu, že je tam to slovo že byly protěžovaný a že měla o nějaké peníze na hodinu víc, protože taková praxe byla, že pokud byl někdo v partaji, tak měl nějaký výhody i ten běžný člověk, oproti nestraníkovi, měl.

(úryvek 1, rodič Ca)

Srovnání nám ukazuje i jeden z rozdílů v recepci dokumentů, který se objevil mezi generací studentů a generací jejich rodičů a prarodičů, a sice to, že studenti logicky odvozovali identity mluvčích z jim zprostředkované známé typologie postav a situací (disidenti byli proti režimu a měli se špatně, byli zavíraní; komunističtí funkcionáři podporovali režim a měli se dobře, byli mocní). Studenti častěji konstruovali možné světy, zatímco obě generace pamětníků se spoléhaly na své vlastní vědomosti (a nevědomosti) o minulosti a možné světy navrhovaly méně. Dalším dokladem, že pamětníci měli menší tendenci konstruovat možné světy, může být i to, že když student ve výše uvedeném komentáři činí předpoklad, že kuchařka se neslučuje s komunistickým funkcionářstvím, tak pamětníci naopak činit předpoklady nechtějí: „Jsou to kuchařky, takže nepředpokládám, že je to bolševickéj funkcionář“ (rodič Cb) nebo „nepředpokládám, že nějaká kuchařka byla v nějaké vyšší funkci nebo něco takovýho“ (rodič D). Negativní forma odpovídá využití encyklopedie k vyloučení určitých možností jako nerealistických.

Rozvinutých konstrukcí možných světů se v komentářích nevyskytuje mnoho. Relativně nejvíce je jich v komentářích k úryvku 17:

„...abych začal uvažovat o vstupu do komunistické strany. No a tak samozřejmě, že to bylo otázka jako diskuzí, jednak u nás doma a jednak jako s rodiči. No a oni říkali, tak jako dávali sebe za příklad, že prostě, pokud se takto nestane, tak, že prostě ohrozím vlastně.“ (disident)

Z úryvku je zřejmá situace (diskuze o vstupu do KSČ), ale jen těžko lze odhadnout politickou příslušnost vypravěče. Pravděpodobně proto si výrok vyžádal relativně velkou kooperační aktivitu, která se nezastavila u encyklopedie, ale vedla až ke konstrukci možných světů, ze kterých bylo následně možné odvodit status vypravěče.

R: Protože se tady mluví o ohrožení v případě, že do té strany nevstoupí. ...takže jako by... mě to vede k tomu, že ti rodiče... teda asi nebyli komunisté tělem i duší, že by vyloženě prostě s tou ideologií natolik souhlasili, že by chtěli vstoupit do strany, ale spíš protože to dělali z nějakých, já nevím, praktických důvodů.

T: Mhm.

R: Ze svojí bezpečnosti. A to, že ten člověk... vlastně... taky o tom nemluví nějak přesvědčeně, že chce o tom diskutovat, že chce prostě probrat pro a proti proč do té strany vstoupit anebo nevstoupit. Tak proto... si myslím, že to byl obyčejný člověk.

(úryvek 17, student C)

R: Tady to je výrok obyčejného člověka, kterej tady uvažuje (důraz) o vstupu do komunistické strany, diskutuje o tom s rodiči. To je člověk, kterej tam nakonec třeba vstoupí, ale tak nějak jako... eště ke všemu říká, že když tam nevstoupí, tak se ohrozí, čili jako je vidět že, žádný velký nadšení pro tu ideologii nemá, že to prostě když, v případě, že to udělá, tak že to udělá z existenčních důvodů, je to ta běžná situace obyčejného člověka, kterej prostě nakonec tam pak vstoupil do té strany ale tak, no. Ale asi se pak nikam nedral k nějakým vysokým funkcím, že jo. Protože prostě jenom chtěl mít pokoj.

(úryvek 17, rodič D)

R: Abych začal uvažovat o vstupu do komunistické strany, čili to je mladý člověk z rodiny, jeho rodiče jsou buď sami, sami komunisté anebo lidé, kteří si uvědomují, že by v tom komunistickém režimu by měl podstatně usnadněn, více usnadněnu kariéru, kdyby byl členem komunistické strany.

T: Hm.

R: Čili on ještě není členem komunistické strany, proto to není K[omunistický funkcionář], disident to rozhodně není, poněvač o tom uvažuje, takže je to naprosto jasné. Óčko [obyčejný člověk].

(úryvek 17, prarodič Db)

Povšimněme si, že interpretační kooperace hladce probíhá nezávisle na tom, je-li skutečný vypravěč v souladu se čtenářskou identifikací. V případě úryvku 17 žádný ze čtrnácti respondentů neurčil vypravěče správně jako disidenta, protože všichni zjevně považovali za velmi nepravděpodobné, že by byl svými rodiči přemlouván ke vstupu do komunistické strany a že by váhal. Nicméně i taková interpretace je možná: rodiče nestraníci (z donucení kvůli svému „buržoazního původu“) pragmaticky radí synovi vstoupit do KSČ; jemu se ale nechce, protože s komunistickou politikou nesouhlasí (i vzhledem k osudu svých rodičů), a později ho tento nesouhlas dovede až do řad aktivních odpůrců režimu – disentu. Existují i další verze možných světů, které by vedly k disidentskému statusu vypravěče, například se stal komunistou v 60. letech minulého století, po roce 1968 byl vyloučen a v 80. letech podepsal Chartu a začal se politicky angažovat. Nicméně jak plyne z komentářů, respondenti je považovali za relativně nepravděpodobné. Dostáváme se tak opět k zjištění, že při recepci dokumentu paměti se aktualizují jen některé z responden-

tům dostupných encyklopedických kompetencí a jsou konstruovány jen nejpřavděpodobnější, nejsnazší či nejpřijatelnější možné světy. V důsledku toho se do velké míry znevýznamňuje rozdíl mezi respondenty s rozsáhlými encyklopedickými znalostmi založenými na vlastní zkušenosti a těmi, kteří takovými znalostmi nedisponují a kooperaci provádějí na základě relativně omezených zprostředkovaných vědomostních zdrojů (studenti).

### 3.3.2.7 *Shrnutí sémioticky orientované obsahové analýzy*

Sociologicky vzdělanému čtenáři sémiotický přístup jistě připomíná přístupy fenomenologické sociologie a etnometodologie. Například konstrukce hypotézy významu úryvku z jednotlivých slov a slovních spojení je podobná „dokumentární metodě interpretace“ [Garfinkel 1967: 39–40 a 76–103]. Čtenáři se snažili zkonstruovat smysl úryvku tím, že jej nebo jeho části stavěli do pozice dokumentu (stopy) určitého smysluplného vzorce jednání či myšlení, který byl implikován odhadovanou identitou mluvčího. Když v úryvku našli slovo nebo slovní spojení, které nešlo napoprvé považovat za dokumentující navržený vzorec jednání vážící se k dané identitě, snažili se ho reinterpretovat tak, aby smysl úryvku zůstal zachován, popř. hledali možnosti, jak identitu vypravěče této nové situaci přizpůsobit. V obecné rovině bychom dokonce mohli prohlásit, že analýza čtenářské semiózy je jakousi etnometodologií čtení, protože jde o výzkum racionálních vlastností indexikálních výrazů obsažených v textu a jiných praktických činnostech, kterými čtenář dosahuje porozumění tomu, co text znamená (srovnej Garfinkelovu definici etnometodologie [Garfinkel 1967: 11]); podobný přístup lze nalézt i v [Livingston 1995; McHoul 1982]. Sémiotický přístup je v souladu s etnometodologickým také v reflexivitě, tj. v tom, že metody (autorské) produkce významu v textu považuje za identické s těmi, které umožňují jeho porozumění (čtenářem).

Pro analýzu čtenářské encyklopedické kompetence při identifikaci vypravěčů se zase nabízí Schützův koncept typizací a příručního balíku vědění vzniklého sedimentací jednotlivcových dřívějších zkušeností. Podobně jako jednotlivé encyklopedické kompetence jsou čtenářskou kooperací s textem aktivovány, tak i celek vědění každého člověka je „spící“ a teprve při jednání se určitá část aktivuje podle principu relevance [Cox 1978: 90].

Sémiotická obsahová analýza splňuje nároky definice obsahové analýzy, tj. na základě zkoumání textů přináší inference o okolnostech užití textů. Technika sledování kooperace s texty sice zkoumá texty jen nepřímou, sledováním způsobů, jak jejich čtenáři vypovídají, že jim rozumějí. I tak přináší hodnotná zjištění o vytváření obsahů textů v souvislosti s okolnostmi jejich užití.

## 3.4 **Strojový čtenář – computer assisted text analysis (CATA)**

Když se výzkumník rozhodne analyzovat nějaký text nebo soubor textů za účelem generování jejich obsahu, narazí vždy na zásadní otázku, která se týká kódování, tzn. způsobu, jakým převést „pomíjivé, nestrukturované nebo neurčité, nicméně jinak zcela smysluplné fenomény do pojmů jazyka dat, ve kterém mohou být analyzovány pomocí náležitých technik“ [Krippendorff 2004: 125]. Nejběžnější postupy kódování textu se zakládají na manuálním přiřazení určité oblasti textu určitému kódu lidským kodérem, například výrok „s dcerou jsme se neustále hádaly“ by mohl být podle předem určených instrukcí kódován jako „konflikt v rodině“. Tento postup má svou silnou stránku: kodér je vyškolený člověk, je schopen pochopit význam analyzovaného textu a více či méně validně i reliabilně přiřazovat kategorie; neboli je schopen se vypořádat s přirozenými jazykovými fenomény polysémie (víceznačností) a konotace. Je schopen odvodit, že v textu „mám dceru ráda, ale občas se spolu dostaneme do křížku“, výraz „dostat se do křížku“ znamená mít spor, a tudíž úsek správně přiřadí kódu „konflikt v rodině“. Na druhou stranu manuální kódování má i slabé stránky. První souvisí s tím, co bylo uvedeno už v úvodu k této knize – množství textu k analýze. Manuálně lze kódovat několik biografických rozhovorů do cca dvou desítek kategorií, ale už ne všechny články obsahující slovo nespravedlnost v českých celostátních denících v posledním desetiletí, kterých je přes deset tisíc. Jen samotné přečtení by zabralo týdny. (I když by bylo teoreticky možné zaměstnat desítky vyškolených kodérů, nevyhne se problému s reliabilitou takového kódování; „intercoder reliability“ je ovšem problém, do kterého se zde nemůžeme pouštět a čtenáře proto odkazuji na [Krippendorff 2004, kap. 11].) Druhý problém spočívá v samotném kódovacím klíči. Klasická obsahová analýza počítá s předem vytvořeným kódovacím klíčem. Například ve výzkumu mediální reprezentace nespravedlnosti můžeme kódovat podle principů nespravedlnosti vytvořených v nějakém předchozím kvalitativně orientovaném výzkumu na dané téma. Co však v případě, že nechceme aplikovat na našich datech nějaký již vytvořený kódovací klíč, ať už z důvodu, že máme pochyby o jeho použitelnosti na naše data nebo nás zajímá induktivně orientovaný explorační přístup, například o čem se v člancích o nespravedlnosti nejčastěji píše, v jaké souvislosti se koncept objevuje? V takovém případě je manuální kódování problematické právě z téhož důvodu, který z něj činí silný nástroj pro teoreticky zakotvené kategorie – pro neodstranitelný sklon kodérů interpretovat text vždy v kontextu. Výrok „vyrostl jsem v pocitu, že je nespravedlivé, když člověk, který má víc odvahy než ostatní, stojí na okraji“ může být v publikovaném rozhovoru se známým politikem kódován jedním kodérem do oblasti

nespravedlnost-politika, jiný do nespravedlnost-socializace/výchova, jiným třeba do oblasti nespravedlnost-pocity/emoce. Různost zvolených kontextů či oblastí významu ovšem není chybou kodérů, nýbrž vlastností lidského čtení a interpretace textu. Čtenář s textem kooperuje a jeho význam svým čtením spoluvytváří. I když jen málokdy se čtení textů kodéry diametrálně liší, zhusta se můžeme setkat s tím, že si kodéři nedokážou vybrat z možných verzí dominantní kontext a v důsledku toho nevědí, jakým kódem nebo kategorií mají daný text kódovat. V případě exploračního výzkumu to je zásadní překážka neboť kodérům nelze poskytnout obecný návod, jak v situaci kontextové nejistoty postupovat.

Z obou výše uvedených problémů – neschopnosti lidských kodérů zpracovat velké množství textů a rozhodnout o kontextu v případě exploračního přístupu – vedlo výzkumníky v textové analýze k hledání automatického způsobu induktivně orientovaného kódování. Zadání bylo jednoduché: a) jak kódovat velké množství textu, který je k dispozici v elektronické podobě a b) jak zaručit kvalitu kódování, aby zachycovalo relevantní atributy textu efektivně a spolehlivě. Odpověď na první úlohu je jednoduchá – kódovat musí stroj. Odpověď na druhou otázku závisí na tom, co považujeme za relevantní atributy textu.

### 3.4.1 Frekvenční analýza textů

Jestliže nás jako relevantní atribut určitého textu zajímá jeho téma (o čem text je), zjištěná nejfrekventovanější plnovýznamová slova se zpravidla k tématu vážou. Již tradičně bývá frekvence výskytu určitého prvku v daném prostředí – nejen v textové analýze – považována za indikátor dominantní povahy tohoto prostředí (mnoho stromů značí les, mnoho vojáků armádu nebo bitvu). Kupříkladu nejčastější významová slova této kapitoly [3.4] jsou *text*, *slovo*, *spoluvýskyt*, *korpus*, *graf*, která bezesporu její téma indikují. Metoda analýzy textů založená výhradně na samostatně posuzovaných frekvencích jednotlivých slov však poskytuje pouze velmi přibližná zjištění a hodí se jen na určité druhy textů, zejména odborných, které jsou tematicky vyhraněné. Pro analýzu neobdobných textů, například biografických rozhovorů nebo mediálních sdělení, není vhodná, protože tyto texty se frekventovanými slovy příliš neliší. Také při výzkumu novinových článků bychom sice dokázali tematicky odlišit články o ekonomice od článků o sportu, ale to k sociologicky zajímavým zjištěním nepostačuje. V současnosti navíc již existují sofistikované metody, jak téma textu nebo jeho shrnutí z frekvencí slov, jejich vzájemných vztahů a klasifikací, odhadnout; pro přehled viz [Berry, Kogan 2010;

Dale, Moisl, Somers 2002; Song, Wu 2008]. Z odborného pohledu však nejde o techniky analýzy, které by měly přinést nějaké nové vědění nebo odhalit nějaké netriviální struktury, jako spíše o metody zpracování textových dat za účelem automatického získání informací v textu již obsažených. Toto jejich zaměření je zřejmé z jejich anglických názvů *natural language processing* nebo *information retrieval*. Nicméně mohou sloužit jako pomocné techniky při sociologickém výzkumu vycházejícím z velkého množství textových dat. Jako příklad můžeme uvést analýzu proměn témat v on-line diskuzních skupinách nebo sledování objevujících se témat ve webozích.

Pokud chceme jít dál než je zjištění tématu v nějakém textu nebo souboru textů, tzn. chceme-li se pustit do objevování nového, potom se pro nás stane relevantním atributem textu spíše jeho struktura nebo struktury (vzorce) než téma. Posunuje nás to z oblasti, kterou bychom označili jako čtenářskou, do oblasti ne-čtenářské: zjištění tématu textu dokáže člověk-čtenář, zatímco zjištění vnitřní struktury rozsáhlého souboru textů je čtenářem jen těžko postižitelné. Tradiční obsahová analýza si sice činí ambice odhalit „fenomény, které nejsou přímo pozorovatelné,“ ale spoléhá v naprosté většině případů buď na kodéry-čtenáře, kteří čtením odhalují a kódují mikro- či makro-témata, případně jejich hodnocení (negativní, pozitivní, neutrální), nebo na ex-ante vytvořené kategorizační slovníky a tezaury, které čtenářské chápání předpokládají [Carley 1993]. To znamená, že stojí buď na čtenářských, nebo teoretických předpokladech.

Induktivně orientovaná počítačová textová analýza (*computer assisted text analysis*, CATA) čtenářskou dovednost nepotřebuje, přesněji nepotřebuje ji v analytické fázi výzkumu. Předpokládá minimální znalost textů, které vstupují do analýzy, a snaží se explorativním způsobem zjistit jejich strukturní charakteristiky nebo vzorce, které se v nich objevují. Za základní jednotku textu je považováno slovo a strukturou textu se rozumí soubor vztahů mezi slovy, resp. vztahy mezi různými strukturami a vzorci slov. Jednu z takto orientovaných metod textové analýzy představím v následujícím oddíle.

### 3.4.2 Metoda sledování spoluvýskytů slov – příklad analýzy biografických vyprávění

Jako ilustrace CATA by mohl posloužit výzkum mediální komunikace aktivistických skupin, který budu popisovat v kapitole 4.5; v jeho rámci jsem se spolu se spolupracovníky snažil popsat a prozkoumat aktivistický diskurz. V této části jsem se ale rozhodl předvést sledování spoluvýskytu slov jako specifické techniky obsahové analýzy na výzkumu přepisů biografických vy-

právění. Cílem je prostřednictvím analýzy vyprávění jako textů získat informace o okolnostech jejich vzniku. Konkrétně se výzkumná otázka týká rozlišování soukromého a veřejného, tedy zda a případně jak rozlišují pamětníci ve svých životopisných vyprávěních sféru soukromého a veřejného života a zda se u různých sociálních skupin vyprávěčů tyto sféry liší. Často totiž můžeme slyšet, že státní socialismus prakticky vytěsnil autentickou veřejnou sféru a že jediný možný prostor pro smysluplnou seberealizaci člověka, zvláště v posledních dvaceti letech socialismu, se nacházel v soukromí, v rodině, a u nás se projevoval zejména chatařením, kutilstvím a nejrůznějšími koníčky.

Nejprve se věnujme samotné pojmové dichotomii veřejného a soukromého. V souvislosti s pádem státního socialismu ve střední a východní Evropě se otázka rozlišení veřejného a soukromého objevuje zejména v ekonomickém kontextu, tzn. ve vztahu k masivní privatizaci do té doby státem vlastněného majetku [např. Mertlík 1997; Mlčoch 1998]. Samotné označení „privatizace“ jasně vymezuje orientaci tohoto procesu, kdy majetek přechází z rukou státu do rukou soukromých, ať už fyzických nebo právnických osob. Druhá oblast odborného diskurzu, ve které má dichotomie veřejné/soukromé privilegované postavení, je oblast genderu, přesněji genderově podmíněných změn ve společnosti po roce 1989. Interpretace těchto změn vycházejí z představy, že státní socialismus dal vznik specifickému genderovému řádu [Vodochodský 2007; Havelková 1993; Wagnerová 1995]. Klíčová pro tento řád byla skutečnost, že ženy byly v důsledku tlaku na jejich všeobecnou zaměstnanost vtlačeny do veřejného prostoru, zatímco zrušení soukromého podnikání a nomenklaturní systém obsazování řídicích pozic uzavřel mnoha politicky neangažovaným mužům možnost profesní kariérní seberealizace. Po roce 1989, k překvapení feministek ze zemí se silným feministickým hnutím, neprojevovaly ženy v postkomunistických zemích na rozdíl od mužů tak masivně ambice po veřejné kariéře, naopak, mnohé pozitivně oceňovaly možnost (nikoli nutnost) opustit nudné či špatně placené zaměstnání a věnovat se více rodinnému životu v širokém smyslu slova [Fodor 2002; Havelková 1994; Heinen 1997]. Třetí oblast, kde se polarita veřejného a soukromého diskurzivně objevuje, i když ne v plně explicitní podobě, je oblast politická. Jedná se o případ často diskutované „nepolitické politiky“ nebo „antipolitické politiky“ [např. Havelka 1998; Otáhal 1998], pod kterou se dnes rozumí především odpor vůči demokracii spočívající výhradně na soutěži politických stran a důraz na nezastupitelnou roli tzv. občanské společnosti. V období státního socialismu však termín označoval především Havlovu představu děláni politiky v posttotalitním systému skrze „život v pravdě“ prostřednictvím „mravní rekonstituce společnosti,“ tzn. politiku vycházející z privátního světa

jednotlivce [Havel 1990: 129]. Jelikož veřejná politika byla v reálném socialismu zbavena jakéhokoliv smyslu kromě reprodukce moci, bylo podle Havla i dalších třeba hledat útočiště v morálním zázemí člověka jako jednotlivce.

Cílem tohoto stručného a nutně zjednodušujícího představení tří oblastí užití pojmové dichotomie soukromého a veřejného bylo ukázat, že hledat nějaký společný univerzální základ či výklad by bylo marné. Soukromé může znamenat podle kontextu například formu vlastnictví, formu skupinového bytí nebo prostor existenciální základny smyslu lidského bytí.

Kromě mnohoznačnosti je rozlišení soukromého a veřejného podle antropoložky Susan Gal (a nejen podle ní) ve své podstatě ideologické; jde o případ „jazykové ideologie,“ jak to nazývá [Gal 2005], podobně jako dichotomie práce/domácnosti, racionality/emoci, mužského/ženského apod., a za jejím použitím je třeba hledat zájmy a mocenské vztahy. Některé oblasti sociálního života jsou prohlašovány za veskrze soukromé, protože není žádoucí, aby vstupovaly do tržních vztahů (např. rodinná péče, výchova), zatímco jiné oblasti jsou deklarovány jako veřejné a tržnímu prostředí vystaveny být mohou (placená asistence u potřebných, vzdělání). Konkrétní podoba rozlišování tak odráží konkrétní zájmy sociálních skupin. Další charakteristikou pojmové opozice veřejného a soukromého, kromě již zmíněné obsahové vágnosti a ideologické zakotvenosti, je to, že jde o „fraktální rozlišení,“ tzn. jednotlivé části opozice lze opět dichotomicky rozdělovat podle stejného klíče [ibid.]. Jestliže domov je soukromý prostor, tak v jeho rámci máme též veřejné a soukromé oblasti (např. obývací pokoj či kuchyně versus ložnice). A rámci ložnice jsou šuplíky nočního stolku soukromější než jeho vrchní deska. Podobně lze fraktálně rozlamovat na veřejné a soukromé politiku, práci, veřejné prostory apod. Například v zaměstnání existují oblasti soukromí zaměstnanců (osobní skříňka, vlastní kancelář) a oblasti veřejné (místnost pro setkání s klienty) a o proporci těchto oblastí se mohou vést spory (sledování elektronické pošty zaměstnavatelem, zaznamenávání telefonických hovorů).

Stojíme nyní před otázkou, jak přistoupit k výzkumu soukromého a veřejného prostřednictvím biografických textů. Jak jsem se pokusil ukázat, nelze předpokládat, že jsou skutečnosti, které by byly ze své podstaty soukromé nebo veřejné; platí opak: soukromé a veřejné je a) předmětem a výsledkem vyjednávání a mocenských tlaků a b) závislé na perspektivě a kontextu. Z toho důvodu je nutným metodologickým předpokladem výzkumu relační a komparativní přístup, který bude brát soukromé a veřejné jako důsledek určitých vztahů a interakcí.

Pro svůj výzkum soukromého a veřejného před rokem 1989 jsem vyzkoušel orálně historická biografická vyprávění pamětníků. Biografické rozhovory zde mají tu přednost, že zachycují veřejné a soukromé skrze subjektivní

zkušenost. Na druhou stranu jistým omezením je to, že nejsou historickými daty v pravém slova smyslu, protože se jedná o současná vyprávění, minulé životy prezentované z dnešní pozice a pro aktuální účely, takže jsou patrně zejména ty skutečnosti, které kontrastují vůči dnešku.

S vědomím těchto omezení jsem se pokusil zjistit, jak se v rozhovorech objevuje rozlišení soukromého a veřejného, přičemž jsem využil toho, že vypravěče lze rozdělit podle pohlaví a podle profese (na muže a ženy a příslušníky dělnických profesí a tzv. inteligence). Předpokládal jsem, že vzájemné porovnávání těchto skupin může odhalit sdílené a variabilní aspekty studované dichotomie.

#### 3.4.2.1 Vytvoření korpusů

Přepsané rozhovory pro analýzu jsem získal ze sbírky *Biografie příslušníků dělnických profesí a inteligence*, která byla vytvořena v Centru orální historie Ústavu pro soudobé dějiny AV ČR a se kterou jsem pracoval již v předchozím výzkumu. Jedná se o soubor cca sta biografických rozhovorů s lidmi rozmanitých profesí, mužů i žen, z různých částí republiky, z měst i venkova. Hned na začátku jsem se rozhodl, že budu vycházet z rozdělení vypravěčů, které poskytuje sbírka, tzn. na muže a ženy a na příslušníky dělnických profesí a příslušníky tzv. inteligence, což jsou lidé s vysokoškolským vzděláním pracující v odborných a manažerských funkcích. Tím vznikly čtyři skupiny, jež jsem označil jako: muži-dělníci, ženy-dělnice, muži-inteligence a ženy-inteligence.

Korpusy rozhovorů pro jednotlivé skupiny jsem vytvořil sloučením všech rozhovorů stejné kategorie vypravěčů do jednoho velkého souboru. Výsledkem byly čtyři velké soubory. Poté jsem z nich poloautomaticky odstranil otázky tazatelů, které by mohly ovlivnit výsledky systematickou chybou. Při práci s mediálními texty je nutné zase odstranit (nejlépe poloautomaticky pomocí vytvořeného skriptu nebo makra v textovém editoru) průvodní informace k článkům (jméno autora, popis autora, popis fotografie, název rubriky či sekce apod.) nebo opakující se záhlaví a zápatí stránek. Takto vytvořené a očištěné korpusy jsou připraveny pro analýzu.

#### 3.4.2.2 Frekvenční slovníky

Vytvoření frekvenčního slovníku je prvním krokem analýzy. Programů, které jsou schopny počítat frekvenci slov v textovém souboru, existuje celá řada, např. Atlas.ti, TextStat, AntConc. Pro každý korpus postupujeme stejně. Nejprve spočítáme frekvence jednotlivých slov, která seřadíme od nejčetnějších po nejméně četná. Protože nás ve vyprávění zajímají významové oblasti, vybereme do slovníku jen plnovýznamová slova nejčetnější, zpravidla podstatná jména. Jde nám totiž o to, abychom v následujících krocích mohli pra-

covat jen se samotnými slovy, bez ohledu na bezprostřední kontext, v němž se v textu vyskytla.

Při sestavování slovníků stojíme před dvěma otázkami: jaká slova a hlavně kolik slov do analýzy zahrnout. Mezi nejčastější slova v obou korpusech se řadí také slova označující čas nebo trvání (*dneska, tenkrát, tehdy, rok, nikdy, pořád*), která by teoreticky mohla indikovat ve vyprávění časové struktury, nicméně mimo svůj konkrétní kontext výskytu jsou jen těžko interpretovatelná. *Tehdy* může znamenat v roce 1948 stejně jako v roce 1989. Podobně jsou na tom častá slovesa *dělat, jít, myslet*, která sice mají vlastní význam, ale pro analýzu, jež se ve vyprávěních snaží odhalit životní sféry, nejsou dostatečně jednoznačná. Co se týče množství slov zahrnutých do analytického slovníku, jsme omezeni na jedné straně zobrazovací schopností výsledných grafů, které jsou čitelné zhruba do padesáti až šedesáti zobrazovaných slov, a jednak velikostí korpusu, přičemž platí, že čím větší korpus, tím více slov má smysl do slovníku zařadit. U malých korpusů hrozí, že relativně málo frekventovaná slova zahrnutá do analýzy budou v grafu tvořit přirozené outliers stojící mimo ostatní struktury, a tudíž nebude možné je smysluplně interpretovat. Protože v našem případě jsou korpusy biografických vyprávění použitých dostatečně veliké (přes půl milionu slov každý), vybral jsem padesát nejčastějších slov, výhradně podstatných jmen.

Technicky je procedura výběru slov do slovníku složitější než jen vypsání padesáti nejčastějších. Vzhledem k tomu, že čeština má ohebná slova, je prakticky potřeba vybrat nikoli nejčastější jednotlivé formy slova, ale slovní základy neboli lemmata. Nejlépe by se pracovalo s plně lemmatizovanými korpusy, ve kterých bychom vybrali lemma, například *škola*, a automaticky by se do slovníku zařadily i všechny gramatické tvary tohoto slova (*škola, školy, školou, školami* atd.). V nelemmatizovaných textových korpusech postupujeme tak, že nejprve vybereme cca trojnásobné množství nejčastějších slov, tato převedeme na slovní základy (např. *babič, sestr* apod.) a poté zjistíme výskyty všech slov odvozených od těchto slovních základů. V některých případech ale takto postupovat nemůžeme, protože slovní základ je společný různým slovům. Když zůstaneme u členů rodiny, kořen *mám* slov *máma, mámou* atd. je totožný s první osobou singuláru slovesa *mít* a tudíž bychom sčítali frekvence dvou různých slov. V takovém případě musíme postupovat odlišně. Použijeme k tomu program COOA [Hájek, Hájek 2009], který jsem pro účely analýzy spoluvýskytů vytvořil a který je volně ke stažení. Tento program umožňuje počítat frekvence jak slovních kořenů, tak i několika forem téhož slova v případech, kde slovní kořen využít nelze. V případě slova *máma* kumulujeme četnosti jednotlivých pádů jednotného a množného čísla, ale i synonymních výrazů *matka, mamča, mamka, mamina* apod.



Při tvorbě slovníku se musíme vypořádat ještě s jednou obtíží, a to jsou slova nesoucí více významů. Snad největší potíže činí slovo *stát*, které je ve-lice časté a znamená a) politickou instituci, b) vzpřímený postoj, c) cenu („stát tolik a tolik“) nebo d) dosažení určitého stavu či pozice („stát se prezidentem“). Podobně víceznačnými slovy jsou slova například *strana* (politická, strana sporu, „na druhou stranu“, „pokračování na další straně“), *společnost* (firma, skupina, národ, organizace), *svoboda* (hodnota, vlastní jméno) nebo *demokracie* (politické uspořádání, část názvu politické strany). Jediným způsobem, jak se s významovou mnohoznačností (polysémií) vypořádat, je manuálně pozměnit slova tak, aby byl každý význam rozlišitelný, resp. aby byl rozlišitelný význam, který chceme zařadit do slovníku, od významů, které sledovat nechceme. V případě biografických vyprávění jsem všechna slova *stát*, kdy se nejednalo o politickou instituci, ručně změnil na *\_stát*, slova *strana*, kde se nejednalo o politickou stranu na *\_strana* a *společnost* jako firmu na *\_společnost*. K tomu, abychom zjistili, v jakých různých kontextech a tudíž i významech se dané slovo v korpusu vyskytuje, použijeme prakticky ve všech programech pro textovou analýzu dostupnou funkci KWIC, znamenající keyword in context, s jejíž pomocí se nám zobrazí seznam výskytů hledaného slova a jeho variant v kontextu okolních slov tak, jak se v korpusu vyskytuje.

#### Tabulka 4 Prohlížení kontextu slova stát pomocí funkce KWIC

ip, že dělníci předstíraj, že pracují a **stát** předstírá, že je za to platí. A to byl vtip jako no a ten, kdo neměl známosti, tak musel **stát** frontu. Ne. To byl takovej průměr. To bych řekl, kne „stát je nejhorší hospodář“, no ale **stát** jsem já jako ministr, protože je reprezentuju, ta ebo je pak nějak obnovili, takže každej **stát** socialistickej, nebo každá ta prostě společnost u e zas otázka stylu práce, protože každý **stát** asi si kontroluje své nepřátele nebo pochybovače se týká umění nebo zločinnosti. Když se **stát** otevře tomu světu okolnímu, tak už je součástí té rajdaj někde v Paříži a stará se vo to **stát**, to jde z našich kapes, že jo. Jo a to je to, co enalo to jít si stoupnout do té fronty, **stát** se bojovníkem frontovým a prostě si vystát tu kní vis šesti hrnečků a že jsem tedy okradl **stát** atd. Takže hledalo se leccos a nevím. No a najedn potom tvořit a zpívat, tak vlastně ten **stát** nebo ten režim rozbouřoval daleko víc než kdokoli i hodně vysoké plat. Takže měl jsem se **stát** docentem, ale nějak jsem pověsil tu kariéru, já j takový epizody, i takový se můžou teda **stát**. Ale to nevím, jak toto budete psát. To. (smích)

Tak můžeme zkontrolovat, že pozměněné slovo *\_strana* znamená vždy politickou stranu, nebo *bratr* znamená příbuzného a nikoli člena nějaké organizace (skauti, strana lidová, některá církev). Tato kontrola by měla být co nejdůkladnější, aby se předešlo vzniku artefaktálních výsledků zpětně jen těžko odhalitelných.

#### Tabulka 5 Kontrola poloautomatických úprav mnohoznačného slova strana (*\_strana* = politická strana; *-strana* = ostatní významy)

uze volebních ale měla by být průběžná *\_strana* by si měla vždy zachovat akceschopnost. V informa k volebním kampaním neexistuje jako by *\_strana* hledala jen politicky a odborně připravené osobno já. Na politické scéně chybí levicová *\_strana* která by jim dala víru ve spravedlnost pořádek a t že je nám jedno že nám bude vládnout *\_strana* která dokazuje jak jsou jí cizí zájmy obyčejných u jí cizí zájmy obyčejných zaměstnanců *\_strana* která je na druhé *-straně* barikády. Pokud se nic tuace nejvážněji svědčí o tom že by se *\_strana* mohla dostat na sestupnou dráhu což se jako první o odborářům běhat mráz po zádech. Tato *\_strana* má po posledních vyhraných volbách v Senátu Parla žití jen nahodil. Na druhé *-straně* má *\_strana* málo odborníků pro komunální sféru. STYL stranick a ře-šení netradičních problémů. Tím že *\_strana* nemá dlouhodobou strategii a neracionalizovala sv úlohu aparátu za velmi důležitou žádá *\_strana* nemůže bez jeho kvalitního složení existovat. Něk volebního systému kdy je volena určitá *\_strana* není ani on osobně zvolen. Je dosazen *\_stranou*. P

Posledním krokem při vytváření analytického slovníku je doplnění jed-notlivých slov o synonyma vyskytující se v korpusu, například *táta* a *otec* nebo *firma*, *společnost* a *podnik*. Výsledný slovník vypadá takto:

#### Tabulka 6 Ukázka prvních šesti položek slovníku pro vyhledávání spoluvýskytů slov v přepisech biografických vyprávění. Hvězdička je zástupný znak a označuje jakékoliv další znaky.

1. auto, autě, auty, autu, aut, autům, autem, autech, auta, aut\*, vůz, vozu, vozem
2. babič\*
3. bratr, bratra, bratrem, bratrovi, bratrův, bratři, bratrů, bratřím, bratry, bratrech, brách\*
4. brig\*
5. byt, bytů, bytě, byty, bytu, bytům, bytem, bytech, bytov\*
6. člověk\*
7. ...

#### 3.4.2.3 Sledování spoluvýskytů slov

Jak bylo uvedeno výše, technikou, jak zviditelnit struktury textu v korpusu, je sledování spoluvýskytů vybraných slov. Obecný princip je popsán Ido Daganem [Dagan 2000]. V originálním pojetí je sledování spoluvýskytů slov, tj. kolokací, lingvistickým nástrojem ke zjištění jejich významu a dalších charakteristik [Čermák, Šulc 2006; Firth 1957]. Cílem většiny nelingvi-stických kolokačních studií je ze vzájemné četnosti spoluvýskytu slov odvodit jejich vzdálenost sémantickou [Bernard, Ryan 1998; Mohammad, Hirst 2005; Mohr 1998]. Spoluvýskytů lze ovšem sledovat i z jiných důvodů než je zjištění jejich významu. Stačí, když Firthovu distribuční hypotézu („you shall know a word by the company it keeps“ [Firth 1957: 179], viz též kap. 2.5.3) obrátíme: zajímá nás, jak jsou určitým mluvčím nebo skupinou mluvčích sdružována slova, jejichž význam v textu alespoň přibližně známe. Lze to for-

mulovat i negativně, tak, že zkoumáme, která slova lidé společně nikdy nebo téměř nikdy nepoužívají. Přitom nám, jako sociálním vědcům, nejde o poznání jazyka a vztahů slov v jeho rámci, ale jak je jazyk používán sociálními aktéry, jaké sociální skutečnosti lze skrze určité promluvy (texty) ozřejmit.

Od teoretických úvah se přesuňme k technické rovině metody. Máme-li připraven analytický slovník 20–50 nejfrekventovanějších slov (tj. záznamových jednotek, recording units), je na řadě rozhodnout, v jakých kontextových jednotkách (context units) budeme spoluvýskyty sledovat. Podobně jako při vytvoření slovníku i zde provádíme analytické volby podle povahy dat a výzkumných otázek. Technicky vzato, můžeme volit mezi kontextovými jednotkami přirozenými nebo umělými [pro přehled viz Krippendorf 2004: kap. 5]. Přirozenou kontextovou jednotkou je věta, odstavec či celý článek/rozhovor; jednotkou umělou je námi definovaný úsek textu, například stanovený počet slov, vět, odstavců nebo oblast mezi specifickými znaky či slovy (např. osobními zájmeny). Obecně platí, že v krátkých kontextových jednotkách (např. deset slov, věta) se více zvýrazní obecně jazykové (gramatické a stylistické) jevy, například sousloví, spojení subjektu s aktivitou apod., zatímco v jednotkách globálních (např. odstavce či články) převažují tematické aspekty textu v korpusu [Dagan 2000: 463–464]. Zajímají-li nás slovní spojení a těsné vazby slov (tzv. kolokace), pak je vhodné sledovat spoluvýskyty slov v krátkých kontextových jednotkách, zjišťujeme-li sémantická pole, jsou vhodnější středně dlouhé až dlouhé kontextové jednotky (od věty po několik vět). Často je ale naše volba omezena povahou textů, jimiž disponujeme. V případě analýzy webových stránek nebo biografických rozhovorů nelze jako kontextové jednotky použít odstavce, protože nejsou vždy zjistitelné.

Analyticky nejoblíbenější kontextovou jednotkou je věta definovaná zpravidla jako úsek textu nacházející se mezi dvěma tečkami, resp. několika po sobě jdoucích vět – z toho důvodu, že věta či souvětí většinou vyjadřují určitou myšlenku. Několik po sobě jdoucích vět se často vztahuje k jednomu tématu, i když na tento vztah záleží na typu textu. Pro analýzu biografických vyprávění jsem jako kontextovou jednotku zkusil nejprve jednu, tři, pět a deset vět, abych zjistil, jaká délka kontextové jednotky je pro zachycení variability korpusů nejvhodnější. Nakonec jsem analýzu spoluvýskytů provedl na kontextové jednotce pěti vět.

Pro výpočet spoluvýskytů slov z analytického slovníku jsem použil již zmíněný program COOA. Program počítá párové četnosti dvojic slov obsažených ve slovníku (přičemž se všemi slovy v řádku nakládá jako s variantami téhož slova) v rámci kontextových jednotek a výsledné četnosti normalizuje zvoleným koeficientem podobnosti. Podle mých zkušeností i experimentálních studií je nejrobustnější Jaccardův koeficient podobnosti (S) [Chung, Lee

2001]. Počítá se podle rovnice  $S(x, y) = a / (a+b+c)$ , kde  $x, y$  jsou proměnné (slova ze slovníku),  $a$  je počet společných výskytů  $x$  a  $y$  v korpusu,  $b$  je počet samostatných výskytů  $x$  v korpusu a  $c$  je počet samostatných výskytů  $y$  v korpusu; Jaccardův koeficient je považován za vhodnou míru asociace párově se vyskytujících objektů. Výstupem programu je jednak přehled frekvencí jednotlivých slov (položek) slovníku v korpusu, dále frekvence slov v kontextových jednotkách a hlavně matice normalizovaných vzdáleností mezi slovy. Volitelně si můžeme pro kontrolu vypsat všechny kontextové jednotky s vyhledávanými slovy.

#### 3.4.2.4 Vizualizace spoluvýskytů

Získaná matice vzdáleností slov v korpusu je sice přímo interpretovatelná a jednoduše můžeme zjistit, která slova jsou si v korpusu blízko, tzn. vyskytují se relativně často spolu v kontextových jednotkách, a která slova naopak mají spoluvýskytů minimum, ale cílem metody je získat celkový obraz vzdáleností slov. Toho je možné dosáhnout různými klasifikačními nebo škálovacími technikami, které vycházejí z distančních matic, z nichž nejběžnější jsou shluková analýza (cluster analysis) a mnohorozměrné škálování (multidimensional scaling, MDS). Shlukovou analýzu je vhodné použít tehdy, když chceme odhalit oddělené skupiny slov vyskytujících se v textu blízko sebe, například při zjišťování množství diskretních témat textu (monotematické vyprávění vs. polytematické), a následně můžeme testovat diskretnost shluků analýzou variance. Obecně řečeno, shluková analýza se hodí tehdy, pokud chceme nalézt *lokální* minima distancí, dvojice, trojice blízkých slov, které mají ostatní slova jako své satelity [Aldenderfer, Blashfield 1984; Everitt et al. 2011]. Naproti tomu technika mnohorozměrného škálování je procedurou, která zjišťuje *globální* minimum distancí celého souboru slov, tzn. je optimalizovanou projekcí  $n$ -rozměrného prostoru matice vzdáleností (či podobností) do prostoru s nižší dimenzionalitou [Cox, Cox 2010; Hebrák et al. 2005: 145; Shepard 1980]. Její použití je na místě tam, kde klademe větší důraz na interpretaci celkové struktury textu než na interpretaci jednotlivých vztahů mezi slovy jako v případě shlukování. To bylo i cílem výzkumu, na kterém ilustruji popisovanou metodu.

Výsledkem MDS je „mapa“ bodů (slov), v níž jsou si bližší ty, které jsou si bližší rovněž v korpusu (jak již bylo uvedeno, blízkost v textu se zde rozumí frekvence spoluvýskytů v úseku pěti po sobě jdoucích vět). Poněvadž vzdálenosti mezi slovy v textu (spoluvýskyty) nemají metrický charakter, nabízí se jen nemetrická varianta MDS, která ordinalizuje vzájemné vzdálenosti mezi slovy. Z hlediska počtu zobrazených dimenzí se přidržuji, jak je doporučováno [Cox, Cox 2010: 88], dvourozměrného zobrazení, které sice

mívá relativně vysokou hodnotu stresu (vzdálenosti mezi prvky v zobrazené dvourozměrné konfiguraci jen částečně odpovídají vzdálenostem v původní n-rozměrné konfiguraci), ale na rozdíl od zobrazení třírozměrného poskytuje snadnější interpretaci zobrazené konfigurace.

#### 3.4.2.5 Interpretace získaných grafů

Mnohorozměrné škálování je technika především vizualizační, tzn. její funkcí je poskytnout grafické zobrazení jako podklad k interpretaci strukturálních charakteristik dat. Zobrazení má podobu grafu znázorňujícího body na ploše a připomíná proto mapu. To svádí k intuitivnímu čtení, které se zaměřuje na vzájemnou blízkost či vzdálenost jednotlivých slov tvořících shluky. Správná interpretace, zvláště při vyšších hodnotách stresu, však spočívá ve sledování celkových strukturálních charakteristik, jakými jsou polarita, rozptýlenost/sevrěnost, oddělenost/spojitost, jádro/okraj. V předchozích výzkumech se nám ukázaly jako relevantní dominantní rysy polarita [Hájek et al. 2006a; Hájek, Kabele 2010] a jádro/okraj [Hájek, Bayer 2007].

Kromě celostního čtení konfigurace je třeba při interpretaci grafů vzít v úvahu následující skutečnosti:

- a) Jednotlivé dimenze nemají substantivní smysl a graf lze proto libovolně otáčet a převracet; velmi četná slova, která prostupují rozsáhlé oblasti textu (např. slova *lidé* a *člověk*), jsou zobrazena blíže středu grafu, de facto vymezují střed konfigurace, a jejich vzájemné pozice jsou velmi stabilní; zatímco slova relativně méně častá a frekvenčně nesvázaná s jinými slovy se nacházejí na okrajích grafu a jejich pozice se při změně parametrů zobrazení může změnit. Krajním případem by bylo takové slovo, které by bylo přítomno v každé kontextové jednotce a které by v grafu logicky muselo mít pozici zcela uprostřed, protože jeho vzdálenost k ostatním slovům (tj. spoluvýskyt) by byla dána jen jejich relativní frekvencí v textu. V důsledku tohoto trendu se někdy při interpretaci využívá rozdělení konfigurace na výseče mající společný střed, poněvadž vzdálenost slov od středu vytvořeného nejfrekventovanějšími slovy vypovídá zejména o jejich relativní frekvenci v korpusu, nikoli o jejich specifické distribuci (viz Obr. 2).
- b) Interpretace je pochopitelně ovlivněna významem jednotlivých slov, nicméně v grafu nelze nahradit slova tečkami nebo čísly za účelem dosažení objektivnější interpretace nezávislé na našem jazykovém porozumění. I tak by bylo sice možné odhalit izolované shluky, ale nikoli například polaritu pole, kdy na jedné straně grafu se nacházejí slova vážící se například k politickému diskurzu a na druhé straně ke slovům odkazujícím na každodenní život. Proto, i když musíme vzít při interpretaci v úvahu

význam sledovaných slov, nepojímáme vzdálenost prvků v textovém prostoru jako primárně sémantickou, ale jako vzdálenost diskurzí.

- c) Konečně je třeba uvést, že věcná interpretace konfigurací slov v korpusu zobrazených prostřednictvím MDS nemá žádnou „technickou“ oporu nezávislou na studovaném problému [Shepard 1980]. V tomto smyslu nemůže být ani „objektivní“ a závisí na schopnosti výzkumníka vynalézt smysluplné vysvětlení získané konfigurace slov. Velkou výhodou mají v tomto směru srovnávací analýzy několika korpusů, poněvadž se při interpretaci můžeme opírat o diskuzi jejich specifických i sdílených rysů.

Vraťme se k mému ilustrativnímu příkladu biografických rozhovorů. Na základě vytvořených frekvenčních slovníků jsem pro každý korpus vybral padesát nejfrekventovanějších slov (včetně jejich forem a vyskytujících se synonym) a z nich sestavil analytické slovníky. S využitím programu COOA jsem spočítal vzájemnou vzdálenost mezi slovy v kontextové jednotce pěti vět a matici spoluvýskytů normalizoval Jaccardovým koeficientem podobnosti. Matici jsem použil jako vstupní data pro nemetrické mnohorozměrné škálování v programu SPSS. Pro výpočet dvojrozměrné konfigurace jsem aplikoval proceduru PROXSCAL pracující s maticí podobností. Ostatní parametry byly standardní, tak, jak je modul nabízí; počet iterací byl 1000 náhodných startů a proceduru jsem několikrát (obvykle 5–10 ×) opakoval za účelem získání konfigurace s nejnižším stresem. Výsledný graf pro korpus biografických rozhovorů žen dělnic je na obrázku 1, graf mužů příslušníků inteligence je na obrázku 2.

Jak interpretovat graf spoluvýskytů vybraných frekventovaných plnovýznamových slov ve vyprávěních žen dělnic (Obr. 1)? Nejdříve zkusíme najít střed grafu, který by měl být v okolí nejfrekventovanějších slov. To jsou v tomto případě slova *dítě* (1087 ×), *škola* (734 ×) a *práce* (716 ×). Pozice těchto slov je opravdu ve středu grafu, tzn. konfigurace není excentrická. Dále nás zajímá hustota slov, zda lze identifikovat oblasti nahuštění, které by mohly indikovat nějakou diskurzí strukturu, například výrazné téma, které by bylo pro tyto vyprávěčky charakteristické. Z grafu je patrné, že největší hustota je v jeho středu, který tvoří jakési jádro a okolo něj je prstenec ostatních slov, jež jsou více či méně rovnoměrně rozptýlena. Tato struktura je vyznačena elipsami. „Naslepo“, bez znalosti slov, které tuto diskurzí strukturu tvoří, bychom ji mohli interpretovat tak, že vyprávění žen-dělnic má určitou osu procházející napříč jednotlivými vyprávěními a tvořenou 10–15 vysoce frekventovanými slovy. Na tuto osu jsou navěšena témata, která jsou vzhledem k prázdnému prostoru oddělujícímu jádro od prstence slov kolem něj diskurzí kvalitativně odlišná, tzn. ve vyprávěních hrají sekundární roli.



a vytváří jen úzkou výseč. Tento rys výrazně odlišuje tuto skupinu vypravěčů od skupiny předchozí, kde slova vztahující se k základním biografickým epizodám domova, dětství, vzdělání, rodiny a zaměstnání tvořila osu vyprávění. V biografických vyprávěních mužů příslušníků inteligence taková osa přítomna není a vyprávění přecházejí od školního vzdělávání k práci ovlivňované politickými okolnostmi a sféra domova a rodiny zaujímá jen relativně malé místo, které s politikou a zaměstnáním úzce nesouvisí.

Co z této interpretace grafů spoluvýskytů nejčastějších významových slov vyplývá pro výzkumnou otázku – zda a případně jak rozlišují pamětníci ve svých životopisných vyprávěních sféru soukromého a veřejného života a zda se u různých sociálních skupin vypravěčů tyto sféry liší? Pochopitelně zde prezentovaná analýza byla ryze ilustrativní a na jejím základě můžeme činit jen hypotetické závěry. Hlavním zjištěním je, že ve vyprávěních není veřejné a soukromé – ať již je definujeme jakkoliv – diskurzivně výrazným polarizujícím prvkem. Diskurzivní sféry veřejněpolitického a privátně-rodinného jsou v diskurzu propojeny prostřednictvím líčení školního vzdělávání a pracovního uplatnění. Zvláště slova vztahující se k zaměstnání (*dílna, vedoucí, pracovní, provoz, ředitel* apod.) se jak u dělnic, tak u mužů příslušníků inteligence nachází v blízkosti slov vztahujících se ke komunistické straně a jejím organizačním jednotkám a členům. Poukazuje to na skutečnost, že často zmiňovaný útěk do privátního světa v 80. letech minulého století nevytvořil výraznou oblast soukromí, která by dala v pamětnických vyprávěních vznik rozeznatelné diskurzivní sféře postavené na roveň líčení vzdělávání nebo zaměstnání, ve kterých se privátní a politické mísí. Pro bližší zjištění povahy dichotomie soukromého a veřejného bychom museli analyzovat jednotlivé životní sféry a v jejich rámci sledovat distinktivní znaky této dichotomie.

### 3.4.3 Shrnutí metody sledování spoluvýskytů slov

Mám-li rekapitulovat hlavní výhody a omezení představené metody, její silnou stránkou je schopnost zpracovat velké množství textu bez nutnosti zapojení lidských kodérů a přitom zachytit relevantní aspekty dat, tj. přirozeně se vyskytující diskurzivní struktury. Její hlavní slabinou je silná závislost na frekvenčním rozměru dat, tzn. jevy a struktury, které se neprojevují frekvenčně zřetelnými rysy, nejsou touto metodou zachytitelné. Tato slabina ovšem není tak významná, jak by se mohlo na první pohled zdát. V úvodní pasáži jsem diskutoval častý problém textových analýz spojený s kódovacími strategiemi. Lidští kodéři jsou prostřednictvím svého porozumění textu schopni zachytit mnohem více relevantních aspektů textu, zvláště ve vztahu k jeho významo-

vému a strategickému použití. Na druhou stranu mají opodstatněný sklon nalézat více možných způsobů čtení textu a tím „komplikovat“ rychlé kódování většího množství textu. Metoda sledování spoluvýskytů slov čte text tak, že zviditelňuje frekventované struktury spojené s použitým lexikem (slovy). Ovšem vzhledem k tomu, že pracuje s velkým množstvím textů zpravidla podobného charakteru, relevantní struktury by měly být přítomné ve velké části textů a tak i kvantitativně postižitelné. Ve zde uvedeném ilustrativním příkladu ukázala rozdílná distribuce spoluvýskytů frekventovaných slov na přítomnost diskurzivní osy v biograficky orientovaných rozhovorech žen dělnic, která zahrnovala hlavní životní etapy, zatímco v případě mužů příslušníků inteligence taková struktura přítomna nebyla a rozhovory se vyznačovaly spíše biograficky relevantními, ale přeci jen více tematickými diskurzivními oblastmi. Z tohoto hlediska by se metoda neměla používat na textová data, která jsou buď příliš heterogenní, že nelze předpokládat opakující se diskurzivní struktury (pokud však nechceme analyzovat právě důsledky této heterogenity, jako např. v [Hájek et al. 2006a], nebo na data sestávající z několika málo souvislých textů, kde opakující se lexikálně postižitelné diskurzivní struktury nebudou kvantitativně výrazné (např. porovnání dvou biografii).

## 3.5 Souhrn obsahové analýzy

V rámci této kapitoly jsem se věnoval jen několika možnostem, které obsahová analýza jako technika zkoumání textů nabízí. Připomenuli jsme definici z počátku kapitoly – obsahová analýza je výzkumná technika pro vytváření replikovatelných a validních úsudků z textů na okolnosti vytváření jejich obsahů – důležité bylo upozornit na povahu obsahu, který je obsahovou analýzou zkoumán. Obsahová analýza je totiž také určitá forma „čtení“ a proto obsah textu spoluvytváří, nikoli jen odhaluje.

Klíčovou procedurou obsahové analýzy je kódování. Lidské čtení může poskytnout základ pro adekvátní odborné kódování, protože člověk čtenému textu rozumí v jeho kotextu i kontextu, čte kooperativně. To ovšem může být i slabina, protože – jak jsem ukázal v příkladu analýzy čtení úryvků vyprávění – čtení produkuje množství významů. Z toho plyne metodologický problém jak naučit kodéry číst neznámý text stejně. Kvalitativní obsahová analýza využívající postupu zakotvené teorie se snaží tento problém vyřešit prostřednictvím několika úrovní kódování, zejména kombinací otevřeného a axiálního kódování. První kódování ukazuje na všechny možnosti čtení v daném kontextu, které analytika napadnou; druhý typ kódování hledá mezi záplavou možností ty nejvýznamnější vztahy, jež by měly dávat dohromady

smysluplný celek (obsah). Obsahová analýza formou lidského čtení je velmi časově i kvalifikačně náročná a hodí se tehdy, máme-li k dispozici jen omezené množství textů.

Pro velké textové korpusy je možné lidského čtenáře vycvičit tak, že bude číst operativně, tzn. víceméně jako stroj. Nicméně i tak je kódování a koordinace kódů velice náročná a reliabilita zjištění není zaručena. Proto jsem ukázal techniku, která je strojovým čtením, nicméně která generuje lidsky srozumitelný obsah analyzovaných textů. Výsledkem techniky sledování spoluvýskytů slov jsou grafické konfigurace vzájemných vzdáleností klíčových slov v textu, jež lze interpretovat jako netriviální reprezentaci obsahu textu. Slova v grafickém zobrazení vymezují – i když jen nedokonale – určitá témata nebo diskurzí struktury. Tím je generován obsah textů a skrze něj lze činit úsudky o okolnostech jejich vzniku. Na příkladu dichotomie soukromého a veřejného jsem se pokusil ukázat, jak vypravěči ve svých životních vyprávěních soukromé od veřejného výrazně rozdělovali, resp. nerozdělovali.

## Kapitola 4. Analýza diskurzu

Analýza diskurzu neboli diskurzí analýza<sup>8</sup> je definována jako analýza jazykového jednání, mluveného i psaného, které přesahuje hranice jednotlivých vět. Zaměřuje se zejména na to, jak je význam konstruován a interpretován prostřednictvím jazyka používaného v různých sociálních situacích a kontextech.

V poslední době existuje v českém akademickém prostředí určitá averze vůči slovu diskurz, která však pramení spíše ze špatné informovanosti než z vlastní problematiky diskurzu a jeho analýzy. Situace je zde podobná situaci na konci 80. let minulého století se sociální konstruktivismem, který byl považován za jistou formu degenerace sociologického uvažování, jež se točí v kruhu (zkoumá konstrukce konstrukcí konstrukcí, aby jeho výsledkem byla další konstrukce). Podobně i o analýze diskurzu bývá prohlašováno, že je degenerativní, že žádné nové vědomosti nepřináší, protože je ve výsledku autoreferenční (studujeme diskurzy o diskurzech mající původ v jiných diskurzech a výsledkem může být pochopitelně zase jen další diskurz). Zmnožování diskurzů podle kritiků nepřináší nové vědění, stejně jako zmnožování sociálních konstrukcí. Jak jsem poznamenal výše, jde o nedorozumění, které odjakživa provází všechny reflexivní a tudíž relativizující přístupy; tam, kde byl dříve jasný pořádek (realita a její reprezentace), je najednou zmatek – nelze jednoduše rozlišit, co je realita a co její reprezentace. Proto se v první části této kapitoly pokusím o konceptuální vyjasnění pojmu diskurz, aby se v následující části snížilo riziko nežádoucích nedorozumění.

<sup>8</sup> V české literatuře se lze poměrně často setkat i s termínem diskur(s)ivní analýza [např. Šmídová-Matoušová 2012]. Já se přidržuji termínu diskurzí jako označující skutečnosti týkající se diskurzu; diskurzí analýza, událost či praxe se vztahují k diskurzu, jsou tedy diskurzí. Termín diskurzivní na druhé straně označuje „vyjádřený promluvou“, jako například diskurzivní rozepře.