

# VYUŽITÍ DISKURZIVNÍ ANALÝZY V APLIKOVANÉM ANTROPOLOGICKO-SOCIOLOGICKÉM VÝZKUMU S POMOCÍ STATISTICKÝCH PROGRAMŮ HAMLET 2.0 A TEXTSTAT-2.9C

Barbora Nohlová – Pavel Mikolin

*Katedra sociologie, andragogiky a kulturní antropologie, Filozofická fakulta,  
Univerzita Palackého v Olomouci  
nohlova@sociofactor.eu, mikolin@sociofactor.eu*

**The use of discursive analysis in applied  
anthropo-sociological research with the help of  
the statistical programs Hamlet 2.0 and  
TextSTAT-2.9c**

*Abstract*—This contribution will provide information about the possibilities and limitations of the use of discursive analysis of textual data with the help of the statistical program Hamlet 2.0. It consists of two mutually interconnected parts. In the first part of the contribution, the theory of the method of discursive analysis will be discussed. The second part will be focused on the application of the statistical program Hamlet 2.0 in practice. This will be shown using the example of a partial discursive analysis of state administration and self-government relating to Islam in the Czech Republic, which was part of the project „Islam in CZ: The establishment of Muslims in Czech public spaces“. Readers will be apprised of all steps in the analysis using the program Hamlet 2.0, which include selection of the texts, their preparation for analysis, the data analysis itself and possibilities for the final interpretation of results. On the basis of this practical example, the advantages and disadvantages

of this program for applied anthropo-sociological research will also be presented.

*Keywords*—Discursive analysis, Hamlet 2.0, Islam in CZ, Muslims, state administration, self-government

## ÚVOD

NÁSLEDUJÍCÍ text se věnuje praktickému využití analýzy diskurzu s ohledem na možnosti využití programů TextSTAT-2.9c a Hamlet 2.0 ve výzkumné praxi. Ukazujeme v něm možnosti zpracování a vyhodnocení kvalitativních dat na konkrétním příkladu – analýze diskurzu státní správy a samosprávy, který se váže k etablování islámu v České republice. Přitom vycházíme z teorie analýzy diskurzu tak, jak ji nastínil Michel Foucault. Jeho přístupu se však věnujeme pouze okrajově, protože naším cílem je především poukázat na vyhodnocení zkušeností, které jsme nabyli při aplikaci zmíněné metody při využití programů

TextSTAT-2.9c a Hamlet 2.0 v rámci zpracování dat, která jsme získali prostřednictvím kvalitativních polo-strukturovaných rozhovorů.

Naše bádání se zaměřilo pouze na jednu dílčí aktivitu v rámci komplexního výzkumu,<sup>1</sup> který sleduje etablování islámu v České republice. Jak uvádí Topinka (2010), islám se v České republice etabloval především v poválečném období; pro polistopadové období jsou ve vztahu k rozvoji muslimské obce charakteristické dvě zásadní změny. Muslimové založili několik organizací, které byly časem zaregistrovány a získaly právní status, také se jim podařilo vybudovat institucionální, organizační a kulturní zázemí svých sdružení a organizací. Druhou významnou změnu, která vedla k dynamickému rozvoji muslimské obce, zapříčinila imigrace. Počet imigrantů od roku 1989 trvale narůstal a imigrace, jejíž součástí byli také muslimové, ovlivnila strukturu muslimské „komunity“ v České republice, ale i dění uvnitř ní. Přesto si islám zachovává v české společnosti charakter cizokrajnosti, povědomí Čechů o jeho realitách se ukázalo jako velmi malé. Veřejné mínění o muslimech se utváří pod vlivem médií, která produkují stereotypní obraz islámu.

Topinka (2010) si dále všímá formování různých diskurzů, které vznikly v čase, a řadí sem například specializované diskurzy (policejní, lázeňský, uprchlický apod.) nebo protiislámský diskurz, který ve druhé polistopadové dekádě produkuje výpovědi o islámu jako o fundamentalistickém protimoderním náboženství, které popírá liberální demokratické hodnoty a reprezentuje hrozbu pro společnost. Dále píše: „... největší diskurzivní střety se odehrály kolem výstavby mešit v Brně a Praze a kolem žádosti o přiznání oprávnění k výkonu zvláštních práv, kterou Ústředí muslimských obcí podalo v roce 2006. Oba tyto střety byly analogií k francouzské ‚šátkové aféře‘, protože v obou případech se jednalo o první významné pokusy o deprivatizaci náboženství.“ (Topinka 2010, 256)

Komplexní výzkum má za cíl zjistit podoby a procesy týkající se etablování islámu ve třech rovinách: státní správy a samosprávy (institucí), občanské společnosti s důrazem na veřejné mínění a imigrantské „komunity“. Naším úkolem bylo metodou analýzy diskurzu vyhodnotit data, která jsme získali prostřednictvím rozhovorů, které sledovaly, jak je islám vnímán a interpretován na úrovni státní správy a samosprávy, a jaké obsahy zkušeností se na této rovině utvářejí. Získaná data byla primárně vyhodnocena metodou zakotvené teorie,<sup>2</sup> naším úkolem ale bylo aplikovat jiný přístup – diskurzivní analýzu a nahlédnout na tatáž data z odlišné výzkumné perspektivy. Konkrétně se jednalo o nalezení objektů, typů výpovědí, pojmů a témat, která se objevují ve výpovědích sociálních aktérů, kteří patří do veřejné správy. Naším úkolem bylo tedy popsat, jak diskurz formuje výpovědi, jaké vztahy nastoluje mezi jeho subjekty a jaké důsledky to může mít. Jak je zřejmé, projekt využívá několik metod výzkumu a analýza diskurzu je pouze jednou z nich.

V souvislosti s touto metodou jsme si stanovili několik výzkumných otázek. Níže uvádíme nejdůležitější z nich:

1. Jak je islám vnímán a interpretován na úrovni státní správy a samosprávy?
2. Jaké obsahy zkušeností se na této rovině utvářejí?
3. K jakým jiným diskurzům je odkazováno, když sociální aktéři hovoří o islámu / muslimské komunitě / jednotlivých muslimech?

#### VÝZKUMY SOUVISEJÍCÍ S ISLÁMEM A MUSLIMY V ČR

Vnímání muslimů v dnešní době překračuje sociální rozměr a přerostlo do roviny etnicity. Například v kontextu nezaměstnanosti a nedostatku pracovních pozic v českém prostoru jsou muslimové zmiňováni jen

výjimečně. Základní a nejdůležitější charakteristikou, která je přisuzována muslimských přistěhovalcům, je náboženská orientace, která se také ukazuje jako nejvíce stigmatizovaný fakt. Výše jsme zmínili, že na formování takového názoru ve většinové společnosti mohou mít vliv média (Topinka 2010). Výzkumný projekt Islám v ČR: etablování muslimů ve veřejném prostoru navazuje na výzkum Integrační proces muslimů v ČR – pilotní projekt 2006/2007, avšak vzhledem k námi zvolené metodologii jsme věnovali pozornost také dalším výzkumům provedeným v ČR. V rámci naší analýzy jsme pracovali především s mediálními analýzami zabývajícími se prezentací islámu a muslimů v ČR, protože ty zaměřují svoji pozornost na procesy vzniku a formování názorů příjemců na konkrétní tematiku.

Na xenofobní medializaci se zaměřují monitorovací centra Evropské unie, častěji se však objevují výzkumy zaměřené na konkrétní menšiny v jednotlivých zemích, kdy je jen málokdy přihlédnuto k mediálnímu vlivu. S tím, jak v Evropě narůstá počet muslimů, objevuje se i více mediálních analýz, jež se jim věnují. V České republice se toto téma otevírá jen velmi pozvolna, pozornost médií je častěji věnována početnějším a viditelnějším menšinám, např. Romům a Ukrajincům, se kterými má majorita více zkušeností.

Častěji je možné narazit na práce, které vznikají v multikulturních a integračních centrech. Z posledních roků je tedy třeba zmínit například *Jak se (ne)píše o cizincích v českém tisku* (Klváčová a Bitrich 2003), nebo *Obraz cizinců v médiích* (Kaderka a Karhanová 2001). Tématu mediálního zobrazování se věnuje Martina Křížová, která v roce 2006 vypracovala pro multikulturní centrum analýzu *Neviditelná menšina* (Křížková 2007), která se věnuje výhradně muslimům. Podobnou práci napsala téhož roku i Eva Burgetová (Burgetová 2006). Před dvěma lety také vyšla knižně baka-lářská práce Tomáše Melichárka, která mapuje mediální diskurz týkající se muslimů

v Brně. Všechny tyto práce se věnují islámu a muslimům v ČR. Jak jsme zmínili již výše, zajímalo nás v naší analýze především to, jak je islám vnímán a interpretován na úrovni státní správy a samosprávy, jaké obsahy zkušeností se na této rovině utvářejí a k jakým dalším diskurzům zástupci státní správy a samosprávy odkazují, když hovoří o islámu, muslimské komunitě nebo jednotlivých muslimech. Výše uvedeným výzkumům jsme věnovali pozornost především proto, že svým zaměřením nejvíce odpovídali jak našemu tématu, tak i zvolené metodologii.

#### DISKURZ A DISKURZIVNÍ ANALÝZA

Nejdříve je nutné objasnit, co chápeme jako diskurz. Pokud se budeme zabývat slovníkovými definicemi, zjistíme, že v angličtině a francouzštině (*discourse, discours*) označuje diskurz rozhovor, diskuzi nebo rozpravu nad nějakým tématem (srov. *Discourse (Oxford Dictionaries)* 2015).

V rámci lingvistiky však pojem diskurz původně označoval jazykové útvary nad úrovní věty. Takové pojetí diskurzu je však velmi obecné, níže tedy uvádíme to pojetí diskurzu, s nímž jsme v našem výzkumu pracovali my.

Diskurz se vztahuje k jistému způsobu užívání verbálních a vizuálních kódů, a to při produkci či posilování určitých významů, které se vztahují k nějakému předmětu. Určuje tak přístupy k danému tématu. Diskurz je také souborem možných způsobů, jimiž je možné nakládat s významy v konkrétní oblasti, vytváří hodnotová poselství (Burton a Jiráček 2003, 304).

Podle Phillipse a Jørgensenové (2002, 1) je diskurz specifickým způsobem hovoření o světě a porozumění jemu či jeho aspektům. Dále formulují čtyři hlavní aspekty, jež je možné aplikovat na diskurzivní analýzu. Jedná se o (Phillips a Jørgensen 2002, 5–6):

1. kritický přístup k běžně rozšířenému vědění;

2. kulturní a historickou specifičnost;
3. vztah sociálních procesů a vědění;
4. vztah mezi sociálním chováním a věděním.

O diskurzivní analýze také hovoří jako o metodě, která je zároveň svojí vlastní teorií; v souvislosti s tím podle nich zahrnuje další čtyři domény, těmi jsou: filozofické premisy, teoretické modely, metodologické postupy a specifické postupy analýzy (Phillips a Jørgensen 2002, 4).

Podle Cholta (2006, 18–19) je diskurz vyjádřením jistého způsobu užívání komunikačních prostředků, a to jak verbálních, textových a vizuálních, tak i jiných, při kterých dochází k produkci sdělení (významů), jež se vztahují k nějakému tématu. Diskurz tedy znamená to, jak je přistupováno k určitému tématu, jakým způsobem se o tématu hovoří, a jak se nakládá se souborem možných významů a kategorií, které se k takovému tématu vážou. Tyto významy a kategorie jsou však sociálně podmíněné, proměnlivé v čase a jsou veskrze konstruované. Definiční rysem diskurzu je pak jisté makrotéma. Cholt (2006, 19) vnímá diskurz jako pravidelný a systematický proces. Definiuje tento proces jako časově rozptýlený výskyt událostí a výpovědí, jež jsou vázány k danému tématu, a které se střetávají v diskurzivních polích. Výsledný celek nemusí být nutně konzistentní, avšak je v něm možné identifikovat jisté myšlenkové vzorce, typy reprezentace a argumentační strategie, vytváření identit a pozicování aktérů, stejně jako ideologičnost. Cílem diskurzivní analýzy je na základě analýzy dat identifikovat a kategorizovat výpovědi. Avšak to není jediný cíl, diskurzivní analýza se také snaží o odkrytí procesu jejich utváření. Výsledkem pak tedy není pouze soubor znaků – typologie, které odkazují k obsahům nebo reprezentacím takových kategorií, ale taktéž odkrytí praktik a podmínek, které tyto kategorie a jejich objekty a subjekty systematicky vytvářejí. Výstupem by měl nakonec být text přístupný a srozumitelný. Jeho obsah by měli

být schopni interpretačně sdílet i ti, jichž se přímo týká (např. minority). Podle Cholta (2006, 19) lidé pomocí jazyka vyjadřují své názory, čímž konstitutivně vytvářejí reprezentaci světa.

Tato definice se velmi blíží pojetí diskurzu Michela Foucaulta, jehož práce byla také naším teoretickým východiskem. V našem výzkumu tedy pracujeme zejména s koncepcí diskurzu, kterou Foucault představil ve své práci *Archeologie vědění*. Foucault svým pojetím diskurzu klade na čtenáře velmi vysoké nároky, proto bylo nutné pracovat i s jinými texty zabývajícími se diskurzem, ty jsme již velmi zjednodušeně představili v předchozích odstavcích.

Nyní se pojďme zaměřit přímo na Foucaultovo vymezení diskurzu. My jej uvádíme pro lepší pochopení v simplifikované formě, aby bylo patrné, z jaké pozice jsme k výzkumnému tématu přistupovali. Cholt (2006, 22–23) charakterizuje Foucaultův přístup jako bezpochyby historický, zaměřený tedy na historická data. Pro Foucaulta je diskurz systémem pravidel, která vymezují prostor, který zde existuje pro tvorbu aktuálních výpovědí. Pravidla strukturace výpovědí určují diskurz zevnitř. Diskurz také určuje vztah k jiným diskurzům, a tak vyznačuje své vlastní hranice. V centru pozornosti je samotný „prostor“, v rámci něhož výpovědi vznikají, nejedná se tedy o aktuální diskurz jako o proces, ale o tvorbu výpovědí. Při provádění analýzy, jež má za cíl popis určitého diskurzu, je nutné vycházet z konkrétních výpovědí, z nichž každá byla vyprodukována lidmi za určitých podmínek. Centrem Foucaultova zájmu je taková realita pravidel, tedy řád diskurzu, jež dává vzniknout konkrétním výpovědím (Foucault 1976).

Foucault (2002, 44–45) vidí místo příčin a důsledků nahodilosti. Jednotlivé historické konfigurace jsou pro něj pouze jedněmi z možných, avšak zároveň není ani jedna lepší než druhá (zdržuje se jejich hodnocení). Foucault je také skeptický vůči poznání. Snaží se vyvarovat hodnoticích soudů.

Foucaultův koncept diskurzu se opírá o archeologický přístup, který se zaměřuje na „časové snímky“ historie, a genealogický přístup, který klade důraz na procesualní stránku vědění/moc.

Diskurz vnímá jako produktivní. Foucault neuznává, že by diskurz byl „uvnitř“ před jazykem, tedy před tím, než je verbalizován slovy. Zároveň se podle něj ani objekty nenacházejí „vně“ diskurzu, v nějaké vnější sféře. Foucault se zaměřuje na „prostor“ diskurzu mezi „vně“ a „uvnitř“ a distancuje se od toho, že by diskurz byl utvářen pouze slovy, nejedná se podle něj o čistě lingvistický fenomén. Při zkoumání diskurzivního pole bychom se neměli zaměřit pouze na jeden konkrétní text, ale je nutné uvědomit si, že takový text má vztah k dalším textům, neboť existuje v síti textů a vztahů. Koncept ne-diskurzivního poskytuje částečně odpověď na problém hranic diskurzu. Diskurzivní pole je tvořeno vymezeným prostorem, je určováno souhrnem všech uskutečněných vyslovených či napsaných výpovědí. V tomto smyslu je tedy ohraničeno, je tvořeno konečným a omezeným prostorem výpovědí; je však také nutné vzít v úvahu, že výpovědi v čase přibývají, diskurzivní pole tedy není statické. Diskurz tak není nikdy úplný, není definitivně ukončen. Například fyzické tělo jako takové diskurzivní není, avšak tělo nemůže fungovat v ne-diskurzivním prostředí. Pojem „tělo“ je výsledkem diskurzu a stejně tak jsou i přírodní fenomény post hoc ospravedlněním ustanovených faktů. Podle Foucaulta (2002) není důležitý vztah výpovědi ke skutečnosti, tedy faktům. Je nutné zaměřit se nikoliv na jejich pravdivost, ale na jejich vzájemnou podmíněnost a koexistenci.

Jak již bylo zmíněno, Foucault vyzývá k tomu, abychom se při zkoumání diskurzivního pole nezaměřili čistě na jednu konkrétní výpověď. Výpovědi jsou v prostoru rozmístěny různým způsobem; naším cílem má být nalezení pravidelností mezi nimi, tedy objevení takzvaných diskurzivních formací.

Klíčová jsou také pravidla, podle kterých dochází k formování diskurzu. Těmi Foucault rozumí podmínky, za nichž jsou výpovědi v diskurzivním poli rozmísťovány. Na prvním místě Foucault hovoří o objektech diskurzů, měli bychom tedy sledovat, jaké se v diskurzivním poli objevují objekty. Na druhém místě pak Foucault uvádí typy výpovědí, ty jsou organizovány a propojeny prostřednictvím pojmů. Nakonec hovoří o tématech, teoriích nebo koncepcích. Diskurz, který je, jak již bylo zmíněno výše, uzavřeným celkem výpovědí, se vztahuje k jiným diskurzům. Buď je vymezen analogicky, tedy je v něm možné nalézt shodu některých předmětů, nebo je vymezen komplementárně či opozičně. Konečně, Foucault rozlišuje tři „prostory“, které obklopují výpověď (Foucault 2002, 134–162):

1. kolaterální prostor – je vymezen dalšími výpověďmi spadajícími do stejné množiny;
2. korelativní prostor – je vztahem výpovědi vůči objektům a subjektům;
3. komplementární prostor – představuje nediskurzivní formace.

Diskurzivní analýza pak zkoumá reprezentace a interpretace světa, odhaluje také, které výpovědi (promluvy a texty) jsou v diskurzu produkovány. Dále si klade za cíl zjistit, které výpovědi (promluvy, texty) jsou produkovány v konkrétním diskurzu, s jakými pojmy je v jeho rámci operováno a do jakých vzájemných vztahů jsou ve výpovědním prostoru uváděny, dává tedy konkrétní promluvy a texty do širších souvislostí.

#### ANALÝZA DISKURZU STÁTNÍ SPRÁVY A SAMOSPRÁVY

Výše jsme stručně přiblížili Foucaultovo pojetí diskurzu, které je pro náš výzkum s těžejší. Při diskurzivní analýze jsme zvolili nepříliš známé statistické programy TextSTAT-2.9c a Hamlet 2.0, jež nám byly doporučeny kolegy, kteří s těmito programy

již dříve pracovali (srov. *Analýza mediálních sdělení o přírodních vědách: Monitoring médií* 2007). V nich jsme tedy provedli kvantitativní analýzu textů.

Kvantitativní analýza textů dokáže odhalit formativní účinky diskurzu na uspořádání výpovědí, jejichž projev je možné vidět ve specifickém rozložení slov a slovních skupin ve zkoumaných textových dokumentech. Hamlet 2.0 je statistický program pracující s textovými soubory, ve kterých počítá slova obsažená v analyzovaném textu a vytváří z nich pak slovník. Po vypočítání frekvence slov vytvoří program matice. Matice jsou složeny z analytických jednotek, každá jednotka je tvořena shlukem slov, které dělí, nebo naopak spojuje vzdálenost mezi jednotlivými slovy. Hamlet 2.0 je schopný tuto vzdálenost vypočítat, čímž vzniká určitý význam, tedy kontext. Funkce multidimenzionálního „škálování“ (multidimensional scaling) je pak z matice schopná vytvořit graf, který zobrazuje jednotlivé shluky slov. Interpretace grafu, tedy odhalení významu shluků slov, je hlavním výstupem analýzy; graf však zobrazuje, jaký konkrétní diskurz je vytvářen.

Text, který jsme v programu Hamlet 2.0 analyzovali, tvořil soubor získaných rozhovorů. Pro jejich sběr jsme zvolili několik oblastí státní správy a samosprávy, u kterých bylo vysoce pravděpodobné, že jejich individuální aktéři budou přicházet do kontaktu s muslimy.<sup>3</sup> Předmětem naší analýzy pak byla data získaná prostřednictvím rozhovorů vedených se zástupci státní správy a samosprávy v období od 1. 1. 2014 do 30. 4. 2014. Celkem jsme analyzovali 60 kvalitativních polo-strukturovaných rozhovorů, z nichž 22 pokrývalo oblast samosprávy a 32 oblast státní správy. Zbýlých šest rozhovorů bylo provedeno se zástupci neziskového sektoru. Tyto rozhovory tvořily základ pro analýzu diskurzu státní správy a samosprávy.<sup>4</sup>

Prvním krokem analýzy byla úprava rozhovorů. Ty byly sloučeny v ucelený textový soubor, tzv. korpus. Všechny rozho-

vory jsme překopírovali do jediného souboru formátu DOC, a následně je manuálně upravili a odstranili jsme ty části, které byly pro analýzu irelevantní. Máme zde na mysli především interpunkční znaménka, vstupy tazatele do rozhovoru, nebo takové části textu, které nekorespondovaly s předmětem našeho zájmu. Výsledný očištěný korpus jsme pak vložili do programu TextSTAT-2.9c. Takto očištěný korpus čítal 765 530 znaků, včetně mezer.

Program TextSTAT-2.9c je statistický program sloužící k analýze textů. Pracuje s textovými soubory různého kódování a se soubory formátu HTML. Finálním výstupem programu je frekvenční slovník a konkordance, vytvořené z textových souborů vložených do TextSTAT-2.9c. Možnosti programu zde však nekončí, k jeho silné stránce také patří již zmíněná konkordance. Tato funkce umožňuje náhled slova vybraného z frekvenčního slovníku přímo ve větách, v nichž bylo použito. Užitečná je taktéž možnost následného náhledu těchto vět v textu. Další praktické využití programu TextSTAT-2.9c pro nás především spočívalo v jeho schopnosti pracovat s českým textem. Jazykové rozmezí tohoto programu totiž obsahuje i češtinu, nečinilo mu tedy velké obtíže přečíst znaky používané v tomto jazyce.

Jak jsme zjistili, mezi přednosti programu patří také přehlednost a intuitivní ovládání – s jeho pomocí je například možné provést dodatečnou úpravu textového korpusu před tím, než byl vložen do programu Hamlet 2.0 za účelem provedení finální fáze analýzy.

Hamlet 2.0, jak již bylo zmíněno výše, je rovněž statistický program sloužící k počítačové analýze textů. Primárním účelem tohoto programu je procházet textové soubory za účelem vytvoření slov či kategorií slov v daném slovníku, a počítání jejich společných frekvencí na základě specifických kontextuálních jednotek nebo na základě vět nebo jako kolokace na základě daného rozpětí slov. Přínos měření empirických vlastností textu je zde užitečně zkom-

binován s možnostmi grafické vizualizace v tomto programu.

Výše zmíněný frekvenční slovník programu TextSTAT-2.9c zobrazuje všechna slova a znaky použité v textu, bylo v něm tedy možné dohledat takové znaky, které byly s ohledem na předmět naší analýzy irelevantní a unikly naší pozornosti při provádění primární úpravy textu. Také funkce konkordance posloužila ve chvíli, kdy jsme již v programu Hamlet 2.0 určovali, která z vybraných slov frekvenčního slovníku vstoupí do finální části analýzy tak, aby mohla být zahrnuta ve výsledném grafu. Zobrazení slova ve větě přímo v textu umožňovalo důkladné zvážení jeho vazby vůči zkoumanému tématu.

Hamlet 2.0 a TextSTAT-2.9c nejsou jedinými dostupnými softwary. Dalším programem, který lze využít pro analýzu korpusu, je například Antconc, jehož výhodou je, že je bezplatný.

Dále lze jmenovat programy MAXQD, jehož uživatelské prostředí přeložil do češtiny Ladislav Toušek z Katedry antropologie FF ZČU, nebo Voyeur Tools, který není nutné stahovat, protože se jedná o online software. Za zmínku také stojí program HK Coder, který původně sloužil pouze k analýze textů v japonštině, ale v současnosti už je vhodný pro širokou škálu jazyků.

Nabízí se však otázka, proč byly pro analýzu zvoleny právě programy Hamlet 2.0 a TextSTAT-2.9c. Kvůli časové náročnosti nebylo možné vyzkoušet všechny výše zmíněné programy, proto bylo třeba jejich výběr omezit. Kromě krátkého popisu a představení funkcí jsme o většině programů neměli více informací, rozhodli jsme se proto opřít se o doporučení kolegů, kteří měli již s programy Hamlet 2.0 a TextSTAT-2.9c zkušenosti.

Přestože je rozpětí funkcí programu TextSTAT-2.9c značné a považujeme jej za nesmírně nápomocný a praktický, bylo pro účely naší zvolené analýzy nutné doplnit ho dalším programem, v našem případě Hamletem 2.0. TextSTAT-2.9c totiž nedisponuje, na rozdíl od Hamleta 2.0,

pro nás klíčovými funkcemi; především postrádá funkci multidimenzionálního škálování, stejně jako není schopen vygenerovat graf, jehož vytvoření bylo pro kompletní provedení naší analýzy nezbytným krokem. TextSTAT-2.9c je tedy vynikajícím doplňujícím výzkumným nástrojem, ale jeho samostatné použití je pro tento typ analýzy nevhodné.

Po sekundární úpravě textu jsme hotový korpus převedli z formátu DOC do formátu TXT tak, aby bylo možné jeho vložení do statistického programu Hamlet 2.0.<sup>5</sup> Postup práce s textovým korpusem v Hamletu 2.0 je na začátku poměrně snadný, těžkosti nastávají až u zdánlivých drobností, které jsou však schopné provedení analýzy značně zkomplikovat. Takovým příkladem mohou být jazykové konfigurace. V českém prostředí existuje pouze anglický manuál, ve kterém však některé informace chybí. Program je také především určen pro práci s anglickým textem, informace týkající se jazykových konfigurací cizojazyčných textů jsou tak uvedeny až na jeho úplném konci. Pokud si tedy výzkumník důkladně nepročte celý manuál, první těžkosti mohou nastat právě v okamžiku sestavování frekvenčního slovníku.

Práce v Hamletu 2.0, jak již bylo zmíněno výše, začíná vložением upraveného korpusu do programu. Jakmile je toto provedeno, přichází na řadu vytvoření již několikrát zmíněného frekvenčního slovníku. Frekvenční slovník se poté pomocí příslušné funkce v Hamletu 2.0 sloučí s textovým korpusem, ze kterého jsme do slovníku vkládali námi vybraná slova. Jednalo se o takové pojmy, které jsme měli v úmyslu analyzovat a později vytvořit jejich grafickou podobu tak, aby bylo možné provést interpretaci výsledků.

Tato slova jsme vložili do slovníku tak, že jsme nejdříve vypsali jejich základní tvar (kořen slova) a teprve k němu jsme vpisovali příslušné, s tímto kořenem související pojmy. Zde začalo docházet k prvním problémům. Hamlet 2.0, jak jsme již uvedli, nedisponuje českým jazykovým rozhraním.

Při načtení frekvenčního slovníku se některá slova nezobrazila správně. Pro bližší představu, například slovo „veřejnost“ bylo přečteno pouze jako „ejnost“, zbytek slova zanikl jako fragment slova v podobě „ve“.

Program Hamlet 2.0 nabízí pro takovou situaci řešení, v tomto bodě je však, jak jsme zmínili již výše, třeba dbát zvýšené pozornosti. Na modelovém příkladu ukážeme, k jak závažným komplikacím může v tomto bodě práce s daty dojít.

Chybné zobrazení slov ve frekvenčním slovníku totiž vytváří značný problém. Nabízí se zde však možnost řešení, a to v pomoci si programem TextSTAT-2.9c tak, že si zobrazíme frekvenční slovník v něm, a z něj pak vypisujeme námi požadovaná slova do slovníku v Hamletovi 2.0, a kombinujeme je tak s těmi, která nám zobrazuje tento program (doporučujeme vypsát slova z obou frekvenčních slovníků). Tento postup má však jistá omezení. Jedním je ne zcela snadné korektní dohledání takových slov, která nejsou zobrazena ve slovníku Hamleta 2.0 v celku, neboť při něm může dojít ke značnému zkreslení výsledné analýzy (myslíme si například, že fragment již výše uvedeného slova „ejnost“ mohl být „veřejnost“, ale ve skutečnosti jsou do jeho četnosti započítána všechna slova, která byla podobným způsobem programem „roz-krojena“, a my tedy ve výsledku nevíme, zda je uváděné číslo relevantní, tudíž zda se slovo skutečně vyskytlo tolikrát, aby pro nás bylo důležité umístit ho do analýzy). Jakmile vidíme, že stejné slovo obsahuje v programu TextSTAT-2.9c jinou četnost, než jaká je uvedená u Hamleta 2.0, je jasné, že je někde chyba. Přemýšleli jsme, jak takovou situaci vyřešit, a nabídla se nám následující možnost, která se sice jevila zdlouhavá, ale ze začátku se zdálo, že náš problém brilantně vyřeší.

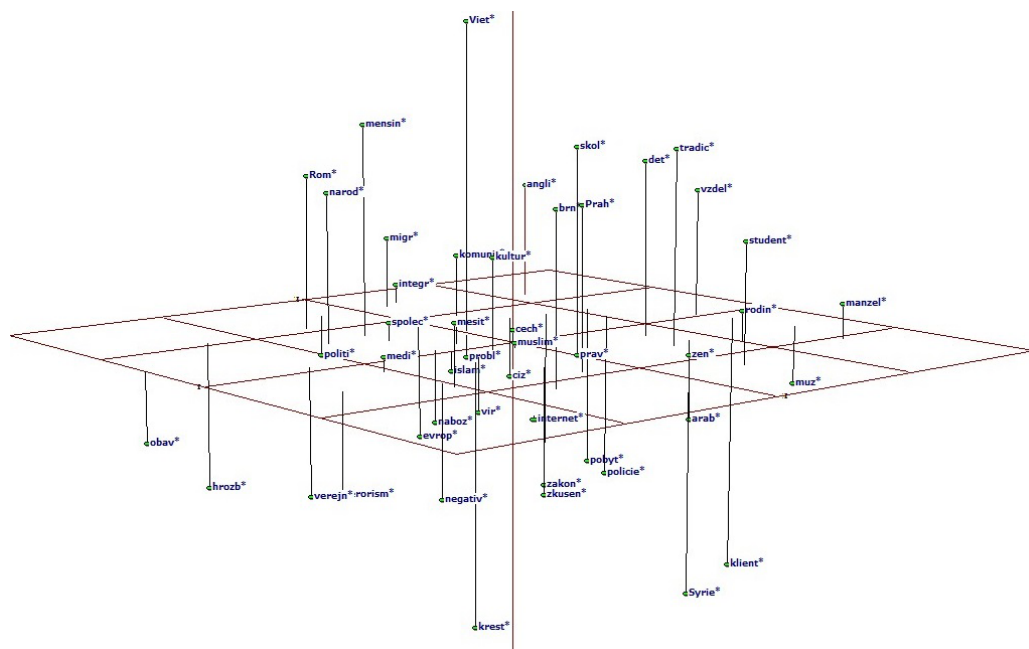
Náš další postup byl tedy následující. Přepsali jsme námi pro analýzu vybraná slova do takové podoby, aby neobsahovala žádné české spřežky. Vyšli jsme z předpokladu, že pokud vybraná slova upravíme tak, aby měla „anglický tvar“, tedy neob-

sahovala žádné, programu neznámé znaky, budou zobrazena správně, a naše nesnáz se vyřeší. Výhodou takového postupu byla taktéž dvojitá kontrola dat, jež jsme vkládali do analýzy, protože každé slovo prošlo několikanásobným zvážením. Nevýhodou však byla jeho časová náročnost, se kterou bylo nutné počítat; v našem postupu bylo totiž nutné všechny pojmy, které jsme měli v úmyslu použít k analýze, nejdříve vypsát z frekvenčního slovníku programu TextSTAT-2.9c do pracovního souboru ve formátu DOC, a poté je znovu manuálně přepsat v již existujícím textovém korpusu připravenému k analýze tak, aby neobsahovala spřežky. Alespoň pro částečné urychlení tohoto postupu jsme využili funkce „najít a nahradit“ v programu Word Office 2003. Teprve s takto upraveným korpusem jsme dále pracovali v programu Hamlet 2.0. podle již výše zmíněného postupu. Ten fungoval dobře, přinejmenším to zpočátku vypadalo a nebyť našeho „pokusu“ v závěru naší analýzy, o kterém ještě budeme hovořit dále, nebyli bychom nikdy přišli na to, že je něco špatně.

Vypsali jsme tedy všechna slova v jejich základním tvaru do námi vytvořeného slovníku v programu Hamlet 2.0 a přidali jsme k nim jejich synonyma. V okamžiku, kdy je v programu ručně sestaven slovník, je uživatel v půlce pracovního postupu. Je zde však třeba krátce zmínit, že mimo komplikace související s jazykovou konfigurací existuje mnoho jiných záležitostí, díky kterým se práce v programu Hamlet 2.0 může stát zbytečně komplikovanou.<sup>6</sup>

Vraťme se však nyní do fáze, kdy máme slovník úspěšně hotový. V tuto chvíli již před uživatelem leží pouze provedení samotné analýzy dat. Hamlet 2.0 pracuje na principu výpočtů vzdáleností slov v textu, která byla zvolena pro analýzu. Tím, že je takto zpracuje, umístí je na základě jejich vzájemné vzdálenosti a četnosti do konečného grafu, což umožňuje uživateli názorně vidět vztahy, které daná slova mezi sebou mají. Pro naši analýzu jsme zvolili 47 hesel v základním tvaru a celkový slovník syno-





Obrázek 1: Konfigurace slov ve výpovědním prostoru

nym čítal 927 slov v odvozených tvarech výchozích hesel.<sup>7</sup>

V okamžiku zadání pokynu k provedení analýzy začne Hamlet 2.0 zpracovávat námi zadaná slova ve slovníku spolu s textovým korpusem, ze kterého byl vytvořen frekvenční slovník. Doba zpracování informací záleží na výkonnosti počítače, délce zadaného korpusu a slovníku, pohybuje se tedy v rozmezí několika desítek minut až několika hodin.

Hamlet 2.0 v této fázi spočítá výskyty a spoluvýskyty slov ve všech jednotkách. Z četností spoluvýskytů vznikne symetrická matice, kterou je nutné normalizovat za pomoci Jaccardova koeficientu podobnosti (Everitt a Rabe-Hesketh 1997). Pro převedení matice vzdáleností do prostoru je nutné použít funkci mnohorozměrného škálování (MDS). Výsledkem je pak mapa slov, jež jsou v grafu zobrazena jako body. Ta slova, která se v textech k sobě vyskytují nejbližše (jedná se o jejich spoluvýskyt v jedné větě), se také v prostoru nacházejí ve vzájemné blízkosti. Naopak ta slova, která jsou v textech vzdálená, jsou v prostoru zobrazena jako vzdálená od sebe. Vý-

sledek je pak uveden v trojrozměrném prostoru. Níže přikládáme takové grafické znázornění k nahlédnutí.

Slova, jež se nacházejí blízko středu os, jsou pro analýzu ústřední, jedná se totiž o slova, která se v textech vyskytovala nejčastěji. Naopak slova, která jsou od něj vzdálená, takový význam nemají, jedná se o slova, která se nevyskytovala v textech tak často. Výhodou grafického znázornění v programu je také to, že je s grafem možno pohybovat v prostoru (otáčet s ním a posouvat ho), je tedy možné podívat se na pojmy (a především na jejich vzdálenosti) z mnoha úhlů, což napomáhá pochopení toho, proč se v prostoru nacházejí právě tam, kde se nacházejí.

Posledním krokem analýzy je pak interpretace grafu. Při ní je nutné pracovat s pravidlem, že slova vyskytující se v podobných kontextech mají příbuzný význam, proto je z grafu možné vyčíst základní strukturu pojmů, které budou interpretovány. Vzdálenost jednotlivých prvků v grafu vnímáme jako vzdálenost diskurzivní; při interpretaci je tedy nutné posoudit diskurzivní vzdálenost mezi slovy a také

jejich umístění v prostoru. Výzkumník pak hledá vysvětlení pro jednotlivé konfigurace ve výpovědním prostoru.

#### VÝSLEDKY ANALÝZY

Při interpretaci grafu je nutná podrobná a hluboká znalost textů, ze kterých analýza vychází, jinak nejsme schopni vztahům porozumět. Bez důkladné znalosti textů, v našem případě přepsaných rozhovorů, není výzkumník schopen plně pochopit významy, které se v grafu nacházejí.

Než zde představíme samotnou interpretaci grafu, rádi bychom vysvětlili, jakými pravidly se naše interpretace řídí. Při interpretaci jsme se snažili odhalit základní strukturu slov ve výpovědních prostorech. Všimli jsme si rozmístění slov v prostoru, a to z hlediska jejich polarity, rozptýlenosti/sevřenosti, oddělenosti/spojitosti a jádra/okraje. Dalším důležitým aspektem, na který bylo nutné brát ohledy, byla vzdálenost slova v grafu od jeho středu, protože slova, která se v textu vyskytují nejčastěji, se, jak již bylo zmíněno výše, nacházejí nejbližší středu, tedy oblasti průseku os. Tato slova si mohou, avšak nemusí, být sémanticky blízká, na druhou stranu vysoká četnost určitého slova v souboru výpovědí je indikátorem jeho významné role. Proto byla i tato zobrazení ve studii zohledňována. Interpretaci dále tedy ovlivnil význam jednotlivých slov a vzdálenost mezi slovy nebyla chápána jako sémantická, ale jako diskurzivní. Protože interpretace výpovědních prostorů zobrazených prostřednictvím MDS nemá žádnou technickou oporu nezávislou na studovaném problému, závisí interpretace konfigurací slov na schopnosti výzkumníka nalézt smysluplné vysvětlení těchto konfigurací (srov. Hájek, Kabele a Vojtíšková 2006, 279–280; *Analýza mediálních sdělení o přírodních vědách: Monitoring médií 2007*, 45–46).

Vzhledem k tomu, že je tento článek zaměřený především na popsání a zhodnocení práce se statistickými programy TextSTAT-2.9c a Hamlet 2.0, následující interpretace je pouze částí výsledků celkové analýzy, která ve svém plném rozsahu není zařazena; a to s ohledem na svůj značný rozsah a zaměření tohoto článku.

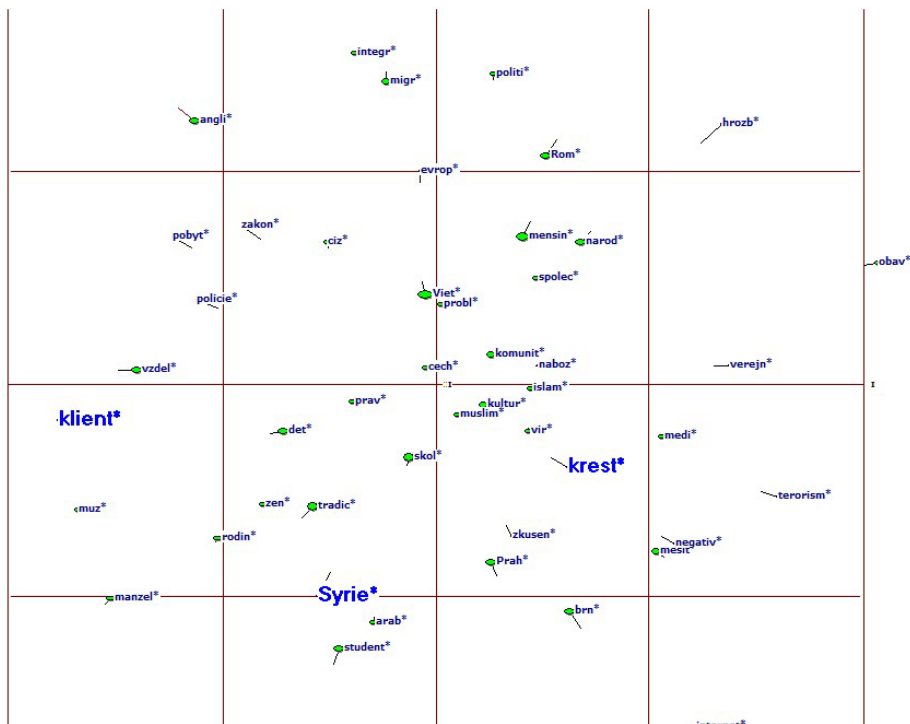
Jak již bylo zmíněno výše, je v programu Hamlet 2.0 možné využít více rozměrného zobrazení konfigurací slov ve výpovědním prostoru. V Hamletu 2.0 je možné s grafickým zobrazením prostorově manipulovat, rozmanitě graf natáčet a prohlížet si ho z nespočetného množství úhlů, což patří k další, nesporné výhodě programu.

V grafu jsou znázorněny základní tvary slov (např. Viet\*), avšak při tvorbě frekvenčního slovníku v programu Hamlet 2.0 byla k tomuto základnímu tvaru přiřazena i slova významově blízká (Viet\* – vietnamský, vietnamskou, Vietnamci apod.). Uvádíme zde interpretaci shluku ústředních pojmů, které se nacházejí nejbližší průseku os, a byly tedy nejfrekventovanější se vyskytujícími slovy v jednotlivých rozhovorech.

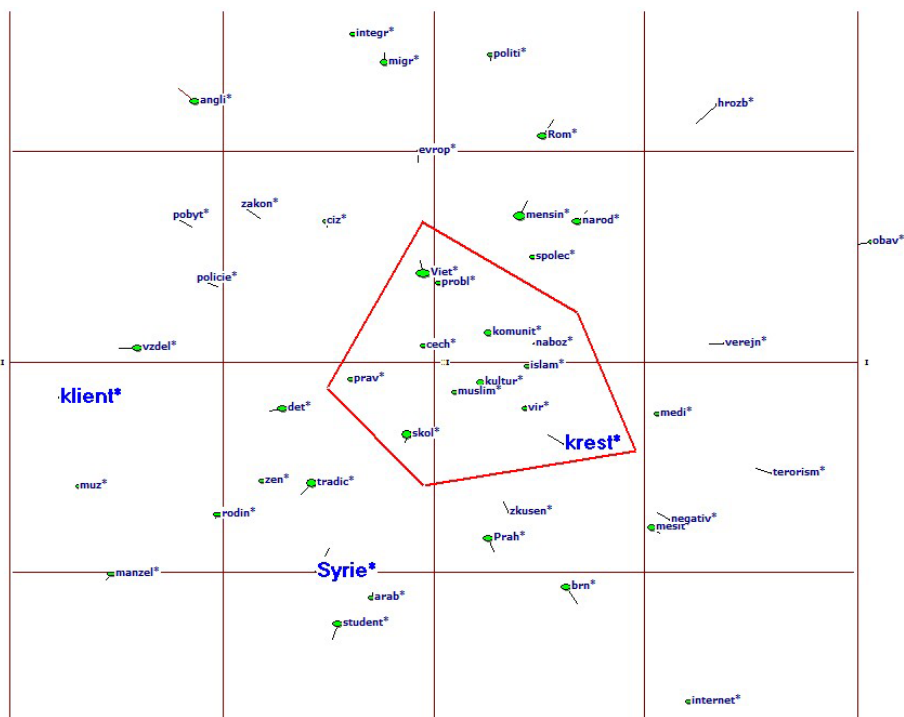
#### INTERPRETACE ÚSTŘEDNÍCH POJMŮ ANALÝZY STÁTNÍ SPRÁVY A SAMOSPRÁVY

Nyní přistupme k samotné interpretaci pojmů z grafu. Nejbližší průseku os se v našem grafickém zobrazení nacházely následující pojmy:

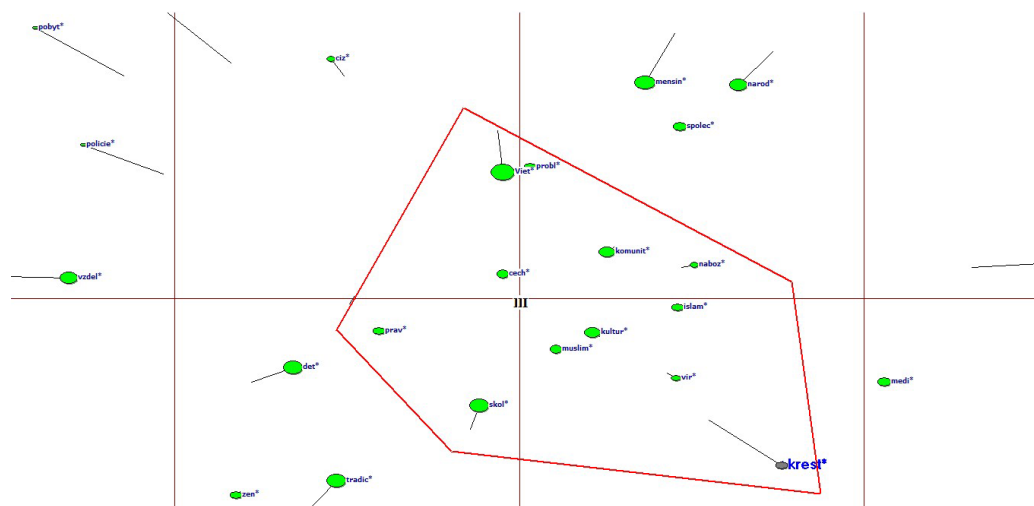
- ÚSTŘEDNÍ POJMY HLAVNÍ (nejbližší k průniku os) – cech\*, muslim\*, kultur\* – tyto pojmy je možné považovat za centrální pojmy analýzy
- ÚSTŘEDNÍ POJMY VZDÁLENĚJŠÍ – prav\*, skol\*, vir\*, islam\*, naboz\*, komunit\*, probl\*, Viet\*, krest\*
- JAKO CELEK – cech\*, muslim\*, kultur\*, prav\*, skol\*, vir\*, islam\*, naboz\*, komunit\*, probl\*, Viet\*, krest\*



Obrázek 2: Konfigurace slov ve výpovědním prostoru, pohled shora



Obrázek 3: Ústřední pojmy analýzy, pohled shora



Obrázek 4: Ústřední pojmy, detail

Pojmy jsme interpretovali jak samostatně, tedy ve smyslu toho, v jakých kontextech se tato konkrétní slova objevovala ve výpovědích našich participantů, tak v souvislosti s dalšími pojmy, v jejichž blízkosti se společně nacházela v rámci výpovědního prostoru. Samotnou interpretaci doplňujeme taktéž výpověďmi našich participantů.

Prvním, interpretovaným pojmem, bylo slovo *cech\**, vyskytující se nejbližže průsečků os. Jednalo se o nejméně frekventovanější se vyskytující pojem, jenž byl pro tuto analýzu ústřední.

Tento pojem se vztahoval především k českému prostředí, do kterého je celá analyzovaná problematika zasazená. V souvislosti s tímto slovem je však viditelná ještě jedna rovina. Pojem *cech\** udává soubor norem a hodnot vyznávaných na území ČR jejími obyvateli. Je zde tedy možné vidět jisté srovnání s tím, jak česká veřejnost vnímá a hodnotí islám, a naopak to, jak jej hodnotili sami participanté. Většina z nich se shodovala na tom, že česká většinová společnost je laděna značně protiislámsky. Z jejich profesního pohledu se však v souvislosti s muslimy jednalo pouze o jednu z mnoha skupin cizinců, kteří do naší země cestují, popřípadě imigrují. Muslimy a islám v České republice nepovažovali partici-

panti za hrozbu, respektive ne o nic větší, než jakou by mohli být pro Českou republiku a její obyvatele jiné národnostní či etnické menšiny. V tomto kontextu participanté také velmi často hovořili o tom, že muslimská komunita v Česku je poměrně nenápadná a nevýrazná, a to především ve srovnání s komunitou romskou či vietnamskou, které participanté také velmi často zmiňovali.

Dalším interpretovaným pojmem byl pojem *muslim\**. Tento pojem se vztahoval k příslušníkům islámského vyznání, a to jak těm, kteří se nacházejí na území ČR, tak k těm pobývajícím v zahraničí. Ve výpovědním prostoru se tento pojem nachází v blízkosti pojmů *vir\** a *kultura\**. Pokud se podíváme na graf, uvidíme, že pojem *muslim\** se nachází přímo v centru těchto tří výše zmíněných pojmů – *cech\**, *vir\** a *kultura\**.

Pojem *kultura\** byl participanty zmiňován především v souvislosti se specifickým systémem hodnot a norem, které jak obyvatelé ČR, tak zástupci menšin v ČR zastávají. Znázorňuje odlišnost a jedinečnost každého daného etnika. V této oblasti se participanté shodovali na faktu, že kulturní rozdíly, se kterými se při práci se svými klienty setkávají, jsou poměrně značné, což samotnému cizinci stěžuje jeho

úsilí o integraci, respektive začlenění. Samotní pracovníci státní správy a samo-správy však tyto kulturní rozdíly nepovažovali za nepřekonatelný problém, naopak v něm mnohdy shledávali možnost obohacení naší kultury a společnosti. Naše společnost by se tedy, dle jejich názoru, měla více zaměřit na poznávání a chápání těchto rozdílů, a napomoci tak cizincům nejen z muslimských zemí přiblížit a pochopit naši kulturu. V grafu je patrná i blízkost pojmu *spolec\**, který v tomto kontextu poukazoval na fakt, že česká společnost se kulturních odlišností bojí. Místo snahy o pochopení a porozumění ostatních kultur Češi často stereotypizují, k čemuž v negativním slova smyslu přispívají zejména média a nedostatek objektivních informací. Možnost, jak české společnosti přiblížit cizí země a kultury, viděli zaměstnanci státní správy a samosprávy v pořádání kulturních a multi-kulturních akcí, na kterých se sami podíleli a často je přímo organizovali. V případě muslimů byli další možností kulturní akce, pořádané přímo jejich náboženskou komunitou v ČR.

„Když nám ty děti pod vlivem maminek utíkají se někam modlit, tak to je takový jako nezvyklý, že na to musí mít prostor. Nebo takový ty obyčejný věci, když jsme někde na výletě a děti potřebují na záchod a ty holčičky dvanáctiletý mi řeknou, že na záchod do přírody nejdou, protože z domu mají, že se to nedělá, ale neví proč, tak je to takový i spíš někdy úsměvný, někdy je to takový, že mě to rozčílí i vnitřně, protože to komplikuje nějaký situace...“ (participant)

„Myslím si, že je to dané tou mediální reprezentací toho islámu. Jednak je tam velká kulturní odlišnost, islámská kultura je prostě jiná než naše a ten mediální obraz, který tu ti muslimové mají, se spojuje s terorismem a tu kulturu prezentuje jako ne-integrovatelnou. Média to podávají tak, že je to nemožné, aby tu ti lidé mohli v klidu žít, a pořád se poukazuje na všechny problémy, které jsou ve Francii v Londýně a já nevím kde. A všechno se to strašně stereotypizuje na etnický původ těch obyvatel, ja-

kože je to kvůli tomu, že jsou muslimové. A vůbec už nikdo neřeší, že je to třeba kvůli tomu, v jaké jsou ti lidé sociální situaci. Takže nevím, určitě tam hraje roli to, že ta kultura je pro běžného člověka v mnohém nepochopitelná.“ (participant)

Pojem *vir\** chápeme jako popis dvou náboženství, která přicházela v našich rozhovorech ke slovu – *islám* a *křesťanství*. Stejný význam nese v naší analýze i pojem *naboz\**, umístěný v přímé blízkosti pojmu *islam\**. *Náboženství*, jako slovo, bylo více zmiňováno v souvislosti s islámem, než slovo *víra*, které bylo zmiňováno ve stejné míře jak u křesťanství, tak u islámu. Je však více než možné, že se jedná pouze o různé použití slov, jež nemá žádný hlubší význam.

Pojmy *krest\** a *islam\** jsou tedy v grafu ve stejné vzdálenosti od pojmu *vir\**. Je důležité uvědomit si jejich význam především ve vztahu k jejich umístění vůči dalším pojmům v grafu. Tato důležitost bude detailněji rozvedena níže.

Dovolíme si teď tedy, v návaznosti na připojená grafická znázornění, krátké shrnutí výše zmíněného.

Úplně nejbližší průniku os se nachází pojem *cech\**, který odkazuje na české prostředí a kulturu. Nejbližší k němu je umístěn pojem *muslim\**. Oba tyto pojmy vymezují okruh problematiky, o níž jsme s participanty hovořili. Dále se zde nacházejí pojmy *kultur\**, *naboz\** a *vir\**. Všechny jsou umístěny v těsné blízkosti pojmů *cech\** a *muslim\**, a taktéž pojmu *krest\**.

Křesťanství bylo participanty stavěno do protikladu vůči islámu, a to ve více rovínách. To souviselo s evropskou křesťanskou kulturou a jejím vymezením vůči islámské kultuře, stejně jako s křesťanskou kulturou v ČR, která se nyní sice výrazně neprojevuje, ale dle našich participantů je základem toho, co dodnes formuje české hodnoty a co je hluboko zapsané v české historii a kultuře.

„No a samozřejmě náboženství křesťanství, že jo. To je něco, co máme archetypálně v sobě a křížové výpravy a tady ty věci si myslím, že zanechaly v každém z nás tady

*v České republice ať už podvědomý, nebo vědomý obraz o tom, jako jak by měl správný křesťan k islámu přistupovat, no. A tohleto je podle mě směs, která se odráží jakoby v celkovém v tom pojetí té komunity u nás, no.*“ (participant)

*„Tak Evropa je katolická, stála na křesťanskou, na křesťanských základech, a oni, že jo, mají strach, že za chvíli tady budou muslimové... zahrnou celou Evropu no...“* (participant)

Pojmem, který dosud nebyl zmíněn, je pojem *komunit\**. Zástupci menšin se často sdružují do komunit, které jim pomáhají lépe se vyrovnat s prostředím cizí země, a přestože nyní žijí v prostředí s odlišným systémem tradic a hodnot, umožňují jim tyto komunity podržet si i nadále svoji autenticitu. V grafu je také možné vidět, že je slovo v přímé blízkosti pojmů *naboz\**, *islam\**, *muslim\**, *kultur\** a *cech\**, příslušníci islámského vyznání, muslimové, se tedy v českém prostředí sdružují do komunit, které jim umožňují určité zachování jejich náboženské a kulturní jedinečnosti. Jejich vyznání a kultura, které jsou spojené s islámem, byly našimi participanty vnímány v jistém kontrastu vůči českému prostředí, spojeném s křesťanstvím. Je však nutné podotknout, že tuto odlišnost zaměstnanci státní správy a samosprávy nepovažovali za problematickou, nebo snad nebezpečnou pro kulturu v České republice. Tyto vlastnosti jsou islámské kultuře, víře a náboženství, podle participantů, připisovány českou většinou společností, nahlížející na islám a muslimy s odstupem, někdy až xenofobně.

*„My věříme tomu, že xenofobie je založená na tom, že se lidé bojí něčeho, co neznají, a že ve chvíli, kdy to poznají, tak bude šance, že se to změní. Jaké jsou moje zkušenosti, jaká je současná nálada v ČR, jaká je současná politická nálada, vemte si našeho prezidenta, kvůli jeho frčům xenofobním s námi některé arabské země přerušily diplomatické styky.“* (participant)

Zaměstnanci státní správy a samosprávy často vyjadřovali obavu, že vzhledem k od-

lišné kultuře a vyznání nositelů islámského vyznání a obyvatel ČR může docházet ke kulturním a náboženským střetům. Zde byl důležitým pojmem pojem *prav\**.

Tento pojem byl v souvislosti s oblastí státní správy a samosprávy zmiňován ve dvou rovinách. První z nich bylo použití výrazu ve smyslu ochrany lidských práv obyvatel ČR, zahrnujících především právo na svobodu vyznání a ochranu českých tradic. Mnohem výrazněji však z rozhovorů vyplývalo profesionální jednání zaměstnanců s jejich klienty, ve snaze hájit jejich lidská práva v českém prostředí. Pro zaměstnance státní správy a samosprávy byl typický shodný přístup ke všem osobám cizí státní příslušnosti ve snaze zajistit jim co pokud možno nejlepší možné služby a pomoci jim vyřešit jakoukoliv problematickou situaci. Náboženské vyznání klientů pro ně nebylo rozhodující, mnohem důležitější pro ně byla země, protože z takového určení vyplýval okruh problémů, se kterými za nimi klienti docházeli.

Na tuto skutečnost nám v grafu opět poukazují pojmy *naboz\** a *kultur\**. Tyto pojmy mezi sebou vyjadřují popis nejmarkantnějších odlišností českého a islámského obyvatelstva, ale také snahu o zachování spravedlivého přístupu k oběma jmenovaným a zachování jejich lidských práv. Ve svém zaměstnání se setkávají s pestrou škálou cizinců, kdy každý z nich je nositelem specifického kulturního a náboženského dědictví své země. Úkolem státních zaměstnanců není cokoli posuzovat a hodnotit. Jejich hlavním zájmem je pomoci přistěhovalcům s co nejrychlejší adaptací na zdejší prostředí, podmínky a legislativu. V tomto snažení se řídí stanovenými postupy, které nijak nesouvisí s vyznáním ani kulturou země, ze které cizinec pochází. Země původu pouze často předpovídá okruh problémů, se kterými jsou konfrontováni a se kterými se snaží vypořádat.

Pro dokreslení obrazu zbývají pouze tři ústřední pojmy: *Viet\**, *problem\** a *skol\**.

Vietnamská menšina je v ČR velmi výrazně zastoupena a participanti ji ve svých

výpovědích často zmiňovali. Její zástupce považovali za osoby, u kterých proběhla úspěšná integrace do českého prostředí. Je tedy značně zajímavé, že se v blízkosti tohoto pojmu vyskytuje nejbližší právě slovo *problem*\*. Toto spojení slov poukazuje na výpovědi participantů, kteří vietnamskou menšinu považovali za dobře integrovanou, popřípadě velmi nenápadnou, alespoň co se vystupování na veřejnosti týče. Výjimkou je pouze značná aktivita v oblasti obchodování. Jinými slovy se participanté téměř shodovali v tom, že muslimové a islám nepředstavují, alespoň za současného stavu, pro Českou republiku problém, stejně jako je tomu v současné době s vietnamskou komunitou. Se spojením pojmů *problem*\* a *Viet*\* bylo tedy operováno dvěma způsoby; participanté srovnávali absencí problémů s vietnamskou komunitou s absencí problémů u muslimské komunity v ČR. Úspěšná integrace vietnamské komunity do české společnosti byla často zmiňována jako předobraz úspěšné integrace muslimů. Stejně tak je však nutné podotknout, že někteří participanté vnímali rostoucí počet Vietnamců jako potenciální hrozbu, alespoň v oblastech s jejich větší koncentrací. Například v městské části Praha 12, kde se nachází největší středoevropská vietnamská tržnice.

#### MOŽNÉ KOMPLIKACE PRÁCE S PROGRAMEM HAMLET 2.0

Výše jsme představili grafické výstupy programu Hamlet 2.0 spolu s tím, jak vypadá jejich následná interpretace. Nyní se však vraťme k práci s programem samotným.

Po provedení interpretací (tehdy jsme s Hamletem 2.0 pracovali již téměř pět měsíců a měli jsme pocit, že ho konečně ovládneme natolik, aby mohl vzniknout tento text) jsme se ve snaze o provedení kontroly naší analýzy rozhodli, že zkusíme provést mírně odlišný postup, který měl dle našeho úsudku přinést stejné výsledky, jaké nám poskytl graf předešlý.

Tentokrát jsme do manuálního slovníku programu Hamlet 2.0 vepsali slova pouze tak, jak je zobrazil program TextSTAT-2.9c (tedy z jeho frekvenčního slovníku), nepoužívali jsme při tom vůbec frekvenční slovník zobrazený Hamletem 2.0. Předpokládali jsme, že program buď analýzu vůbec neprovede (především kvůli tomu, že nebude schopný najít slova se spřežkami, kterými jsme mu nyní zaplnili slovník), nebo výsledky nebudou dávat smysl. K našemu překvapení však měla analýza nejen hladký průběh, ale výsledný graf se lišil tak nápadně, že nás donutil zvažovat, který z výsledných dvou grafů je správný. Napadlo nás ověřit, jestli jsme postupovali správně. Uvědomili jsme si nepříjemnou možnost, že pouhým přepsáním slov vstupujících do analýzy, si nemusel program Hamlet 2.0 poradit s ostatními slovy v textu (nemusel je být schopen korektně zobrazit a přečíst).<sup>8</sup>

Výše v textu jsme uvedli, že je při práci s Hamletem 2.0 nutné změnit jazykové rozhraní, ve kterém program čte zadaný text tak, aby byl schopný pracovat s českým jazykem. Tuto informaci je opět nutné vyhledat v manuálu, stejně tak jako přesný způsob postupu, kterým toho lze dosáhnout. Pro takové, z pohledu českého jazyka správné zobrazení, je však nutné manuálně vepsat do tabulky k tomu určené ty znaky, které chceme, aby Hamlet 2.0 rozeznal.

Je však ještě jeden klíčový okamžik, kdy je ještě možné v průběhu analýzy jazykové rozhraní změnit ještě před tím, než program vygeneruje finální graf. Hamlet 2.0 na tuto příležitost upozorní při zpracovávání dat, a to otázkou, zda chce uživatel použít to jazykové rozhraní, které je automaticky nastaveno (v našem případě angličtinu). Pokud v této chvíli odmítnete, budou vám postupně nabídnuta pole, do kterých je možné podle klíče (opět uvedeného ke konci manuálu) zadat jeden po druhém ty znaky, které chcete, aby program přečetl. Nejedná se o nikterak složitý postup. Vyžaduje pouze mechanické a časově poměrně nenáročné přepsání znaků se spřežkami, do jejich původní podoby, které jsou následně

upraveny v celém výchozím korpusu. My jsme tento krok nepodnikli, protože jsme netušili, že dojde k takovému rozdílu v našich výsledcích v okamžiku při změně znaků a v okamžiku, kdy tuto změnu neučiníme. Mimo jiné jsme už měli hotový slovník za pomoci programu TextSTAT-2.9c a čtení spřežek programem Hamlet 2.0 jsme nepovažovali za nezbytně nutné.

Přehlédnutí tohoto momentu nás však dovedlo k chybnému výsledku. Ukázalo se, že je potřeba při analýze věnovat pozornost i drobným informacím. Chybu jsme odstranili a výsledek byl po následné kontrole koenečně ten správný.

#### ZHODNOCENÍ VYUŽITÍ PROGRAMU HAMLET 2.0

Závěrem je třeba dodat, že drobné přehlédnutí může mít na výsledné interpretace zásadní vliv. Nutnost začít znovu od začátku celou naši analýzu i s interpretací, která nám zabrala od začátku až do konce pár měsíců, nás naučila, že je třeba velké obezřetnosti při zpracovávání dat a předpokladem je úplné porozumění programovému prostředí. Jsme přesvědčeni, že toto zjištění je užitečné, pohled na různorodé interpretace a výsledky je varováním pro všechny výzkumníky, kteří se potýkají s neznámým programovým prostředím, zvláště tehdy, pokud v nich není zohledněna čeština. Přesto tvrdíme, že Hamlet 2.0 je užitečný a dobrý nástroj. Pro studenty má řadu předností: je bezplatný a volně dostupný. Uživatelé ale může déle trvat porozumění jednotlivým krokům a postupům. Program má také, oproti jiným statistickým programům dostupným na českém trhu, nespornou výhodu v tom, že disponuje možností trojdimenzionálního zobrazení dat, a v dalších přednostech (korespondenční analýza dat, profil textu, porovnání více frekvenčních slovníků a textů mezi sebou navzájem apod.). Pro pokročilejší uživatele tak skýtá velké množství možností, které, podle nás, bohatě vynahrazují počáteční těžkosti způ-

sobené prací s novým neznámým programem.

Je třeba také zmínit, že na trhu existuje program Hamlet 3.0, jehož zkušební verze je možné stáhnout zdarma na měsíc do počítače. Práce s ním však pro nás byla značně nevyzpytatelná; fungoval necelý týden, a poté jsme narazili na potíže a nestabilitu. Na druhou stranu je potřeba uvést, že jazykové rozmezí už zohledňuje češtinu, s čímž souviselo vše, co nám působilo potíže při práci s Hamletem 2.0. K dispozici je jeho placená verze,<sup>9</sup> která se z našeho pohledu vyplatí pouze v případě systematictější a dlouhodobější práce.

Program TextSTAT-2.9c je skvělým výzkumným nástrojem, který lze využít po každé, kdy je nutná rychlá orientace v rozsáhlém textovém materiálu (funkce konkordance, tedy náhledu vybraných slov ve všech větách, v nichž bylo použito). Je nejen schopný pracovat s českým textem, ale také přečíst znaky použité v tomto jazyce. Disponuje přehledností a intuitivním ovládním, a díky schopnosti vytvořit frekvenční slovník je velmi užitečný při kvantitativní textové analýze. Stejně jako Hamlet 2.0 je volně ke stažení, a pro studenty je tedy lehce dostupný.

Oba programy se skvěle se doplňují a práce s nimi pro nás byla velmi obohacující. Stejně tak však výzkumníkům, kteří si zvolí námi popsany typ analýzy a oba výše zmíněné výzkumné nástroje, doporučujeme zvýšenou pozornost při pročitání manuálů a obezřetnost v každé fázi prováděné analýzy. Jejich kombinaci a práci s nimi však při dodržení těchto doporučení vřele doporučujeme.

#### ZDROJE PODPORY

Text vnikl s podporou projektu „Islám v ČR: etablování muslimů ve veřejném prostoru“, který probíhá v rámci Programu bezpečnostního výzkumu České republiky v letech 2010–2015 (VG20132015113).



## POZNÁMKY

<sup>1</sup> Projekt se zabývá podobami, které nabývá etablování muslimů v českém veřejném prostoru. Sleduje způsoby etablování se islámu v české společnosti a jeho důsledky, a přitom vychází z analýzy jednání významných aktérů – veřejnosti, médií, úřadů a dalších. Platí, že muslimská „komunita“ se vyznačuje značnou heterogenitou, stejně jako interakční a etnickou pestrostí, a etablování na veřejnosti tak souvisí se snahou získat veřejný charakter a vliv, což však může vést k řadě sociálních konfliktů (Topinka et al. 2013).

<sup>2</sup> Ukázalo se, že deprivatizační úsilí muslimů se setkává s rezistencí na úrovni každodennosti, samosprávy i státní správy. Povědomí o životě muslimů v české společnosti je velmi nízké. „Pokud ho neprovází osobní zkušenost, tak se aktéři uchylují často k abstraktním a generalizujícím modelům a situaci vnímají především prizmatem událostí, které se odehrávají v zemích západní Evropy, kde je přítomnost muslimů viditelná. To je viditelné ve veřejné správě, v rámci které je samotná téma tabu. Výjimkou jsou velmi specializované organizace, kde je setkávání se s muslimy poměrně frekventované a kde dochází k vnímání heterogenity a jsou přítomny konkrétní obsahy zkušenosti s muslimy a islámem.“ (Topinka et al. 2013, 270–271)

<sup>3</sup> Aktéři pocházeli z krajských úřadů, magistrátů, dále z prostředí Cizinecké policie ČR, Správy uprchlických zařízení Ministerstva vnitra ČR a dalších organizací.

<sup>4</sup> Pro její potřeby nebyly tři výše zmíněné sektory vnímány odděleně, ale jako celek. S ohledem na těsný vztah, co se roviny zájmů, povahy, zaměření a obsahu práce námi oslovených participantů týče, byl takový postup možný, přičemž nehrozilo, že by právě z důvodů rozdílnosti mezi námi dotazovanými participanty ze tří zdanlivě odlišných oblastí došlo k narušení výsledného uceleného výstupu analýzy.

<sup>5</sup> Hamlet 2.0 není schopen pracovat s dokumenty v jiném formátovém rozmezí, než je právě TXT. Provedení tohoto kroku bylo tedy více než nezbytné. Taktéž bylo třeba dbát na kompatibilní kódování vkládaného textu a výše zmíněného programu. Při změně kódování TXT souboru z ANSI na Unicode například došlo při naší práci k tomu, že námi vytvořený frekvenční slovník v programu Hamlet 2.0 již nebylo možné po jeho uzavření a uložení znovu otevřít.

<sup>6</sup> Jedná se skutečně o drobná pravidla, která jsou však poměrně klíčová pro správné sestavení slovníku v programu a jejichž nedodržení může v horším případě znamenat ztrátu dat a nutnost nového začátku, v lepším případě „jen“ ztrátu některých slov, důležitých pro analýzu. Tyto „drobnosti“ bohužel nejsou v anglickém manuálu vůbec popsány a uživatelé začátečníkovi mohou způsobit velké problémy. Většina z nás je například navyklá z běžné práce s textovými dokumenty v počítači používat klávesu DELETE při nutnosti umazat

část textu. Pokud však zapíšete slovo do vámi nahystaného slovníku v Hamletovi 2.0 a zvolíte tento postup, vymaže se vám zároveň s ním úplně vše, co jste dosud do slovníku zapsali. A pokud si vepísované pojmy průběžně neukládáte, nezbuďte vám, než začít svoji práci znovu od začátku. Bohužel, vzhledem k absenci takového upozornění v manuálu, nemá uživatel na podobnou informaci možnost přijít jinak než v okamžiku, kdy k této situaci dojde.

<sup>7</sup> Za takovým účelem nabízí Hamlet 2.0 hned několik kroků analýzy, nebo řečeno přesněji, několik různých analytických výstupů. Zpracuje a zobrazí nejen kvantitativní hodnoty vkládaného textu (počet slov, frekvenci apod.), ale také graficky zobrazí souvislosti mezi jednotlivými skupinami slov, a ve výsledném grafu i mezi slovy samotnými.

<sup>8</sup> Tato domněnka vyvstala z již zmíněného problému, tedy z nesoučinnosti programu se slovy obsahujícími špréžky. Pokud by to byla pravda a program by skutečně tato slova kvůli špréžkám, která by nadále obsahovaly, nezobrazil celá, nemusejí být vzdálenosti mezi nimi správné.

<sup>9</sup> Cena se pohybuje kolem 3000 Kč, pro studenty je o něco nižší.

## POUŽITÉ ZDROJE

- Analýza mediálních sdělení o přírodních vědách: Monitoring médií.* 2007. Viděno 28. července 2015. [http://www.projektmedved.eu/docs/2007\\_medved\\_MONITORING\\_final\\_III.doc](http://www.projektmedved.eu/docs/2007_medved_MONITORING_final_III.doc).
- Burgertová, E. 2006. *Přehled tisku za květen – neviditelná menšina.* Viděno 29. března 2015. <http://www.migraceonline.cz/cz/e-knihovna/prehled-tisku-za-kveten-neviditelna-mensina>.
- Burton, G. a J. Jiráček. 2001. *Úvod do studia médií.* Brno: Barrister.
- Denní tisk.* 2008. Viděno 18. února 2008. [http://www.mafra.cz/denni\\_tisk](http://www.mafra.cz/denni_tisk).
- Discourse (Oxford Dictionaries).* 2015. Viděno 9. září 2015. <http://www.oxforddictionaries.com/definition/english/discourse>.
- Everitt, B. a S. Rabe-Hesketh. 1997. *The analysis of proximity data.* New York: J. Wiley.
- Foucault, M. 2002. *Archeologie věděni.* Praha: Herrmann & synové.
- Foucault, M. 1976. *Birth of the clinic.* London: Routledge.
- Hájek, M., J. Kabele a K. Vojtíšková. 2006. „Zázemí a ‚bojiště‘ v usilování o spravedlnost: Textová analýza odborářské, feministické a lidskoprávní mediální komunikace.“ *Sociologický časopis* 42 (2): 279–280.

Cholt, J. 2006. *Diskurzivní analýza globálních změn klimatu případ emisí – argumentační strategie aktérů*. Diplomová práce. Praha: Fakulta sociálních věd Univerzity Karlovy v Praze.

Jørgensen, M. a L. Phillips. 2002. *Discourse analysis as theory and method*. London: SAGE Publications.

Kaderka, P. a K. Karhanová. 2001. *Obraz cizinců v médiích: Zpráva o projektu za rok 2001*. Viděno 29. března 2015. [http://www.cizinci.cz/files/clanky/121/Obraz\\_cizincu\\_v\\_mediich.pdf](http://www.cizinci.cz/files/clanky/121/Obraz_cizincu_v_mediich.pdf).

Klváčová, P. a T. Bitrich. 2003. *Jak se (ne)píše o cizincích v českém tisku*. Viděno 29. března 2015. [http://aa.ecn.cz/img\\_upload/bacca6403f52f8f0f60667687afefbf5/](http://aa.ecn.cz/img_upload/bacca6403f52f8f0f60667687afefbf5/)

[Jak\\_se\\_ne\\_pise\\_o\\_cizincich\\_2003.pdf](#).

Křížková, M. 2007. „Neviditelná menšina.“ In *Cizinci, naši a média: Mediální analýzy*. Praha: Multikulturní centrum.

Topinka, D. 2010. „Integrace muslimů/imigrantů do ‚české‘ společnosti.“ In *Individualizace, náboženství a identita: Poznámky k současné sociologii náboženství*, ed. D. Lužný a D. Václavík, 253–301. Praha: Malvern.

Topinka, D., T. Janků, L. Linhratová a J. Zádina. 2013. „Muslimové imigranti v České republice: Etablování na veřejnosti.“ In *Fenomén moci a sociálně nerovnosti*, ed. L. Bomba, E. Kővérová a M. Smrek, 238–273. Bratislava: Univerzita Komenského v Bratislavě.