

1. DIGLOSIE A MULTIDIMENZIONÁLNÍ ANALÝZA

1.1 Historické expoé

Pojem diglosie je v souvislosti s českou jazykovou situací používán velmi brzy poté, co byl ustaven ve světové lingvistice Charlesem Fergusonem (1959). Jedno z prvních použití tohoto konceptu (ne-li úplně první) se objevuje ve stati Petra Sgalla (1963) na závěr známé diskuse o obecné češtině.

Sgall ve svém příspěvku tvrdí, že „[v] češtině je situace blízká diglosii, liší se však zejm. tím, že spis[ovná] a ob[ecná] čeština jsou útvary poměrně velmi blízké (jde jen o odchylky v některých složkách) a že se jejich prvky v běžně mluvených projevech (stylu jednacím, konverzačním, popř. i odborného aj., někdy zase obcovacím) často kombinují.“ (Sgall 1963: 246) Zároveň poukazuje na odlišnost české situace od prototypické diglosie, jak ji definoval Ferguson, v tom, že spisovná čeština (SČ) a obecná čeština (OČ), tedy variety V a N, nemají v zásadě odlišné funkce. Dokumentuje to na příkladech OČ prostředků v mluvených odborných projevech (nejspíše šlo o citáty z přednášky V. Skaličky), které byly neuralgickým bodem celé diskuse a jejichž autentičnost byla Sgallovými oponenty zpochybňována: „Kdybychom i v případě, že typičnost takových projevů [tj. spisovných projevů s OČ prvky – VC] bude dalším výzkumem potvrzena, setrvali na třídění útvarů zároveň podle rozdílů v podobách prostředků i v jejich bohatství (jako při diglosii), museli bychom říkat, že OČ ‚přejímá‘ ze spisovné velké množství odborných výrazů aj., spisovná z obecné deminutiva atd. Ale vymezíme-li útvary jen podle prvního z obou rysů [tj. pouze podle odlišnosti prostředků, nikoli bohatství funkcí – VC], pak jsme v situaci výhodnější: můžeme mluvit o obecně českých prostředcích ap., ne o spisovnosti či nespisovnosti projevu jako celku.“ (Sgall 1963: 247, podobně i Sgall et al. 1992: 51). Sgall tedy poukazem na pronikání OČ prostředků do stylově vyšších projevů odmítá představu funkční komplementarity mezi SČ a OČ, ačkoli je to jeden z definičních rysů diglosie (Ferguson 1959: 328, Gvozdanović 2014: 11), a naopak zdůrazňuje především velký funkční překryv mezi SČ a OČ (výše zmíněné odborné termíny nejsou čistě spisovné, deminutiva nejsou čistě obecněčeská).⁷

Zatímco Sgallův postoj se od 60. let v zásadě nemění, stanovisko ostatních účastníků diskusí o stratifikaci národního jazyka prochází určitými vývojovými fázemi. Např. Sgallův oponent v diskusi a nejvyšší lingvistická autorita v 60. letech Bohuslav Havránek s hodnocením české situace jako diglosní rezolutně nesouhlasí (Havránek 1963: 257). Argumentem je mu především významný překryv prostředků obou variet (v morfologii ho odhaduje na 90 %).

7 Srov. s pozdější formulací, v níž sice setrává na výrazném překryvu, nicméně přiznává i určitou míru funkční odlišnosti: „Funkce standardní a obecné češtiny se sice liší v tom, že druhá jmenovaná varieta neslouží jako standard; jevy OČ se však vyskytují často a jsou bez váhání přijatelné v neutrálním stylu běžné mluvy, včetně vyšších stylistických vrstev...“ (Sgall et al. 1992: 215–216, přeložil VC)

František Daneš ve svém příspěvku z roku 1988, který je citován zejména v souvislosti s kategorickým odmítnutím konceptu hovorové češtiny jako samostatného útvaru, uvádí, že česká jazyková situace diglosii připomíná.⁸ Zdůrazňuje přitom především fakt, že SČ neplní funkci běžného dorozumivacího prostředku, kterou má OČ, a dovozuje: „[k]dyby tato obecná čeština představovala relativně jednotnou (územně výrazně nediferencovanou) varietu, představovala by současná česká jazyková situace vyhraněný případ diglosie.“ (Daneš 1999: 292) Není přitom zcela jasné, proč by vnitřní regionální diferencovanost OČ měla v otázce diglosie hrát nějakou úlohu (Daneš sám upozorňuje, že je taková rozrůzněnost u nižší variety v diglosiích poměrně častým jevem).

V 90. letech se autoři opět přiklánějí k hodnocení české jazykové situace jako nediglosní zejména z toho důvodu, že rozdíly mezi SČ a OČ nejsou velké; Bermel (2000: 33) hodnotí českou situaci jako případ ustupující diglosie, Sgall a jeho okruh setrvává na tom, že v české situaci některé diglosní jevy přítomny jsou, ale o „skutečnou“ diglosii nejde (Čermák 2002; Sgall et al. 1992: 215; Sgall & Hronek 1992: 26). Vedle překryvu prostředků a funkcí mezi SČ a OČ je je také uváděn aspekt akvizice SČ.⁹

Uvedení pojmu *diglosie* do české lingvistiky v kontextu debat o míře jazykové regulace a postavení OČ vedle SČ značným způsobem determinovalo jeho přijetí. Přes veškeré kontroverze a problematické aspekty, které jsou s ním spojeny (viz Nekvapil 2017), se jedná o termín k neutrálnímu (bezhodnotovému) popisu sociolingvistické situace daného jazykové společnosti; v českém prostředí je ovšem podvědomě vnímán s negativními konotacemi (viz Cvrček 2008: 32). To není překvapivé, uvědomíme-li si, jaké jsou cíle jazykové regulace v rámci teorie jazykové kultury; konkurence dvou kódů, z nichž jeden je oficiálně regulován („kultivován“) a druhý se vyvíjí a uplatňuje spontánně vedle něj, není zcela ve shodě s ideálem funkčního přístupu k jazykové kultuře, jak byl budován od 30. let. V něm byl spisovný jazyk spojován se všemi důležitými komunikačními funkcemi (Starý 1995: 36) a díky expertně-technokratickému tříbení měl být ve všech ohledech superiorní. Noetický optimismus 30. let, v jehož ovzduší teorie jazykové kultury vznikala, si vůbec nepřipouštěl možnost, že by kód, který je výsledkem vědecky podloženého tříbení, mohl nebýt přijat v některé z komunikačních funkcí (vč. běžné konverzace). Poukaz na diglosii se tak stává implicitním poukazem na nefunkčnost teorie jazykové kultury.

Pokud přijmeme předpoklad v pozadí těchto úvah, že diglosie je v určitém ohledu nežádoucí, otevírá se tím prostor pro intervence motivované snahou o zmenšení rozdílu mezi OČ a SČ, tj. to, co Bermel označuje jako demontáž (*dismantling*) diglosie z oficiálních pozic (Bermel 2014: 26). Otevřeně v tomto směru plédují lingvisté mající názorově ke Sgallovi blízko: např. Čermák (2002) v souvislosti s reformou pravopisu, která může připustit některé prostředky mluvené češtiny do kodifikovaného standardu, a přispět tak k „neprohlubování“ diglosie. Tento postup je v intencích teorie jazykové kultury možný v zásadě pouze za předpokladu, že existuje výrazný funkční překryv mezi OČ a SČ, jinak

8 V příspěvku Daneš prohlašuje, že do souvislosti s českou situací uvedl diglosii poprvé on sám v konferenčním příspěvku z roku 1968, resp. v přednáškovém cyklu z akademického roku 1965/66 (Daneš 1999: 292), což – jak vidíme z našeho expozé – se nezakládá na pravdě.

9 Srov. „Osvojování SČ je tradičně především záležitostí školy, nicméně rozdíl mezi oběma kódy není tak zásadní, aby méně vzdělaní mluvčí nemohli SČ porozumět i bez specifického vzdělání.“ (Sgall et al. 1992: 216)

by totiž požadavek na úpravu kodifikace směrem k připuštění některých OČ prvků nemohl projít přes kodifikační kritérium funkce, jak ho definoval Daneš (1979: 87).

Vedle posunů v lingvistických názorech dochází také k proměnám v jazyce samotném. Proces, který probíhá paralelně se snahami lingvistů o odstranění diglosie a který Bermel nazývá jako zánik (*dissolution*) diglosie (Bermel 2014: 27), souvisí především s proměnami beletrie (srov. Poukarová & Cvrček 2022). Analogicky k pohybům v jiných jazycích, kde v minulosti došlo k zániku diglosie, jako např. v ruštině, japonštině či čínštině (srov. Arokay et al. 2014), sledují lingvisté i v češtině, že se v beletrii – nejprve v překladech – objevuje větší podíl OČ a tím dochází k přirozenému tlaku na normu standardu. V češtině je tento trend o to výraznější, že norma SČ byla tradičně odvozována od literární produkce (srov. Ertlovo kritérium dobrého autora, viz Starý 1995: 122). Tento trend je reflektován a dále využíván při debatách o kodifikaci (srov. vyjádření o významných překladech uvádějících do české beletrie OČ u Sgalla & Hronka 1992; viz 1.3¹⁰).

1.2 Empirický pohled na diglosii prostřednictvím multidimenzionální analýzy

Z historického přehledu uvedeného výše plyne, že ne/přítomnost diglosie v českém prostředí byla předmětem polemik, zřídka však byl na tento fenomén uplatňován empirický pohled (mezi výjimkami můžeme zmínit např. práci Bermel 2000). Komplementární perspektivu může nabídnout multidimenzionální analýza (MDA), jiný vlivný koncept vytvořený za účelem postžení funkční variability v jazyce (Biber 1988: 1995).

Blízkost obou přístupů vyjde najevo, uvědomíme-li si, že koncepty, s nimiž MDA běžně operuje – registr a žánr – lze označit za nadindividuální variety, stejně jako jsou varietami sociolekty a dialekty, s nimiž v případě diglosie operuje sociolingvistika. Registr, jakožto emblematický pojem MDA, lze charakterizovat jako situačně podmíněnou varietu (Biber & Conrad 2009: 5; Mrázková 2016), zatímco (inter)dialekt či sociolekt jsou podmíněné regionálně a/nebo sociálně, což samozřejmě nevyklučuje jejich dis/preferenci v určitých komunikačních situacích.¹¹

Na druhé straně není možné pominout odlišnosti obou perspektiv (např. diachronní či sociálně-kulturní aspekt diglosie, který v typické MDA jako empirické korpusové metodě logicky chybí). Ve výsledku tak pohled na diglosii prizmatem MDA může pouze navštívit otázky související s odlišnými prostředky/rysy konkurenčních variet a jejich funkcemi v synchronním pohledu.

V následujících oddílech nejprve stručně představíme MDA jako výzkumnou metodu, poté popíšeme český model registrové variability, který byl na jejím

10 Za přelomový bývá v pracích Sgallova okruhu často pokládán překlad Salingerova románu *Kdo chytá v žitě*, který do češtiny převedli manželé Pellarovi v roce 1960 a který využíval OČ jako základní varietu.

11 Zároveň je třeba upozornit, že toto pojetí, kdy variety lze nahlížet z různých perspektiv, ale na popis jejich vzájemných vztahů to v zásadě nic nemění, není všeobecně přijímáno. Sgall et al. např. uvádějí, že jazykové útvary (*language formations*) jsou především charakterizovány specifickými prostředky a slouží „poměrně širokým a rozmanitým společenským a stylistickým funkcím“ (Sgall et al. 1992: 50). Naproti tomu styly a registry jsou podle nich striktně funkčně vymezeny.

základě vytvořen, a na závěr tohoto oddílu načrtneme možnosti jeho využití k prozkoumání míry diglosie v současné češtině. Následující sekce se pak bude věnovat výsledkům a jejich interpretaci.

■ 1.2.1 Multidimenzionální analýza

Primárním cílem MDA je identifikovat hlavní dimenze variability v textech na základě analýzy toho, jak jsou využívány variantní prostředky, u nichž lze předpokládat funkční uplatnění (Cvrček et al. 2018). Sekundárně pak můžeme z pozic textů na těchto dimenzích odvodit, jaký je inventář nejpoužívanějších registrů.

Postup při vytváření multidimenzionálního modelu registrové variability předpokládá provedení pěti kroků (detaily viz Cvrček et al. 2020):

1. **Vytvoření diverzifikovaného korpusu** – v našem případě se jedná o 10milionový korpus Koditex (Zasina et al. 2018), který obsahuje texty různých žánrů psaného jazyka, mluvené i webové komunikace (podrobnosti o složení viz Zasina & Komrsková 2019); klíčovým požadavkem při vytváření korpusu bylo jeho pestré složení, korpus proto není tvořený celými texty, ale pouze úseky v rozsahu 2000–5000 slov.
2. Pořízení **soupisu jazykových rysů**, které lze považovat za nositele variability – tyto rysy se rekrutují ze všech rovin jazykového popisu (od fonologických přes morfologické a lexikální až po syntaktické, pragmatické či rysy založené na analýze celého textu). Východiskem pro pořizování tohoto soupisu byly jednak rysy použité v modelech pro jiné jazyky, souborné práce o české variabilitě či stylistice a další nepublikované zdroje (interní seznamy různých jevů užívané pro značkování korpusů ČNK, intuice autorů modelu atp.).
3. Technicky nejobtížnější je **operacionalizace rysů** z bodu 2, jejímž cílem je zjistit kvantitativní charakteristiku rysů v každém textu korpusu (ve většině případů je vyhovující relativní frekvence, v některých případech ale může jít o index, např. lexikální bohatost textu). Tato operacionalizace je pracná sama o sobě, pro věrohodnost celého modelu je navíc třeba u každého rysu zjistit jeho spolehlivost (právě neschopnost uspokojivé operacionalizace byla jedním z nejčastějších důvodů, proč rys ze seznamu vyřadit).
4. Údaje z předchozího kroku jsou **statisticky vyhodnocovány pomocí faktorové analýzy**, jejímž úkolem je zjistit korelaci rysů a na jejím základě vytvořit faktory (latentní proměnné), které jsou následně interpretovány jako dimenze variability; klíčovým vstupním parametrem je počet dimenzí, které mají být faktorovou analýzou extrahovány – model, s nímž tu budeme pracovat, je založen na 8 dimenzích (k motivaci tohoto rozhodnutí viz Cvrček et al. 2021).
5. **Interpretace výsledků** faktorové analýzy se děje pomocí dvou sad informací: jednak jde o váhu rysů v dané dimenzi (*loading*) a jednak o pozici textu na každé dimenzi (*factor score*).

V zásadě lze celý postup shrnout do popisu výzkumného cíle MDA: v průběhu vytváření textu musí mluvčí/pisatel učinit celou řadu rozhodnutí při volbě

z variantních prostředků; tato rozhodnutí nejsou izolovaná a nezávislá, jsou vzájemně provázána a řízena komunikačními funkcemi a záměry. Zatímco komunikační funkce či záměry jsou obtížně empiricky uchopitelné, použité variantní prostředky jsou explicitní, a tudíž je můžeme korpusově analyzovat. Na základě informací o souvýskytu variantních prostředků v textech se MDA pokouší konstruovat síť vztahů mezi jazykovými rysy pomocí modelu, v němž jsou jako dimenze zachyceny hlavní aspekty/charakteristiky variability v daném jazyce.

■ 1.2.2 Český model variability

Analýzou výsledků na korpusu Koditex byl vytvořen multidimenzionální model pro češtinu, který operuje s 8 dimenzemi (podrobnosti k interpretaci výsledků viz Cvrček et al. 2020: 92–105), které byly následně pojmenovány podle krajních poloh, na nichž se texty mohou vyskytovat:

1. dynamický (+) vs. statický (–),
2. spontánní (+) vs. připravený (–),
3. vyšší (+) vs. nižší (–) stupeň koheze,
4. polytematický (+) vs. monotematický (–),
5. vyšší (+) vs. nižší (–) míra explicitní adresnosti,
6. obecný (+) vs. konkrétní (–),
7. prospektivní (+) vs. retrospektivní (–),
8. postojový (+) vs. faktuelní (–).

Každý text je možné umístit do prostoru vymezeného těmito dimenzemi a jejich průsečík přesně charakterizuje registr, který byl k jeho vytvoření použit. Pro další výklad budou klíčové první dvě dimenze, které zároveň vysvětlují největší podíl variability z celého modelu.

První dimenze (**dynamický** × **statický**) je především tvořena opozicemi rysů verbálních (či adverbálních) a nominálních (či adnominálních). Směrem k pozitivnímu pólu této dimenze přispívají rysy jako frekvence minulého času sloves, infinitivů, perfekta u sloves, zájmen pro 3. osobu či různých (adverbiálních) určení času (tedy povětšinou rysy operující na slovese nebo ho rozvíjející). Ke konstituování negativního pólu naopak přispívají rysy jako frekvence různých substantivních pádů, substantivních přívlasků, adjektiv, abstrakt, klastrů adjektiv či substantiv či verbální substantiva (tedy rysy nominální nebo sloužící k nominalizaci/deagentizaci výpovědi). Ačkoli se můžeme v ojedinělých případech setkat s tím, že dynamická akce je zachycena nominálními prostředky (a naopak), ve většině případů jsou rysy z pozitivního pólu asociovány s narativitou či aktivními mentálními procesy (kontemplace, uvozování řeči), zatímco rysy z negativního pólu většinou referují o jevech staticky. Tato opozice verbální × nominální není specifická pro češtinu, objevuje se i v analýzách jiných jazyků a koresponduje se dvěma univerzálně přítomnými strategiemi rozvíjení textu (*clausal-phrasal*): na jedné straně jde o přidávání klauzí (což se projevívá ve větší verbalitě, tedy dyna-

mičnosti textu), na straně druhé se rozvíjí fráze přidáváním dalších jmenných částí (Biber 2014). Jelikož dynamičnost nemusí nutně implikovat pouze naraci, je také svázána s kontemplací (*verba dicendi a cogitandi*) a vytvářením souvětí s hlubokým zanořováním typickým pro mluvené projevy, objevují se mezi dynamickými žánry nejen různé druhy románů a narativních textů, ale také dopisy, webová fóra a většina spontánně mluvených komunikátů. Statické žánry jsou v korpusu zastoupeny administrativními, encyklopedickými či odbornými texty z různých oborů.

Specifičnost české situace¹² tkví v osamostatnění druhé dimenze (**spontánní** × **připravený**), která sdružuje rysy objevující se většinou pospolu s distinkcí *clausal-phrasal*; kombinace těchto charakteristik pak vedla např. Bibera k uplatnění termínů *oral-literate* (Biber & Conrad 2009: 228–230) nebo ještě dříve *involved-informative* (Biber 1988:115) s explicitním zdůrazněním, že tato distinkce se nekryje s rozdílem mezi módem psaným a mluveným.

Druhá dimenze v českém modelu je tvořena na pozitivním pólu třemi okruhy rysů: a) jevy související s *hic et nunc* produkcí projevu (výplňková slova, sémanticky vyprázdněná a neurčitá adverbia, repetitivnost a nízká lexikální bohatost), b) jevy typické pro vytváření a udržování kontaktu s komunikačním partnerem (kontaktní výrazy, demonstrativa, určení místa, expresivní výrazy a interjekce) a c) jevy typické pro OČ (zejm. hláskové změny jako *y > ej, é > í*, které nejsou exkluzivně přítomné pouze v mluveném projevu, ale pronikají i do psaného vyjadřování a mají ustálenou grafickou podobu). Negativní pól je pak tvořen rysy, které odkazují k vyšší míře promyšlenosti a pečlivější stylizaci komunikátu (výskyt a rozmanitost předložek či druhů vztazných vět, vyšší lexikální bohatost a užívání delších slov). Tato dimenze rozděluje texty na připravené (extrémem jsou opět administrativní, encyklopedické a vědeckonaučné texty) a texty spontánní, kam patří zejména mluvené interaktivní projevy (soukromá konverzace, elicitované projevy, veřejné diskuse v TV a rozhlase), z dalších pak scénáře, webová fóra a dopisy.

Ostatní dimenze pokrývají další aspekty variability textů v současné češtině, které nicméně pro účely popisu vztahu k fenoménu diglosie nejsou podstatné.

Pro představu, jaký typ informace model poskytuje, můžeme uvést ukázkou dvou textů, které se nacházejí na opačných extrémech prostoru vymezeného výše zmíněnými dvěma dimenzemi. Na jednom konci škály jsou dynamické a spontánní texty, které nacházíme zejména u mluvených neformálních (neveřejných) a interaktivních promluv:

- (1) *facebook obecně je nebezpečnej s tím že ty nemůžeš ovlivňovat co . oni tam o tobě dají . co oni tam dají za tvoje fotky co oni tam dají za tvoje videa . a jakmile máš nákej profil a někdo si na něho klikne tak už to všechno </s> <s> nande </s> <s> no ale tady </s> <s> jako zvlášť </s> <s> tak taji to má asi vole asi --- každej má svůj profil (Koditex)*

Jedná se o přepis promluvy jednoho účastníka neformálního rozhovoru (včetně pauzové interpunkce), který obsahuje kromě výrazných rysů OČ také velký počet klauzí (vzhledem k rozsahu úryvku), což implikuje velký podíl slo-

12 Vzhledem k původnímu Biberovu výzkumu (Biber 1995) vycházejícímu z aplikace MDA na angličtinu, korejštinu, tuvalštinu a somálštinu.

ves. Naproti tomu statický a připravený extrém reprezentuje např. tento text administrativní povahy:

(2) *Indikátorem je např. celková i měrná spotřeba surovin a energií nebo celková i měrná produkce odpadu v dané technologii. V rámci projektu prevence sledujeme jednak absolutní hodnoty snížení množství produkovaných odpadů a znečištění, jednak změny hodnot indikátorů.* (Koditex)

Ačkoli mají úryvky srovnatelnou délku, druhý je zjevně kondenzovanější, obsahuje méně klauzí, a tedy i sloves, zato má propracovanější nominální fráze (vč. nominalizace pomocí verbálních substantiv). Nepřítomnost rysů OČ ani dalších jevů typických pro mluvené nepřipravené vyjadřování je zjevná.

Na základě pozice textů v dimenzích bylo vymezeno 10 registrů, které si můžeme představit jako shluky textů v určitých oblastech 8D prostoru. Podle převažujících charakteristik těchto klastrů byly registry pojmenovány; pro snazší orientaci jsou rozděleny do dvou makroskupin po pěti registrech podle první a zároveň nejdůležitější dimenze (Cvrček et al. 2020: 119):

1. analýza: statický monotematický registr
2. popularizace: statický polytematický obecný registr
3. žurnalistika: statický nevyhraněný registr
4. fakta: statický polytematický konkrétní registr
5. argumentace: statický kohezní registr
6. anketa: dynamický neadresný registr
7. konverzace: dynamický spontánní registr
8. komentář: dynamický postojový registr
9. scénář: dynamický adresný registr
10. narace: dynamický retrospektivní registr

Smyslem a cílem této klasifikace je přitom poskytnout charakteristiku textů, která by byla komplementární k žánrové specifikaci. Není proto překvapivé, že registrové a žánrové zařazení se leckdy nemusí shodovat; v korpusu nacházíme texty, které mají mluvený žánr, a přitom rysy mluveného jazyka nesdílejí nebo naopak beletristické texty, které napodobují mluvený jazyk poměrně věrně (aniž by se nutně jednalo o scénář):

(3) *Předevěrem jsme s mojí sousedkou uzavřeli výměnnej obchod: její novej sexoň, nějakej Armén, co v Praze dělá pro jedny nóbl aerolinie, mi zdarma sežene otevřenou byznys letenku do vybraný destinace a já na oplátku vysexuju Arménova kamaráda, nezkušenýho zakuklenýho teplouška, co pro tu stejnou leteckou společnost dělá správce sítě. Tenhle ajtáček si prej potřebuje s někým zkušeným ujasnit, jak to vlastně má.* (Koditex)

Ukázka obsahuje poměrně vysoké zastoupení některých rysů OČ, zároveň ovšem přejímá vyjadřovací prostředky běžně užívané v rámci komentářů a diskusních příspěvků na internetu, proto je zařazena k dynamickému postojovému registru.

■ 1.2.3 Využití multidimenzionální analýzy pro výzkum diglosie

V přístupu k diglosii můžeme výsledky MDA využít v zásadě tam, kde jde o srovnání použitých prostředků typických pro jednu z variet a komunikačních situací/funkcí. Vzhledem k tomu, že dimenze variability, jak jsme je představili výše, lze chápat jako svazky rysů, může nám jejich ustavení prozradit něco o tom, v kontextu jakých rysů se objevují rysy typické pro N. Výše jsme zmínili to, že registry, které s pomocí dimenzí ustavujeme, můžeme na obecné úrovni vnímat jako variety; žánry pak reprezentují velmi široce pojatou komunikační situaci. Využití registrů při výzkumu diglosie může spočívat v tom, že analyzujeme ty registry, které obsahují hodně rysů typických pro N, a to s ohledem na žánry (tj. komunikační situace či funkce), které naplňují.

Prakticky se v dalším oddíle zaměříme nejprve na postavení rysů OČ v různých dimenzích, tedy v jakých komunikačních funkcích se OČ objevuje. V zásadě se zde můžeme omezit na druhou dimenzi, kde jsou rysy OČ výrazněji přítomné.

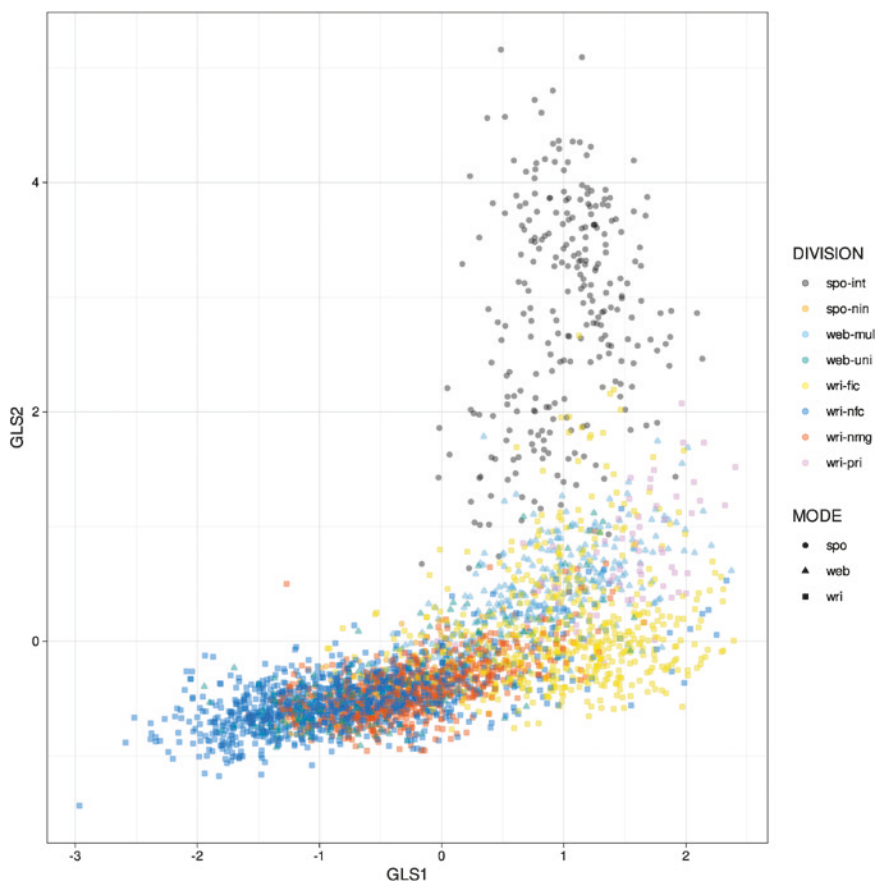
Druhým aspektem hodným prozkoumání je postavení specifických žánrů na jednotlivých dimenzích. Zde se budeme zaměřovat na to, jestli 2. dimenze vyděluje určitý ucelený okruh textů, které bychom mohli definovat nějakou funkcí či komunikační situací, což by odpovídalo funkční komplementaritě předpokládané přítomnosti.

Konečně třetí aspekt hodný prozkoumání souvisí s registry, a především s tím, v jakých žánrech se objevují ty registry, u kterých bychom v souladu s diglosní interpretací předpokládali realizaci v N, tj. v OČ.

1.3 Výsledky

Reinterpretace výsledků české MDA (Cvrček et al. 2020) s ohledem na přítomnost diglosie v češtině nabízí několik podpůrných a několik odporujících argumentů. V zásadě přítomnost z definičních rysů diglosie uváděných Fergusonem cílíme pouze na přítomnost dvou variet konkurujících si v různých komunikačních funkcích (a zbytek aspektů ponecháváme stranou). Tento zjevný deficit (či parciálnost) přístupu zde rozvíjeného je možno kompenzovat silou empirického argumentu, který je proveden na širokém spektru textů reprezentujících různé komunikační situace, jako je tomu v případě MDA založené na korpusu Koditex.

Jedním z nejpřímějších argumentů na podporu diglosní interpretace současné situace je graf reprezentující postavení jednotlivých textů v rámci prvních dvou dimenzí (obr. 1). Viditelná separace šedých bodů, které reprezentují mluvené interaktivní texty (konverzace, elicitované texty a nepřipravené veřejné/vysílané rozhovory), od zbytku korpusu je relevantní indicií o tom, že tyto texty se vydělují od ostatních specifickou komunikační situací (spontánnost), která je také doprovázena specifickými prostředky (mj. i rysy OČ).

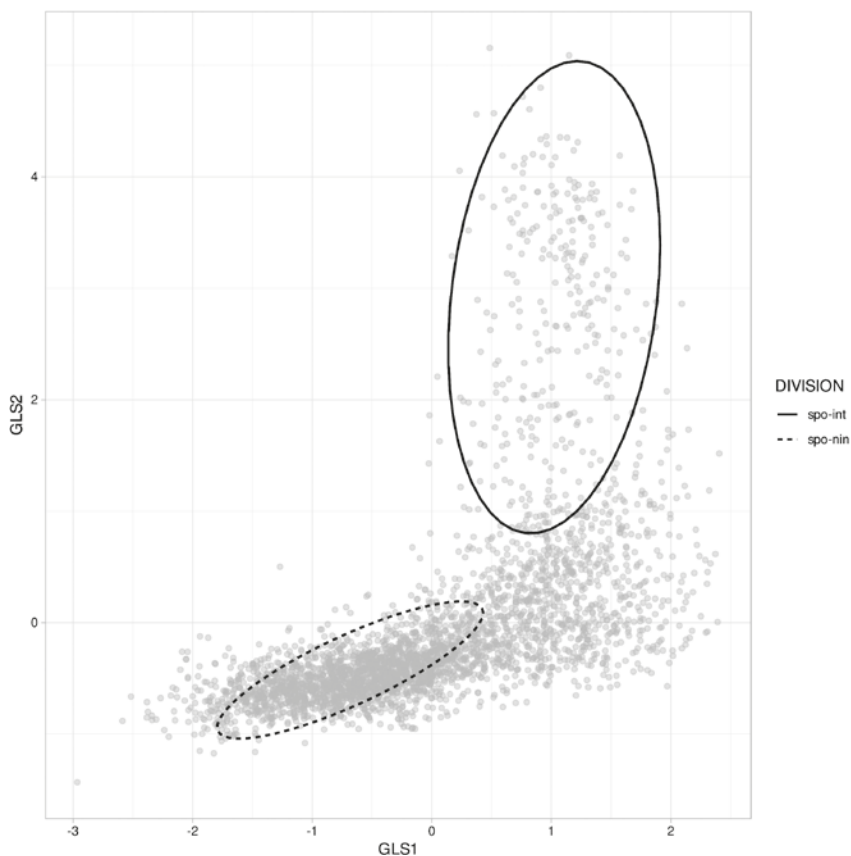


Obrázek 1: Pozice textů v dimenze 1 a 2 v rámci multidimenzionální analýzy modelu registrové variability

Do prostoru, v němž se objevují šedé tečky, se dostávají pouze některé texty beletristické (žluté čtverce), webová vícesměrná komunikace (modré trojúhelníky) a soukromá korespondence (růžové čtverce). Ve všech případech jde o komunikační situace, v nichž je formálnost a připravenost potlačena ve prospěch spontánnosti a kde se zároveň objevují rysy OČ.

Z mluvených žánrů se v horní části grafu neobjevují pouze připravené veřejné projevy (*written-to-be-spoken*), které mohou sice mít svoji akustickou podobu, jazykovým ztvárněním jsou však psané (ozvučený psaný text). Odlišnost těchto dvou mluvených žánrů zachycuje obr. 2. Z něj plyne, že separace dimenze 2 v českém modelu není pouze záležitostí módu (grafického vs. akustického), ale odlišných jazykových prostředků.

S argumenty založenými na obr. 1 úzce souvisí i další indicie ve prospěch diglosní interpretace: specificky české osamostatnění druhé dimenze, která se v analýzách jiných jazyků samostatně neobjevuje, ale vyskytuje se v kombinaci s charakteristikami zastupovanými dimenzí 1. Svědčí to především o určitém rozpojení verbality a spontánnosti v české jazykové situaci. Zatímco v jiných jazycích je míra verbality (nebo *clausal* principu, viz výše) inherentně spojená



Obrázek 2: Pozice mluvených žánrů v prostoru vymezeném dimenzí 1 a 2 – elipsa vykreslená plnou čarou představuje interaktivní mluvené žánry (konverzace, elicitované projevy, televizní/rozhlasové talk-show), elipsa načrtnutá čarou přerušovanou *written-to-be-spoken* projev

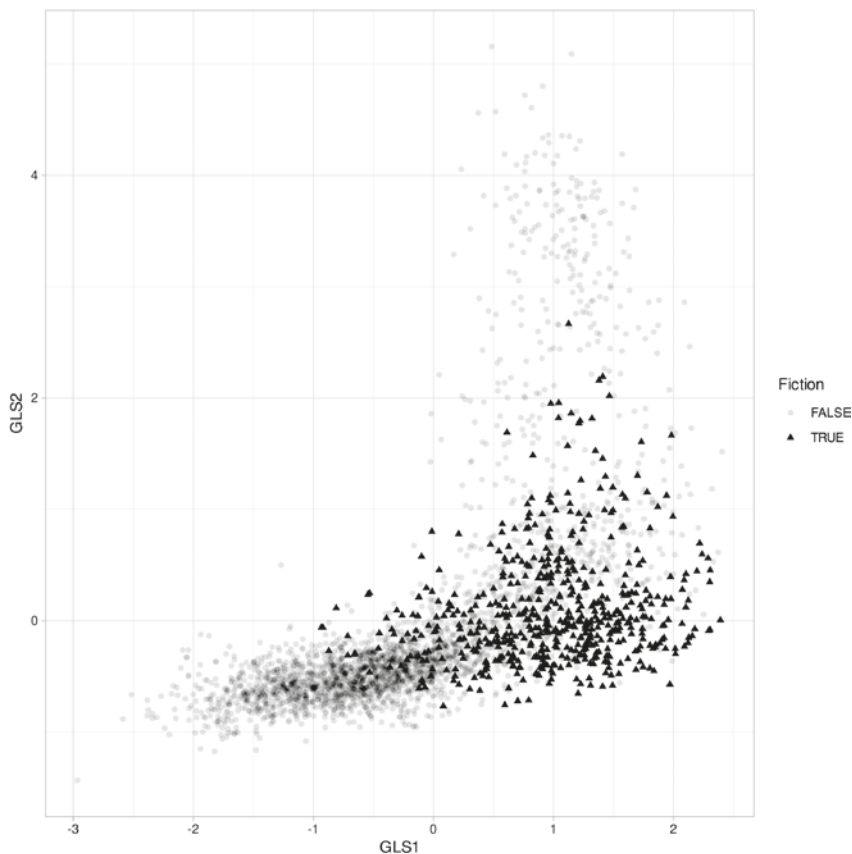
s prostředky charakteristickými pro spontánní komunikaci, v češtině je tato spojitost poněkud oslabena ve prospěch unikátních signálů mluvenosti. Pokud modely jiných jazyků kombinují prostředky, které jsou v českém modelu separovány, znamená to, že odlišnost mluvených textů vnáší do dat takovou míru variability či odlišnosti, která se ve statistickém vyhodnocení promění v samostatnou dimenzi, jež je navíc druhou nejvýznamnější.

Zde se zároveň dostáváme k limitům dat poskytovaných MDA. Z rysů, kterými lze charakterizovat OČ, do modelu vstupují pouze ty, které splňují následující vlastnosti:

- snadná a spolehlivá operacionalizovatelnost;
- frekvence – aby byly rysy použitelné, musí se vyskytovat v určitém množství v textech;
- disperze – aby byly rysy použitelné, musí se vyskytovat v určitém množství textů v korpusu.

Celá řada rysů typických pro OČ jednu z těchto podmínek nesplňuje, např. protetické *v-*, přehláska *ú > ou*, zkrácené préteritní formy typu *upad*, což jsou rysy, které byly v iniciálních fázích vytváření MDA modelu rovněž zvažovány. Jakkoli je tedy problematické ztotožnit pozici textu na dimenzi 2 s mírou OČ prvků, lze konstatovat, že prostředky OČ, které se v ní odrážejí s vysokým loadingem, představují emblematické znaky OČ: zejména diftongizace *y > ej* ve finální i nefinální pozici a úžení *é > í* v koncovce i v kořeni.

Z výsledků, které postihují stav jazyka zhruba po roce 2000, lze odvodit i přechodovou roli beletrie, která se částečně objevuje v území mluvené interaktivní komunikace (viz obr. 3).



Obrázek 3: Pozice beletristických textů v rámci 1. a 2. dimenze (černé trojúhelníky, ostatní texty jsou naznačeny šedými tečkami)

V této souvislosti se musíme vrátit jak k Bermelovu konceptu spontánně ustupující diglosie, tak ke Sgallovu upozornění na klíčovou roli beletrie a zejména překladů na překlenutí rozdílu mezi SČ a běžně mluvenou češtinou (viz 1.1). Z výsledků zde neprezentovaných víme, že současná beletrie nevykazuje výraznější odlišnosti v pozici na druhé dimenzi v závislosti na tom, zda se

jedná o originální český text nebo překlad z cizího jazyka. Můžeme nicméně pozorovat vztlínání některých beletristických textů v obr. 3, které by mohlo být signálem částečného ústupu české diglosie, započatého už v 60. letech. Můžeme se jenom dohadovat, zda by tento ústup nebyl výraznější, pokud by podmínky nebyly jiné. Zejména pokud by kodifikační praxe nebyla z převážně politických důvodů zakonzervována po celé období komunistické totality (viz Cvrček 2006b) a pokud by po pádu totalitního režimu nebyla paralyzována negativním ohlasem *Pravidel českého pravopisu* z roku 1993, což vyústilo ve společenské klima, které nepřeje jakémukoli otevření kodifikačního standardu prvkům OČ.

Z hlediska registrů jsou pro nás podstatné především dva dynamické rejstříky: konverzace (tj. dynamický spontánní registr) a anketa (tj. dynamický neadresný registr). Konverzace se v datech Koditexu uplatňuje z více než 70 % u spontánních interaktivních a neformálních textů, v menší míře pak u textů veřejně vysílaných (TV/rozhlasové talk-show, 19 %) či u textů elicitovaných (5 %), marginálně se uplatňuje na webových fórech a v soukromé korespondenci. Anketa je hlavně využívána u elicitovaných, tedy poloformálních mluvených textů (56 %), vysílaných konverzací (41 %), minoritně pak ve webových fórech a na Facebooku.

Z těchto údajů můžeme vysoudit, že registry s výrazným postavením na pozitivním pólu druhé dimenze, a tedy i s větším zastoupením OČ prvků, se poměrně vyhraněně uplatňují v určitém typu komunikačních situací. Jejich typickými rysy by kromě nepřipravenosti mohla být *hic et nunc* produkce, absence sociálních bariér či potřeba navazovat a udržovat kontakt. (Srov. i kap. 5.)

1.4 Závěr

Shrneme-li pozorování a závěry uvedené výše, je zjevné, že nacházíme argumenty pro i proti Sgallově interpretaci české situace, od níž jsme v úvodu vyšli.

Analýza založená na registrovém modelu MDA poukazuje především na koexistenci dvou variet, jež se uplatňují v disjunktních funkcích (zejména v mluvených interaktivních situacích). Tato odlišnost je nadto potvrzena osamostatněním druhé dimenze, která separuje mluvené texty od ostatních dynamických textů (jako je narace či komentář), což je úkaz specifický pro češtinu.

Naproti tomu slabinou přístupu je fakt, že do modelu vstupovaly jenom tři (zato nejvýraznější) rysy OČ. To není pouze limit výsledků MDA, ale i symptom jazykové situace v češtině – ostatní rysy OČ do modelu nemohly být zařazeny, protože jejich uplatnění je omezené (co do frekvence i disperze). To znovu ukazuje na mnohokrát zmiňovaný aspekt, že rozdíl mezi SČ a OČ není příliš velký, což oslabuje hodnocení situace jako diglosní (vedle jiných aspektů, jako je relativně neproblematická akvizice pasivní znalosti SČ dětmi apod.).

V neposlední řadě lze výsledky interpretovat i s ohledem na probíhající jazykovou změnu, která přispívá k hodnocení situace jako ustupující diglosie; ohlasy tohoto trendu jsou na výsledcích viditelné zejména na pozici beletristických textů, které adoptují rysy mluveného jazyka. Je přitom možné spekulovat, že větší rozšíření rysů OČ do jinak psaných sfér komunikace brání paralyza české kodifikační praxe, která není schopna ustoupit z pozic vyměřených na začátku 90. let minulého století.