

Univerzita Karlova  
Filozofická fakulta  
Ústav českého jazyka a teorie komunikace  
Český jazyk

Disertační práce

Zuzana Laubeová

Mluvenost v dialogické elektronické komunikaci  
Orality in Dialogic Computer-Mediated Communication

školitel: prof. PhDr. Petr Mareš, CSc.

2020

## **Poděkování**

Mé poděkování patří všem, kteří mě podporovali během celého doktorského studia. Především se jedná o mou rodinu, blízké přátele a kolegy z Ústavu Českého národního korpusu. Není třeba je jmenovat, protože poděkování bylo nebo bude součástí našich neformálních setkání. Na tomto místě bych však chtěla jmenovitě poděkovat svému školiteli, prof. PhDr. Petru Marešovi, CSc., za cenné rady a připomínky, které věnoval různým verzím této práce.

## **Prohlášení**

*Prohlašuji, že jsem disertační práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.*

Praha, 29. 3. 2020

Mgr. & Mgr. Zuzana Laubeová

## **Abstrakt**

Tato disertační práce se zabývá problematikou mluvenosti v neformální psané elektronické komunikaci dialogického charakteru. Za zástupce tohoto typu elektronické komunikace byly vybrány internetové diskuse, diskusní fóra a facebookové statusy, které jsou porovnávány s neformálními mluvenými dialogy realizovanými v bezprostředním kontaktu účastníků. Materiálově se práce opírá o korpus Koditex, jehož součástí jsou i tři výše zmíněné žánry, a mluvený korpus Ortofon.

Termín mluvenost může odkazovat jak k realizované komunikační formě – mluvené, či psané (tzv. mediální mluvenost), tak k souhrnu jazykových charakteristik spojených s prototypickým neformálním mluveným dialogem (tzv. konceptuální mluvenost). Předmětem výzkumu je nicméně pouze konceptuální mluvenost, která je pro účely této práce dále zúžena na devět jazykových rysů rozdělených do tří oblastí podle toho, zda jsou předchozími výzkumy považovány za charakteristické projevy dialogičnosti, spontánnosti, nebo snahy o naplnění pragmatických potřeb účastníků komunikace.

Jednotlivé analýzy ukazují výraznou podobnost ve funkčním využití rysů v mluvené i psané komunikaci. Týká se to především rysů spojených se spontánností (opakování stejných slov, přítomnost slovních fragmentů, překlepů a přeráznutí) a naplňováním pragmatických potřeb (diskurzní markery, pronoun non-dropping, iterace grafémů). Z toho lze vyvodit, že vybrané žánry elektronické komunikace vhodně reprezentují neformální elektronickou komunikaci. Rysy představující dialogičnost (otázky, zájmena pro 2. osobu, nominální oslovení) však zároveň poukazují na to, že jsou srovnávány texty s jiným nastavením podmínek produkce. Předmětem srovnání byl dále smích, reprezentovaný v psané elektronické komunikaci emotikony :-) a :-D. Ty jsou využívány nejen jako vizuální projevy pisatelova smíchu, ale především jako indikátory toho, jaká místa textu považuje pisatel za hodna zasmání.

## **Klíčová slova**

elektronická komunikace, mluvený jazyk, neformálnost, spontánnost, dialogičnost, jazykový korpus

## **Abstract**

The subject of this thesis is orality in informal dialogic computer-mediated communication (henceforth CMC). This type of communication is represented by internet discussions, discussion fora, and Facebook posts. All these three genres are compared with informal spoken face-to-face dialogues. The data are covered by the Koditex corpus, which contains the CMC genres, and the Ortofon corpus of spoken language.

The term orality can refer both to the realized communication form – spoken or written (so-called media orality) and to the summary of linguistic characteristics associated with the prototypical informal spoken dialogue (so-called conceptual orality). This thesis focuses only on conceptual orality, which is based on nine language features divided into the three following areas: dialogicity, spontaneity, and fulfilment of pragmatic needs of communication participants.

The analyses show significant similarity between the functional use of features in both spoken and written communication. This applies primarily to the features associated with spontaneity (repetition of the same words, the presence of word fragments, typos and misspellings) and the fulfilment of pragmatic needs (discourse markers, pronoun non-dropping, iteration of graphemes). It follows that the selected CMC genres appropriately represent informal CMC. However, the features of dialogicity (questions, 2nd person pronouns, nominal addressing) highlight that the comparison is conducted on texts with different settings of production conditions. Another subject of comparison was laughter, represented in written CMC by the emoticons :-) and :-D. They are used not only as visual expressions of the writer's laughter, but above all as indicators of the places where the addressee is intended to laugh.

## **Key words**

computer-mediated communication, spoken language, informality, spontaneity, dialogicity, language corpus

# Obsah

Abstrakt .....	4
Abstract .....	5
1. Úvod .....	8
1.1 Téma a cíl disertační práce .....	8
1.2 Struktura práce .....	11
2. Teoretický přehled .....	13
2.1 Uvedení do problematiky .....	13
2.1.2 Jazykový modus (mluvený vs. psaný) .....	13
2.1.3 Mluvenost vs. psanost (podle Kocha a Oesterreichera) .....	14
2.1.4 Komunikační sféra běžného dorozumívání .....	17
2.1.5 Text .....	19
2.1.6 Elektronická komunikace .....	20
2.2 Klíčové aspekty výzkumu mluveného jazyka .....	21
2.2.1 Charakteristiky podle jazykových rovin .....	22
2.2.2 Směry a přístupy zabývající se mluveným jazykem .....	26
2.2.3 Vybrané publikace o mluveném jazyce .....	34
2.3 Klíčové aspekty výzkumu psané elektronické komunikace .....	38
2.3.1 Charakteristiky podle jazykových rovin .....	40
2.3.2 Směry a přístupy zabývající se elektronickou komunikací .....	44
3. Analytická část .....	47
3.1 Data a metodologie .....	47
3.1.1 Popis dat .....	47
3.1.2 Metodologie .....	50
3.2 Frekvenční srovnání dat .....	53
3.2.1 Frekvence slovních druhů .....	54
3.2.2 Frekvence lemmat .....	55
3.3 Rysy spjaté s dialogičností .....	56
3.3.1 Otázky .....	57
3.3.2 Zájmena pro 2. osobu .....	62
3.3.3 Nominální oslovení .....	65
3.3.4 Shrnutí kapitoly .....	71
3.4 Rysy spjaté se spontánností .....	71
3.4.1 Opakování slov .....	72
3.4.2 Fragменты slov, překlapy a přeroknutí .....	80

3.4.3	Shrnutí kapitoly .....	88
3.5	Rysy související s naplněním pragmatických potřeb .....	89
3.5.1	Diskurzívní markery .....	89
3.5.2	Pronoun non-dropping.....	108
3.5.3	Iterace grafémů.....	120
3.5.4	Shrnutí kapitoly .....	126
3.6	Neverbální komunikace na příkladu smíchu .....	126
3.6.1	Smích v korpusu Ortofon .....	128
3.6.2	Emotikony .....	135
3.6.3	Shrnutí kapitoly .....	145
4.	Závěr.....	147
	Seznam použité literatury .....	152

# 1. Úvod

## 1.1 Téma a cíl disertační práce

Vliv internetu a elektronických komunikačních nástrojů na jazykovou komunikaci patří k oblíbeným výzkumným tématům již přes třicet let (v českém prostředí samozřejmě kratší dobu), přesto se nedá říci, že by se jednalo o vyčerpané téma. Naopak zvyšování všeobecné počítačové gramotnosti a čím dál častější užívání internetu v běžné každodenní komunikaci přináší nové možnosti zkoumání.

Termínem elektronická komunikace se obecně označuje jakákoli lidská komunikace zprostředkovaná elektronickými přístroji a přenášená nejčastěji internetem. Důležité však je konkrétní nastavení této komunikace, tedy to, zda se jedná o webovou stránku, na níž si jednotliví uživatelé vzájemně píšou nebo spolu diskutují, zda text na stránce tvoří pouze jeden uživatel jako vlastní článek, zda a (do)kdy lze na texty jiných uživatelů reagovat apod. Před vlastním výzkumem je tedy nutné vybrat si konkrétní žánry elektronické komunikace, což platí samozřejmě i pro výzkum neelektronické komunikace, a zároveň ověřit jejich dostupnost pro následný výzkum. V souladu s tímto hlediskem byl v češtině zkoumán především jazyk e-mailů (např. Čmejrková, 1997), internetového chatu (Jandová et al., 2006), blogů (např. Blinka et al., 2012), internetových diskusí (např. Vepřek, 2018); jazykové srovnání několika žánrů elektronické komunikace (blogů, internetové encyklopedie Wikipedie, internetových diskusí, diskusních fór, veřejných facebookových statusů) pak nově podávají Cvrček et al. (2018b, v tisku). Tento stručný výčet mj. ukazuje posun v uvažování o elektronické komunikaci: starší práce se spíše zaměřují na popis jazykových zvláštností těchto „nových“ žánrů, aktuální studie jsou buď zaměřeny na konkrétní fenomén, nebo mají komparativní charakter (ať už s ohledem na různé typy textů či jazykových prostředků).

Tato práce navazuje na oba přístupy. Zaměřujeme se na analýzu tří vybraných žánrů psané elektronické komunikace (internetových diskusí, diskusních fór, veřejných facebookových statusů) a porovnáváme je s mluvenou komunikací, přičemž vycházíme z předpokladu, že s rozdílným jazykovým modelem či komunikační platformou může dojít ke změně v užívání jazykových prostředků. Právě otázka, zda k něčemu takovému dochází a zda se dá u těchto změn vysledovat podobnost mezi zkoumanými texty, je dalším bodem našeho výzkumu. V této práci zcela opomíjíme neelektronickou psanou komunikaci, např. beletrii, žurnalistiku, soukromé dopisy, přestože jsme si vědomi toho, že by pro úplnost měla být zahrnuta do výzkumu. Není však v našich silách pokrýt tuto oblast ve stejné míře jako mluvenou a psanou



elektronickou komunikaci, proto zájemce o srovnání s neelektronickým psaným jazykem odkazujeme na jiné studie, např. již zmiňovanou kolektivní práci Cvrčka et al. (2018b, v tisku).

Zde zkoumané jazykové projevy jsou charakterizovány neformálností, neoficiálností, spontánností, dialogičností, interaktivností. Dále v textu používáme pro jejich označení pouze první z pojmů. Z hlediska žánrů se tedy budeme zabývat analýzou internetových diskusí, diskusních fór, veřejných statusů na Facebooku a mluvenými dialogy realizovanými mezi mluvčími, kteří se dobře znají. Ačkoli žádný ze zmiňovaných elektronických žánrů není přímým ekvivalentem takto vymezené mluvené komunikace (tím by byl spíše chat), je pro ně charakteristika pomocí pěti výše uvedených pojmů platná (podle výzkumu Cvrčka et al., 2018b, se přiklání k označení dynamických a spontánních textů na rozdíl od textů statických a připravených). Dalším důvodem pro výběr právě těchto elektronických žánrů je jejich dostupnost ve formě veřejného jazykového korpusu (pro češtinu neexistuje žádný korpus chatu).

Úmysl využít pro tuto práci jazykové korpusy byl motivován pragmatickým rozhodnutím přenechat problematiku sběru dat odborníkům. Nutno poznamenat, že zatímco pro analýzu neformálního mluveného jazyka bylo na začátku této práce možné uvažovat o několika mluvených korpusech, z nichž byl nakonec vybrán Ortofon (Kopřivová et al., 2017), neplatilo totéž pro elektronickou komunikaci. V úvahu přicházely buď webové korpusy shromažďující veškeré texty z českých stránek (tj. z domény cz), a to z korpusové rodiny Aranea (*Araneum Bohemicum*, Benko, 2015) či TenTen (*csTenTen*, Suchomel, 2012), anebo webová část korpusu Koditex (Zasina et al., 2018). Výhodou a zároveň nevýhodou dvou prvně zmíněných korpusů je jejich velikost, která se pohybuje v řádech stamilionů až miliard slov,<sup>1</sup> což ale není určující kritérium při výzkumu kombinujícím kvantitativní a kvalitativní metody. Pro tuto práci bylo důležitější najít takový korpus, který by umožňoval pracovat s předem vymezenými žánry ve smyslu jejich dostupného označení na základě webových stránek, z nichž texty pocházejí, a tímto způsobem se dostat k neformální psané elektronické komunikaci.<sup>2</sup> Ani jeden z korpusů Aranea či TenTen tyto podmínky nespĺňuje, proto jsme zvolili korpus Koditex, v němž je uvedeno označení žánrů elektronické komunikace. Jako nevýhodu či omezení u něj lze označit menší velikost – statisíce slov – a důraz na individualitu pisatele, která se projevuje právě

---

<sup>1</sup> Pro korpus *Araneum Bohemicum* jsou dostupné tři verze v různé velikosti: *maximum*, obsahující co největší rozsah dostupných textů, *maius* o velikosti cca 1 miliarda slov a *minus*, který je desetiprocentním vzorkem z verze *maius*.

<sup>2</sup> V době odevzdání této práce byl již zveřejněn korpus NET (Jeziorský, 2019), složený z textů diskusních fór a osobních blogů.

u interaktivních žánrů ve sdružení všech příspěvků téhož pisatele dohromady (viz Cvrček et al., v tisku).

Jakkoli se z výše uvedeného může zdát, že více pozornosti věnujeme elektronické komunikaci než mluvenému jazyku, opak je pravdou. Právě mluvená komunikace je základem, z něhož při srovnání vycházíme a k němuž se vracíme. Tento postup volíme v souladu s pojmem konceptuální mluvenosti (Koch & Oesterreicher, 1985), který v sobě zahrnuje situační a jazykové charakteristiky mluveného jazyka užívaného v běžné, bezprostřední, spontánní komunikaci blízkých osob v neformálním prostředí. Vybrané žánry elektronické komunikace by se daly touto terminologií charakterizovat jako konceptuálně mluvené a mediálně psané (Dürscheid, 2003). Cílem komparace je tedy snaha o zjištění shod a rozdílů mezi mediálně odlišnými, ale konceptuálně blízkými typy textů.

Vlastní výzkum je realizován jako analýza několika vybraných jazykových rysů považovaných za projevy spontánnosti a dialogičnosti. Mimoto analyzujeme rysy, které se nedají jednoznačně přisoudit žádnému z výše uvedených pojmů, protože souvisejí s tím, jakým způsobem si mluvčí přizpůsobuje komunikaci s ohledem na její průběh, obsah či adresáta sdělení (v textu je označujeme za rysy související s naplněním pragmatických potřeb). Při srovnání se tedy snažíme zjistit, zda je tentýž jazykový rys používán v stejné či podobné funkci a zda se nějak liší frekvenčně. Nezabýváme se tudíž jazykovými rysy specifickými pro určitý žánr nebo jazykový modus, ale snažíme se identifikovat shody a rozdíly u úzce vymezené skupiny rysů spojených s vybranými charakteristikami konceptuální mluvenosti. Dále analyzujeme jeden typ neverbálního chování, který se často objevuje v neformální mluvené komunikaci – smích.<sup>3</sup> Za jeho projev v psané elektronické komunikaci považujeme emotikony, např. :-) a :-D, a v jejich analýze ukazujeme, jaké další funkce v textu zastávají.

Jak už bylo zmíněno výše, kombinujeme při analýzách kvantitativní a kvalitativní metody. To se projevuje už v samotné volbě jazykových rysů; část z nich je vybrána na základě výsledků multidimenzionální analýzy českých textů (Cvrček et al., 2018a, 2018b, v tisku) a část podle odborné literatury, v obou přitom byly klasifikovány jako projevy neformálnosti, spontánnosti, dialogičnosti i adresnosti. Vlastní analýzy jsou většinou zpracovány na náhodných vzorcích dat, cílem analýz je totiž posouzení funkce jazykového rysu v daném kontextu, což je provedeno kvalitativní textovou analýzou. Způsob práce s daty, který je vlastní korpusovým

---

<sup>3</sup> To mj. proto, že lidé se smějí rádi a často.

analýzám, kombinujeme zejména s konverzační analýzou a interakční lingvistikou (srov. Clancy, 2016).

Cíle této práce se dají shrnout do následujících bodů:

- 1) porovnání konceptuálně mluvených textů lišících se mediálně, tj. neformálních mluvených dialogů a neformální dialogické (interaktivní) elektronické komunikace;
- 2) určení, který ze tří zkoumaných žánrů elektronické komunikace (internetové diskuse, diskusní fóra, facebookové statusy) se nejvíce blíží neformálním mluveným dialogům;
- 3) ověření toho, zda jsou vybrané jazykové rysy dobrými indikátory dané charakteristiky, tj. čím je motivován jejich výskyt v textech;
- 4) zjištění, jak je vizuální kanál využit v psané elektronické komunikaci k nápodobě mediální mluvenosti (na příkladu smíchu);
- 5) kvalitativní analýza funkcí vybraných jazykových rysů použitých v multidimenzionální analýze češtiny (MDA).

## 1.2 Struktura práce

Tato práce je rozdělena do dvou větších celků rámovaných úvodem (kap. 1) a závěrem (kap. 4). V první, teoretické části jsou nejprve představeny základní pojmy, jež jsou spojeny s problematikou výzkumu mluvené a psané elektronické komunikace: jedná se o jazykový modus, mluvenost a psanost, komunikační sféru běžného dorozumívání, text a elektronickou komunikaci (2.1). Poté následuje stručné představení dosavadního výzkumu mluveného jazyka (2.2) a elektronické komunikace (2.3). U obou jsou nedříve obecně shrnuty oblasti výzkumu a výzkumné otázky, poté jsou představeny konkrétní jazykové charakteristiky, které jsou pro lepší přehlednost rozčleněny podle jazykových rovin (2.2.1 a 2.3.1). Součástí každé z těchto kapitol je oddíl věnovaný směrům a přístupům užívaným při analýze mluvené i psané elektronické komunikace (2.2.2 a 2.3.2). V kapitole týkající se mluveného jazyka jsou ještě představeny tři publikace prezentující mluvený jazyk specificky, mimo rámec psaného jazyka (2.2.3).

Druhá, praktická část disertační práce začíná kapitolou o analyzovaných datech a použitých metodách (3.1). Vlastní výzkum začíná charakteristikou dat podle frekvenčního hlediska, tj. rozborem podle slovních druhů a nejčastějších lemmat (slovníkových tvarů) (3.2). Další analýzy jsou už zaměřeny na konkrétní jazykové rysy sdružené do tří podkapitol podle obecnější charakteristiky, k níž se vztahují: dialogičnost (3.3), spontánnost (3.4), naplňování pragmatických potřeb (3.5). Poslední kapitola analytické části je věnována analýze smíchu jako

příkladu neverbální komunikace (3.6). Podkapitoly v této části a studie věnované konkrétním rysům mají jednotnou strukturu: po krátkém teoretickém uvedení jsou představeny metodologické detaily analýzy, po nich následuje analýza jevu v mluveném korpusu Ortofon a webové části korpusu Koditex věnované internetovým diskusím, diskusním fórům a facebookovým statusům; výsledky analýzy jsou nakonec vzájemně porovnány a podkapitola či studie je zakončena shrnutím.

Poslední kapitola disertační práce usouvztažňuje výsledky jednotlivých analýz a konfrontuje je s odbornou literaturou.

Rádi bychom ještě upozornili na blízkost studie o diskurzních markerech (DM, viz 3.5.1) s kapitolou *Diskurzni markery* (Čermáková et al., 2019) v publikaci *Syntax mluvené češtiny* (Hoffmannová et al., 2019). Pro disertační práci byla využita část výsledků o DM *jo*, u kterého byla hlavní autorkou ve zmiňované kapitole autorka této práce.

### **Formální zásady**

V práci je respektována citační norma APA 7 (užívaná např. v časopise *International Journal of Corpus Linguistics*). Ženská jména jsou v textu souladu s českým územ přechylována, avšak v odkazech a v bibliografii jsou kvůli snazšímu dohledání ponechána v původní, nepřechýlené podobě.

## 2. Teoretický přehled

### 2.1 Uvedení do problematiky

Ústředním předmětem této práce je mluvenost a její projevy jak v mluvené formě jazyka, tak ve formě psané, zprostředkované elektronicky. Samotný pojem mluvenost napovídá příbuznost s mluveným jazykem; z něho by na základě této slovotvorné příbuznosti měly být jednotlivé projevy vyvozeny. Ve výzkumech zabývajících se různými aspekty mluvenosti je primární zdroj, tj. mluvený jazyk, označován různě, např. mluvená forma jazyka, hovorový jazyk, hovorová čeština, běžně mluvený jazyk, mluvený projev, nespisovná čeština, expresivní mluva, neoficiální mluvená komunikace (obsáhleji viz Hoffmannová & Zeman, 2017). Tento stručný výčet ukazuje akcent na různé charakteristiky mluveného jazyka: na způsob realizace/modus/formu a s ním související komunikační kanál, na prostředí, v němž k realizaci dochází, a jeho situační charakteristiku. Důvodem, proč jsme v této práci vybrali termín mluvený jazyk, je především jeho krátkost (oproti např. pojmu mluvená forma jazyka) a explicitní zdůraznění komunikačního kanálu, který je v centru pozornosti této práce.

V následující části se budeme věnovat několika pojmům, které jsou klíčové pro specifikaci mluvenosti v této práci.

#### 2.1.2 Jazykový modus (mluvený vs. psaný)

Tento pojem se vztahuje k existenční formě jazyka. Řečeno pojmoslovím R. Jakobsona (1995 [1960]) se jedná o využití komunikačního kanálu k zprostředkování zprávy mezi mluvčím (vysílatelem) a adresátem (příjímatelem). Podle Jakobsona předpokládá úspěšná komunikace jednak kontakt mezi oběma komunikujícími stranami, jednak užívání téhož kódu, přičemž se u obou předpokládá znalost kontextu nutná ke správnému dekódování zprávy. Kódování zprávy se děje zpravidla vokálně-akusticky nebo graficko-vizuálně, přičemž obě možnosti lze v různé míře kombinovat.<sup>4</sup> Přenos prostřednictvím akustického kanálu je ztotožněn s jazykovým modem mluveným, přenos vizuálním kanálem s jazykovým modem psaným.

V české stylistické tradici je existenční modus jazyka, resp. volba mezi mluveným a psaným kanálem a tedy vokálně-akustickým (a zároveň vokálně-auditivním) či graficko-vizuálním kódem, řazen mezi objektivní stylotvorné faktory.

Vzhledem k inherentní nutnosti přizpůsobit komunikaci vybranému kanálu rozlišují např. Brazil (1995), McCarthy (1998) nebo Miller a Weinert (1998) oba jazykové mody jako

---

<sup>4</sup> Vyjadřování sděleného obsahu vztahuje Grepl (1962: 343) k počtu využívaných sémiologických soustav: v psaných projevech pomocí jedné sémiologické soustavy, kdežto v mluvených kombinaci tří soustav (jazyková, zvuková, parajazyková).

samostatné jazyky včetně vlastní gramatiky. Naopak Vachek (1942) považuje mluvení a psaní za dvě rovnocenné manifestace společného jazykového kódu<sup>5</sup> a každé z nich připisuje komplementární normu.

Pojem komunikační kanál je někdy nahrazován pojmem médium (viz Nöth, 2000: 237), resp. jedním z významů tohoto slova je „prostředek komunikace“, kterým se míní řeč, neverbální komunikace, umění apod. (srov. pojem primární médium, Valček, 2011: 202). Slovník mediální komunikace (Reifová et al., 2004: 140) doplňuje, že média v nejužším slova smyslu jsou „prostředky masové, či mediální komunikace, tj.: a) média tištěná, b) média elektronická, c) média nová,“ která jsou dnes téměř synonymní s počítačovými nebo digitálními médii.

### 2.1.3 Mluvenost vs. psanost (podle Kocha a Oesterreichera)

Tato dvojice pojmů je úzce spjata s existenčním modem jazyka, představuje však širší problematiku determinace komunikačního kódu pro zvolený komunikační kanál. Primárně lze podle Kocha a Oesterreichera oba pojmy chápat mediálně (substanciálně): mluvený text je realizován vokálně-akusticky, psaný graficko-verbálně. Kromě toho mohou být chápány i konceptuálně jako soubory charakteristik typických pro mluvené/psané texty a ovlivněné komunikačním kanálem zvoleným při produkci textu. Tyto soubory distinktivních vlastností a jazykových prostředků se vždy pojí pouze s prototypy mluvenosti a psanosti.<sup>6</sup> Koch a Oesterreicher (1985) dále usouvztažňují konceptuální a mediální mluvenost a psanost s jazykem používaným v komunikaci blízkých osob (*Sprache der Nähe*) a vzdálených, cizích osob (*Sprache der Distanz*).<sup>7</sup>

Příkladem mediálně i konceptuálně mluveného textu by byl rozhovor kamarádek u kávy, mediálně i konceptuálně psaným textem je text zákona. Prezidentský projev vyslovený na Nový rok představuje mediálně mluvený, ale konceptuálně psaný text. Narychlo napsaný vzkaz manželovi či chat dvou přátel na Facebooku jsou příklady mediálně psaných, ale konceptuálně mluvených textů.

V teorii komunikační blízkosti a distance, jak se také o Kochově a Oesterreicherově přístupu referuje, se pracuje především s póly, které opozice mluvený – psaný nabízí. Autoři sami však

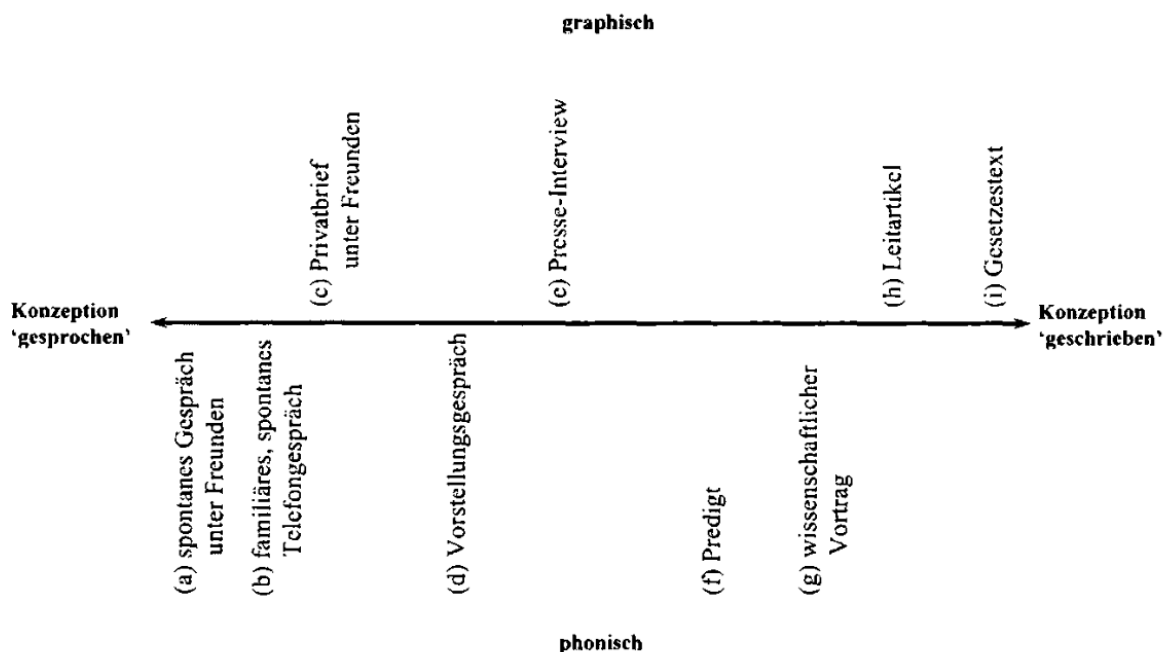
---

<sup>5</sup> „...není soustavy [jazykové – ZL], jež by byla nadřaděna jazyku psanému a mluvenému, nýbrž že sounáležitost těchto dvou soustav je dána prostě tím, že obě slouží jednomu a témuž jazykovému společenství, a to v úkolech, jež se navzájem vzorně doplňují. Vědomí této sounáležitosti vydatně napomáhá okolnost, že mezi oběma soustavami existuje do značné míry korespondence jednotek optických a akustických...“ (Vachek, 1942: 257)

<sup>6</sup> „Der Begriff ‚konzeptionelle Mündlichkeit/Schriftlichkeit‘ zielt also auf Aspekte der sprachlichen Variation, die in der Forschung häufig unscharf als ‚Umgangssprache/Schriftsprache‘, ‚informell/formell‘, ‚Grade der Elaboriertheit‘ usw. erfaßt werden.“ (Koch & Oesterreicher, 1994: 587)

<sup>7</sup> Označení *Nähe/Distanz* se jim jeví jako vhodnější, protože neasociuje pojem médium, jež může být chápán různě (Koch & Oesterreicher, 1994: 588).

otevřeně říkají, že většina textů se k těmto pólům spíše přibližuje, a umísťují je na pomyslnou osu (škálu) mezi nimi (obr. 1).



Obr. 1. Příklad zařazení textů mezi póly mluvenost – psanost podle Kocha a Oesterreichera (převzato z Koch & Oesterreicher, 2007: 349)

Teorie doznala řadu obměn a doplnění svými tvůrci (např. Koch & Oesterreicher, 1994, 2007, 2011), ale i následovníky.<sup>8</sup> Jednou z nich je Ch. Dürscheidová (2003), která teorii aplikovala na elektronickou komunikaci. Upozornila zejména na rozdíl mezi synchronní a asynchronní elektronickou komunikací a upřesnila škálu jednotlivých možností na konceptuální rovině. Mediálně mluvená synchronní komunikace (např. telefonní rozhovor) je vždy bližší ideálnímu pólu konceptuální mluvenosti než asynchronní mluvená komunikace (např. odpověď na vzkaz na záznamníku).<sup>9</sup> U mediální psanosti operuje autorka s pojmem kvazisynchronní komunikace (např. chat na internetu), která je blízká asynchronní mluvené komunikaci, ale přesto vzdálenější od ideálního pólu konceptuální mluvenosti než ona. Nejdále od tohoto pólu stojí mediálně psaná asynchronní komunikace (např. dopis).

### 2.1.3.1 Konceptuální mluvenost

V teorie komunikační blízkosti a distance P. Kocha a W. Oesterreichera (1985) je konceptuální mluvenost spojena s následující řadou proměnných: dialogičnost, neomezené střídání mluvčích v řeči, důvěrný vztah mezi mluvčími, interakce tváří v tvář, neomezený rozvoj témat,

<sup>8</sup> K recepci modelu viz Dürscheidová (2016).

<sup>9</sup> Baronová (2010: 14) považuje pouze telefonní rozhovor a rozhovor z očí do očí za případy skutečné synchronní komunikace.

neveřejnost,<sup>10</sup> spontánnost, silné účastenství (tj. silné zapojení mluvčích do interakce), situační zakotvenost.<sup>11</sup> Konceptuální mluvenost neboli jazyk blízkosti je pojmem odkazujícím k charakteristikám mluvené komunikace odehrávající se každodenně mezi lidmi s blízkými vztahy (partnery, rodinnými příslušníky, přáteli) v neformálním, soukromém prostředí, v němž je větší pozornost věnována významům než slovům. Koch a Oesterreicher (1985) se v první koncepci příliš nezabývají charakteristikou jazykových prostředků, jde jim o postižení vnějších komunikačních podmínek. Podobně, skrze soubor vnějších situačních proměnných, definuje Čermák (2009) tzv. prototypický mluvený jazyk. Charakteristická je pro něj přítomnost následujících rysů: *+spoken, +dialogue, +proximity, +equality, +private, +informal, +interactive, +present, +non-multiple, +spontaneous, +casual, +not aware.*

Hoffmannová (2017b) připojuje k již zmiňovaným proměnným další vnějazykové (nepřipravenost, neformálnost<sup>12</sup>, dynamičnost, subjektivnost, emocionální zaujetí), ale zmiňuje i takové rysy, které se vztahují přímo k jazykovým prostředkům (značná konkrétnost, nízký stupeň organizovanosti a uspořádanosti, fragmentárnost, implicitnost, neurčitost).

### 2.1.3.2 Konceptuální psanost

Proměnné, které podle Kocha a Oesterreichera charakterizují konceptuální psanost, jsou vytvořeny v opozici k proměnným charakterizujícím konceptuální mluvenost. Jedná se o monologičnost, žádné střídání mluvčích v řeči, vztah cizosti mezi mluvčími, časová a prostorová oddělenost, fixní téma, veřejnost, reflektivnost, malé účastenství, vyvázanost ze situace. Hoffmannová (2017b) dodává připravenost, větší formálnost, složitější tematiku, dekontextualizovanost, statičnost, spíše abstraktní, racionální povahu, objektivnost, analytický,

---

<sup>10</sup> Neveřejnost lze ztotožnit s pojmem soukromost (privátnost). Jejich opak – veřejnost (srov. Mareš, 2003) můžeme vztahovat k různým aspektům komunikace: k prostředí, v němž probíhá, ke vztahům mezi jejími účastníky, k přístupu dalších (potenciálních) účastníků. Tyto aspekty jsou vzájemně propojeny: představíme-li si např. narozeninovou oslavu pořádanou doma u oslavence, bude probíhat v soukromém prostředí, kam mají přístup pouze pozvaní hosté; jejich pozvání je implikováno blízkým vztahem k oslavenci. Oslava probíhající v restauraci oproti tomu probíhá ve veřejném prostředí, do kterého mají přístup i nepozvaní hosté a osoby bez blízkého vztahu k oslavenci, kteří se teoreticky k oslavě mohou připojit. Zda tak (ne)učiní, ovlivňují společenské normy a míra, s jakou se s nimi tyto nepozvaní hosté identifikují. Neveřejnost tedy chápeme jako komplexní charakteristiku tří výše zmíněných položek, které spolu kooperují.

<sup>11</sup> Citováno dle Kaderky (2013: 79).

<sup>12</sup> Ve shodě s Marešem (2003) odlišujeme pojem neformálnost od neoficiálnosti. Formálnost tento autor definuje jako zachovávání ustálené, náležité, očekávané podoby řečového chování, ustálených řečových konvencí a textových modelů, udržování distance mezi komunikanty a vzájemného respektu. Neformálnost se tedy projevuje jako nepředpisovost, nsvázanost pravidly, uvolněnost, nenucenost vyjadřování. Také je spojena se spontánním chováním, které navozuje blízkost a důvěrnost, ale může zahrnovat i aspekt neucty ke komunikačnímu partnerovi či lhotejnost. Oficiálnost ztotožňuje Mareš s úředností. Oficiálnost je dána tím, zda je účastník komunikace představitelem určité sociální organizace; jeho příslušná sociální role mu jako autorovi nebo příjemci textu dodává závažné postavení.



logický přístup, vysokou organizovanost, uspořádanost, integrovanost, kontinuálnost, explicitnost a určitost.

Koch a Oesterreicher (1994) také poukazují na existenci frekvenčních rozdílů mezi stejnými jazykovými prostředky v mluveném a psaném jazyce, např. pasivum nebo spojky vyjadřující specifické vztahy se vyskytují častěji v psaném jazyce než v mluveném.

#### **2.1.4 Komunikační sféra běžného dorozumívání**

Jednou ze základních determinant používání jazyka je mimojazykový kontext, tj. charakteristika kdy, kde, jak, s kým a proč jazyková interakce probíhá. Uvědoměním si konstelací těchto vztahů se dostáváme k zobecnění na rovině tzv. komunikačních sfér,<sup>13</sup> definovaných jako „soubor (typů) komunikačních situací a textů příznačných pro určitou sociální sféru“ (Mareš, 2016: 15). Sociální sférou se pak míní soubor verbálních i neverbálních interakcí, charakterizovaných aktéry, jejich vztahy, činnostmi a cíli, tedy přesahem komunikační sféry do sociálního života člověka. Příkladem sociální sféry může být výchova a vzdělání, jejíž komunikační sféru bychom označili za školní komunikaci (viz *Stylistika mluvené a psané češtiny*, Hoffmannová et al., 2016). Pro konceptuální mluvenost je však nejvíce relevantní komunikační sféra označovaná v nejnovější české stylistice (Hoffmannová et al., 2016) za běžnou každodenní komunikaci. Jazykové projevy v této sféře slouží k „prostému sdělování nebo obcování“ (Homoláč & Mrázková, 2014: 5). Cílem běžného dorozumívání je podání či získání informace, výměna názorů a vyjadřování postojů, předávání pokynů, navazování a udržování (mezilidského) kontaktu.<sup>14</sup>

Prototypickým příkladem běžného dorozumívání je mluvený dialog probíhající za fyzické účasti všech komunikantů na tomtéž místě v tomtéž čase, např. rodinný hovor u večeře.<sup>15</sup> Běžná každodenní komunikace však probíhá i psanou formou; *Stylistika mluvené a psané češtiny* zde zmiňuje ručně psané dopisy a elektronické texty e-mail, SMS apod.<sup>16</sup> Do běžného dorozumívání patří podle Hoffmannové a Jílkové (2016) i např. komunikace při nakupování (jídla, jízdenek), dotazování se náhodného chodce na směr cesty, rozhovor nadřízeného s podřízeným, tedy situace vykazující znaky (polo)formálnosti, (polo)oficiálnosti, (polo)připravenosti apod.

---

<sup>13</sup> Můžeme se setkat také s jiným pojmenováním, např. komunikační oblast, komunikační/slohový styl a – zejména v anglosaské literatuře – diskurs.

<sup>14</sup> V teorii funkčních stylů se běžné dorozumívání kryje s prostě sdělovacím stylem (Havránek, 1932). Podrobný přehled o chápání běžného dorozumívání v české a slovenské stylistice viz Hoffmannová a Jílková (2016: 22n.).

<sup>15</sup> Běžné dorozumívání dominuje sociální sféře rodiny a trávení volného času (Homoláč – Mrázková, 2014: 8).

<sup>16</sup> K typologii mluvených textů ve sféře běžného dorozumívání viz Hoffmannová a Jílková (2016: 40).

Určení hranic dané komunikační sféry je značně obtížné. Ačkoli se touto problematikou podrobněji zabývá např. sociologie<sup>17</sup> a její definování není cílem této práce, zmíníme několik skutečností. Běžné dorozumívání bývá vymezováno pomocí kvalit sdělované informace, např. každodennost,<sup>18</sup> běžnost, neodbornost, nespecializovanost, nepřipravenost, pomocí charakteristiky komunikantů s přihlédnutím k jejich vztahu, sociální pozici apod., charakteristiky činnosti, které běžné dorozumívání doprovázejí, a také pomocí charakteristiky komunikačního prostředí, např. neformální, neoficiální, soukromé. Abychom se při popisu jazyka užívaného v této sféře vyhnuli definici kruhem, měli bychom usouvztažnit konkrétní jazykové prostředky s konkrétní konstelací všech složek komunikační situace. I tak budeme postaveni před řadu rozhodnutí, např. jaké prostředky vybrat, jak určit jejich „prototypičnost pro situaci“, zda souvškyt určitých prostředků ještě patří do oné situace, zda s našimi kritérii souhlasí i jiní, zda existuje mezi kritérii nějaká hierarchie atd. Tímto postupem se dostaneme k tomu, že krajní póly (nebo centrum) prototypičnosti schopni určit jsme – s určitou mírou neshody mezi posuzovateli –, ale vždy bude existovat široké přechodové pásmo, do něhož budou patřit jak prostředky neprototypické pro danou situaci či neutrální vůči našim kritériím, tak prostřední část škály, na níž jsou prostředky umístěny subjektivně či vágně. Určit hranice běžného dorozumívání se do jisté míry pokusili Koch a Oesterreicher (1985) teorií o komunikační blízkosti a distance (viz pojem konceptuální mlúvenost v 2.1.3.1).

Pronikání jazykových prostředků z jedné komunikační sféry do druhé je jevem žádoucím, jedná se ostatně o důkaz jazykového vývoje. Tyto přechody úzce souvisejí s pojmy demokratizace, konverzacionalizace či kolokvializace. Všechny tři pojmy odkazují k vývojové tendenci sblížující jazykové kódy užívané v rámci jednoho jazyka v různých komunikačních sférách a textech. Tato tendence je opakem tzv. intelektualizace, snahy vyjadřovat se o složitějších jevech a souvislostech specializovaně (srov. Havránek, 1932; Haller, 1933; Daneš, 1988). Pojem demokratizace chápeme vůči ostatním zmíněným jako nadřazený, zbylé dva jsou jejím specifickým projevem (Fairclough, 1992).

---

<sup>17</sup> Touto problematikou z řad českých sociologů zabývali např. K. Kosík (1966) a M. Petrusek (1992, 2017).

<sup>18</sup> Každodennost zde vztahujeme k takovému jednání a konání, které se pravidelně opakuje (v rozporu se slovtvorným základem pojmu se nemusí odehrávat každý den, ale např. každou sobotu) a po nějaké době může být rutinizováno. Jeho aktéři mu nemusí přisuzovat zvláštní význam, naopak významu nabývá takové jednání či konání, které každodennost naruší, změní. Petrusek (2017) dále specifikuje, že vykonávání těchto činností předpokládá určitý objem vědění fixovaný v tzv. běžném vědomí a vyjadřovaný běžným jazykem. Kosík (1966) spojuje každodennost s důvěrností, familiárností, obeznameností, „domovem“.

Konceptu každodennosti se věnuje např. sociologie, historie, zejména kulturní historie a jí příbuzná historická antropologie. Předmětem zájmu v historii je jedinec a to, jakým způsobem vnímá své postavení a určitou dobu, co ovlivňuje jeho život a čím je obklopen.

Koncept konverzacionalizace byl podle McMahona (1994) vytvořen analogicky ke konceptu gramatikalizace, jímž se míní proces získávání (a ustavování) gramatické funkce lexikální jednotky v určitých kontextech (Hansen, 2017). Konverzacionalizace pak označuje proces pronikání jazykových prostředků typických pro běžnou každodenní, tj. neformální, spontánní, nepřipravenou, komunikaci do takové komunikační sféry, v níž se tyto prostředky vyskytují minimálně nebo vůbec, a proto se stávají nápadnými. O konverzacionalizaci se často mluví v souvislosti s hromadnými sdělovacími prostředky (televizí, rozhlasem, tiskem)<sup>19</sup> a jejich snahou přiblížit se co nejvíce adresátovi a navodit dojem neformálnosti, neoficiálnosti a příjemné atmosféry panující při osobním kontaktu, např. užíváním neformálních pozdravů nebo nespisovných koncovek (Hoffmannová, 2017a). Kolokvializace odkazuje ke střídání a míšení jazykových kódů, zejména spisovného a nespisovného (Hoffmannová, 2008).

Sféra běžného dorozumívání je z hlediska vývojových tendencí nejprogresivnější. Právě v situacích, při nichž ustupuje do pozadí nutnost přesného a jasného vyjádření vznikají nové formy, ať už se jedná o neologismy, jazykovou hru či výpůjčku z cizího jazyka. Tato komunikační sféra však má i své konvence a ustálené vzorce vyjadřování, které jsou ovlivňovány danou skupinou a její příslušností ke kultuře, danou situací a její žánrovou charakteristikou a jazykem samým.

### 2.1.5 Text

Výsledek komunikačního procesu bývá označován různě, nejčastěji se objevují pojmy text, promluva, jazykový projev či komunikát. Pro různá pojetí a jejich distinkci bývá brána zpravidla v úvahu komunikační formu (tj. zda se jedná o výsledek psaní, či mluvení) nebo sémiotické soustavy využití v komunikaci (tj. např. audiovizuální soustava při mluveném dialogu). Pojem s nejdelsí tradicí, která zapřičiňuje i jeho značné rozrůznění, je text (viz např. Adamzik, 2016; Nekula, 2017). Nevýhodou tohoto pojmu je jeho chápání v obecném úzu, v němž je spojen s písmem a psaným jazykem.<sup>20</sup> V případě mluvené komunikace je text předmětem diskuse, nabízí se totiž různé možnosti, co za něj považovat, např. repliky<sup>21</sup> pouze jednoho mluvčího, vzájemně na sebe navazující repliky obou mluvčích. Např. Hausenblas (1984) nepovažuje dialog za text, protože v něm není dodržena jednota subjektu, smyslu a stylu.

---

<sup>19</sup> Konverzacionalizaci veřejných médií názorně demonstruje Čmejrková (2008).

<sup>20</sup> Např. *Slovník spisovného jazyka českého* (Havránek et al., 1960–71) definuje text jako „tištěný nebo psaný záznam jazykového projevu“ zpravidla delší než jedna věta.

<sup>21</sup> Replikou rozumíme nepřetržitý úsek řeči jednoho mluvčího (Nekvapil & Havlík, 2017).

Jiné pojetí textu nabízí Adamziková (2016). Usouvztažňuje ho s pojmem komunikát, který chápe jako nadřazený pro jakoukoli sumu signálů vyskytujících se v komunikaci.<sup>22</sup> Komunikáty dělí podle počtu způsobů přenosu významů na monomediální a multimediální. Monomediální komunikát využívající jazykové prostředky se nazývá text (jiným příkladem je obraz). V případě multimediálního komunikátu se textem označuje pouze jazyková složka, např. u mluveného dialogu probíhající mezi jeho účastníky v tváři tvář je jazykové vyjádření (tj. text) pouze jednou složkou, dalšími jsou např. gesta, mimika, prozódie. Ačkoli si jsme vědomi toho, že toto pojetí textu a komunikátu lze dále problematizovat, pro účely této práce zůstaneme v chápání obou pojmů tak, jak jsme zde představili. Pro výsledek komunikačního procesu probíhajícího jak v mluveném, tak v psaném modu budeme užívat pojem text (s odlišujícím adjektivem, aby byl zřejmý použitý modus).

### 2.1.6 Elektronická komunikace

Tímto termínem bývá označována veškerá komunikace zprostředkovaná elektronickým zařízením, např. počítačem, mobilem, tabletem (Kopecký, 2007). Někdy bývá ztotožňována s internetovou komunikací, ta je však pouze jejím podtypem, neboť ne všechny elektronické texty využívají internet (např. SMS).<sup>23</sup> Internet je příkladem telekomunikační sítě, která je dnes pravděpodobně nejvyužívanější a která postupně přebírá funkce jiných sítí, např. telefonní.

V rámci elektronické komunikace bývají rozlišovány žánry<sup>24</sup> na základě vnětového vymezení elektronického prostředí a jeho možností, např. blog, chat, instant messaging (dále jen IM),<sup>25</sup> e-mail, internetová diskuse, diskusní fórum,<sup>26</sup> tweet. Elektronické texty<sup>27</sup> bývají nejčastěji rozlišovány na synchronní a asynchronní (např. Herring, 2011). Toto dělení akcentuje vázanost textu na čas, v němž je komunikován: synchronní texty (např. chat) jsou založené na

---

<sup>22</sup> V originále „die Gesamtmenge der in einer kommunikativen Interaktion auftretenden Signale“ (Adamzik, 2016: 69).

<sup>23</sup> Ještě nepřehlednější situace panuje v angličtině. Častým anglickým ekvivalentem českého termínu je *computer-mediated communication*, který se ustavil v začátcích výzkumu textů psaných na počítači a zprostředkovaných internetem (tj. e-mail, blog apod.). V současnosti se tímto termínem označují i texty, pro které typicky není počítač potřeba, např. SMS.

<sup>24</sup> I v tomto bodě panuje terminologická nejednotnost, např. Herringová (2008) užívá termín „systémy“, v originále *computer-mediated communication systems*. Detailněji viz např. Heyd (2015).

<sup>25</sup> IM lze chápat jako podtyp chatu, který probíhá zpravidla mezi dvěma komunikanty prostřednictvím speciální aplikace (např. ICQ, Skype, WhatsApp). Chat lze tedy rozdělit podle toho, zda probíhá mezi vícero komunikanty (tzv. webový chat: Jandová et al., 2006) nebo pouze mezi dvěma.

<sup>26</sup> Odlišnost diskusního fóra od internetové diskuse spatřuje Jílková (2016: 116) v tom, že účastníci internetové diskuse se vyjadřují k nějakému jinému textu (např. k publicistickému článku nebo internetové stránce), zatímco v diskusním fóru k předem určenému tématu, aniž by východiskem diskuse byl nějaký konkrétní text. Internetové diskuse se podle autorky v současné době přemístily na sociální síť Facebook či Twitter.

<sup>27</sup> Tento pojem používáme pro obecné označení textů vyskytujících se v elektronickém prostředí, tj. v různých aplikacích, platformách, ale i na webových stránkách.

interakci komunikantů v konkrétním čase, asynchronní (např. blog) na čase, v němž jsou zveřejněny (tj. komunikovány); jsou nezávislé v tom smyslu, že je možné komunikovat je kdykoli. Z časového aspektu se odvíjí blízkost synchronních textů s dialogičností a asynchronních s monologičností.

## 2.2 Klíčové aspekty výzkumu mluveného jazyka

Mluvená forma jazyka se stala předmětem výzkumu až po změně lingvistického paradigmatu, spočívající v přesunu „od zkoumání jazykového systému k výzkumu fungování tohoto systému v reálné situaci“ (Zeman, 2013: 25). Snaha o systematický přístup posléze vedla k rozvoji etnometodologie a konverzační analýzy (viz dále). V českém prostředí lze za jeden z prvních počínů označit referáty (a posléze články) K. Hausenblase *O studiu syntaxe běžně mluvených projevů* a M. Grepla *K podstatě a k povaze rozdílů mezi projevy mluvenými a psanými* (1962),<sup>28</sup> které zazněly na brněnské *Konferenci o historickosrovnávacím studiu slovanské syntaxe a o studiu skladby nářeční*. Studiu mluveného jazyka se předtím věnovali spíše dialektologové, neboť nářečí existuje (většinou) pouze v mluvené formě. J. Chloupek (1962) usouvztažňuje mluvenost a nářečí následovně: „mluvenost není prvkem stylotvorným, a tak jen na vysvětlenou při srovnávání s poměry ve spisovném jazyce můžeme říci, že místní nářečí má pouze jediný soubor výrazových prostředků, totiž ten, který přísluší hovorovému stylu.“ Chloupek dále připomíná, že řada jevů popisovaných jako nářeční je projevem existenčního modu a že je nutno tuto skutečnost při vědecké analýze brát v úvahu.

Při studiu mluveného jazyka se může jevit paradoxním fakt, že se tak děje prostřednictvím transkriptu. Ačkoli přepis vychází ze zvukového záznamu, není nikdy schopen pojmout a převést celý kontext nahrané mluvené situace a všechny zvukové kvality do grafické podoby. Každý přepis je nutně selektivní (viz např. Andersen, 2016), a proto je nutné vycházet při studiu nejen z přepisu, ale i ze zvukového záznamu. Přepis však na rozdíl od zvukového záznamu umožňuje rychlou orientaci a potřebný „pohyb“ po textu.

Výzkum mluveného jazyka je obecně interdisciplinární, neboť slučuje přístupy a poznatky mnoha vědních disciplín (kromě lingvistiky a jejích subdisciplín i např. sociologie, psychologie, kulturologie), směrů a osobností. V této kapitole se zaměříme na charakteristiku jazykových rovin mluveného jazyka; představené jazykové rysy mohou být – řečeno

---

<sup>28</sup> Kromě nich se v tomtéž sborníku věnovali tématu syntaxe mluveného projevu P. Adamec (srovnávající syntaktické prvky v češtině a ruštině), F. Miko (o navazovacích vztazích v hovorovém stylu) a dialektologové, např. J. Skulina, S. Prokešová, A. Vašek.

terminologií Kocha a Oesterreichera – jak mediální, tak konceptuální. Kromě nich zde představíme několik přístupů a publikací, které se mluveným jazykem zabývají.

### **2.2.1 Charakteristiky podle jazykových rovin**

Vlastnosti či charakteristiky mluveného jazyka ve sféře běžného dorozumívání, mezi něž patří např. spontánnost, neformálnost, se projevují na různých jazykových rovinách různými způsoby. Platí přitom, že jedna charakteristika může ovlivnit či determinovat jazykové prostředky na všech rovinách, ale také nemusí. V této části stručně představíme ty jazykové prostředky, které jsou podle lingvistů typické pro jednotlivé jazykové roviny. Na rozdíl od předchozí kapitoly se v této budeme zaměřovat pouze na češtinu.

V řadě prací o mluvené češtině je pozornost věnována jejím různým regionálním varietám. Vliv regionálního faktoru (např. původ mluvčího, původ adresátů, místo, kde se komunikace odehrává) lze pozorovat na všech jazykových rovinách, avšak v naší práci se jím nezabýváme. S odlišností mluvené češtiny, která se v podobě různých nářečí v některých regionech dodnes uchovala, od češtiny ustavené pro psaní je spojen koncept spisovnosti a existence tzv. hovorového jazyka.<sup>29</sup> Ani tato problematika není předmětem této práce. Snažíme se vybrat takové jazykové jevy, které jsou deklarovány jako univerzální pro mluvenou češtinu na celém území ČR.

S charakteristikou jednotlivých rovin souvisí i problematika jejich základních jednotek, která nabývá na důležitosti především v syntaxi.<sup>30</sup> Vymezení základní syntaktické jednotky mluveného jazyka není triviální, jedná se spíše o samostatné výzkumné téma (viz Hoffmannová & Zeman, 2017; Hoffmannová & Havlík, 2017). V následující charakteristice proto není blíže rozváděno. Při samotné analýze mluveného jazyka se vystříháme užití pojmu věta v mluveném jazyce, volíme raději obecnější označení replika či segment.

#### **2.2.1.1 Fonetická rovina**

Zvuková stránka běžně mluveného jazyka je velmi různorodá. Krčmová (2000: 119) považuje za výrazný rys všech běžně mluvených textů variabilitu jejich realizací. Právě při tvoření konkrétních hlásek, hláskových skupin a jejich seskupování do vyšších celků (taktů

---

<sup>29</sup> K problematice hovorového jazyka a hovorového stylu včetně jeho vztahu ke spisovnému jazyku viz Hoffmannová a Jílková (2016: 22n.).

<sup>30</sup> Vzato do důsledku je v mluveném jazyce problematický i koncept slova a jeho částí (morfémů, hlásek). Orientovat se podle výslovnosti mnohdy nelze, protože slova nemusí být vyslovena tak, abychom poslechem odlišili jejich hranice (definované např. pauzou, přerušením plynulého toku hlásek), často navíc dochází k modifikacím shluku hlásek, které při napsání vnímáme jako slovo. Tuto problematiku však v této práci nerozvádíme; podobně jako většina prací o mluveném jazyce analyzujeme transkripty, otázku segmentace řečového proudu na slova ponecháváme na příslušných transkripčních pravidlech.

a výpovědních či promluvových úseků) za využití suprasegmentálních prostředků (např. prozodie, intonace) se ukazuje vliv prostředí, v němž komunikace probíhá, a individualita každého mluvčího.

Signálem pro určení regionálního původu mluvčího jsou především specifické hlásky a hláskové změny.<sup>31</sup> Mezi fonetickými rysy platnými pro mluvenou češtinu ve všech oblastech České republiky bývá na segmentální rovině uváděna především zjednodušená (redukovaná) výslovnost, která může signalizovat např. nedbalost,<sup>32</sup> emocionalitu, obtížnou vyslovitelnost (jak u cizích slov, tak u komplikovaných konsonantických skupin), nízký kontrast mezi sousedními hláskami (Machač & Ziková, 2013: 29). Vyšší stupeň redukce lze také očekávat u dlouhých<sup>33</sup> nebo velmi frekventovaných<sup>34</sup> slov.

Na suprasegmentální rovině bývá jako typický zmiňován nepřízvučný začátek taktu.<sup>35</sup> Nepřízvučným začátkem se může rozumět výskyt příklonek (např. *ho, ji, jsem*) nebo jakýchkoli jiných výrazů, zpravidla jednoslabičných. Krčmová (2000) našla v neformálních nahrávkách z Moravy i z Čech v této pozici výrazy *a, no, jo, šak, tož*. Hoffmannová a Richterová (2015), které analyzovaly klitika *by, si, ti*, považují kladení těchto výrazů na začátek taktu za projev tendence ke zkracování, redukovanému a úspornému vyjádření, jež se projevuje nejčastěji při střídání mluvčích či změně tématu.

Další charakteristika se týká členění textu do taktů měřených ve slabikách. Palková (1997) uvádí pro spontánní monolog průměrnou délku 2,69 slabik. Krčmová (2000) konstatuje přibližně stejnou délku (variuující od 2,5 do 3,1 slabik v taktu) v nahrávkách soukromých i veřejných, dialogických i monologických.

### 2.2.1.2 Morfologická rovina

Významnou odlišností mluveného jazyka od psaného je rozdílná frekvence slovních druhů. Při porovnání procentního zastoupení slovních druhů v psaném a mluveném jazyce je patrný velký

---

<sup>31</sup> Názorný přehled viz v prvním a pátém svazku Českého jazykového atlasu (Balhar & Jančák, 1992; Balhar, 2005).

<sup>32</sup> Neboť mezi známými není potřeba pečlivě vyslovovat. Mluvčí jsou na sebe zvyklí a změna ve výslovnosti stejně jako změna jakéhokoli jiného stereotypního chování si žádá vysvětlení.

<sup>33</sup> Podíváme-li se v korpusu Ortofon např. na výslovnost tříslabičného slova *nějaký*, zjistíme, že je nejčastěji redukováno na dvě slabiky: *ňáký*.

<sup>34</sup> Mezi frekventovaná a zároveň redukovaná slova patří např. kontaktní výraz *člověče*, jehož redukovaná podoba *čéče* je již lexikalizovaná. Za redukci lze označit i krácení vokálů, např. *myslim*. Vztah mezi frekvencí a délkou slova patří stejně jako vztah mezi frekvencí a počtem významů slova k jazykovým pravidelnostem, které nejsou důsledkem gramatiky, ale obecných principů řečového chování (Čech, 2016).

<sup>35</sup> Pojetí taktu (fonetického slova) přejímáme od Krčmové (2000), která jím rozumí jednotku shodnou se slovem v běžném významu (tj. shlukem hlásek ohraničeným v psaném jazyce mezerami) nebo spojení klitik a přízvučného slova.

rozdíl. Podle Cvrčka et al. (2015: 163) patří mezi pět nejčastějších slovních druhů v psaném jazyce podstatná jména, slovesa, přídavná jména, zájmena a předložky, zatímco v mluveném jazyce se jedná o slovesa, zájmena, částice, podstatná jména a spojky. Novější srovnání vytvořené na větším objemu mluvených dat<sup>36</sup> upravuje pořadí pěti nejčastějších slovních druhů následovně: zájmena, slovesa, příslovce, podstatná jména a spojky. Kopřivová et al. (2017a: 63) konstatují dále velký rozdíl ve frekvenci částic (mluvený jazyk: 6,92 % vs. psaný jazyk: 0,97 %) a citoslovcí (mluvený jazyk: 0,34 % vs. psaný jazyk: 0,05 %).

Frekvenční odlišnosti slovních druhů jsou odrazem různých faktorů, např. vyšší výskyt částic v mluveném jazyce bývá vysvětlován větší mírou subjektivitu, častějším uváděním vlastních názorů a postojů, u citoslovcí se může jednat o zvýšenou emocionalitu a hravost, u zájmen a příslovcí o zakotvení komunikace do časoprostoru. Nelze opominout řadu pragmatických faktorů, např. získávání času či přitáhnutí pozornosti, které vedou ke specifickému použití či polyfunkčnosti výrazů v mluveném jazyce oproti psanému.<sup>37</sup> Hranice mezi slovními druhy je u takových výrazů těžko určit (Hoffmannová & Jílková, 2016: 51) uvádí jako příklad výraz *jako*.<sup>38</sup>

Z pohledu flektivní morfologie patří k charakteristikám mluveného jazyka na celém území České republiky především koncovka *-(a)/(e)ma* v I pl. podstatných a přídavných jmen, zájmen i číslovek. Distribuce této koncovky klesá s přibývajícím oficiálností. Podobně je tomu s koncovkou *-i* v N pl. obyvatelských jmen a u substantiv končících *-ista* (Bogoczová, 2000: 106n.). Pro mluvenou češtinu zejména v Čechách je dále typický kondicionálový tvar 1. os. pl. slovesa *být bysme* v opozici oproti *bychom* (např. Čmejrková & Hoffmannová, 2011), jehož užívání také souvisí s nižší mírou oficiálností.

### 2.2.1.3 Lexikální rovina

V souvislosti s lexikální rovinou bývá na předním místě zmiňována role tématu jako činitele ovlivňující výběr konkrétních lexémů (např. Fic, 2000; Čechová et al., 2008). Shromážděním tematicky dostatečně pestrého souboru mluvených textů můžeme oslabit vliv tématu a podívat

---

<sup>36</sup> Cvrček et al. (2015) pracovali s mluvenými daty z korpusů ORAL2006, 2008, PMK a BMK o celkové velikosti přibližně 3,2 mil. slov, zatímco Kopřivová et al. (2017a) použili mluvený korpus ORAL o velikosti 5,4 mil. slov.

<sup>37</sup> Příkladem může být ustavení tzv. diskurzivních markerů, tj. výrazů, které jsou v mluveném jazyce užívány v odlišných funkcích oproti psanému jazyku a u kterých proto nelze jednoznačně určit příslušnost ke slovnímu druhu (Čermáková et al., 2019).

<sup>38</sup> Nabízí se zde otázka, do jaké míry je rozdělení do tradičních slovních druhů účelné. Výhodou je samozřejmě možnost srovnání. Uvážíme-li, že přiřazení ke slovním druhům by mělo odrážet funkci výrazů v kontextu, bylo by nutno nejprve podchytit tyto funkce, tj. provést detailní výzkum mluveného jazyka a ten promítnout do korpusového značkování slovních druhů.



se na nejfrekventovanější lexémy. *Frekvenční slovník mluvené češtiny* (dále jen FSMČ, Čermák et al., 2007) uvádí deset nejčastějších lemmat (tj. výrazů ve slovníkovém tvaru): *ten, být, a, že, já, mít, no, v, na, jako*.<sup>39</sup> Od seznamu deseti nejčastějších lexémů v psaném jazyce (Čermák & Křen, 2004) se kromě pořadí (v psaném je na prvním místě *a*, následuje ho *v*) liší výrazy *já, mít, no, jako*,<sup>40</sup> místo nichž se v psaném jazyce objevují *se, na, z, který*. Toto hrubé srovnání poukazuje především na význam synsémantik při jazykové produkci a na možnou souvztažnost určitých výrazů (*no, jako*) s jazykovým modelem, resp. podmínkami jazykové produkce.

Mluvenému jazyku bývá připisována sémantická neurčitost sdělení. Podle Urbanové (2003) je tato neurčitost, vyplývající z nepřipravenosti a spontánnosti, výhodná pro adresáta, neboť pro něj je snazší orientovat se v lexikálně méně náročném sdělení s méně komplikovanou syntaktickou stavbou. Jako její projev na lexikální rovině můžeme chápat nejen zapojení neurčitých, vágních výrazů a konstrukcí (např. *jako, a tak, takový nějaký*), ale i deiktických výrazů (čímž lze vysvětlit vysokou frekvenci lexému *ten*), které propojují text s konkrétní komunikační situací, jejími účastníky a prostředím. Tvrzení o vyšší míře konkrétnosti mluveného jazyka však opět souvisí se zakotvením textu do situace.

V mluveném jazyce se objevují ustálená spojení lexémů, která nabývají téměř podoby frazému (Müllerová, 2011: 157).<sup>41</sup> Jejich využití je pravděpodobně motivováno omezeným časem pro produkci: čím déle mluvčí přemýšlí o vhodné formulaci, tím pravděpodobněji se slova ujme jiný mluvčí.

I další charakteristiky – stručnost (např. zkratková slova, univerbizace) a expresivita (např. vulgarismy, deminutiva) – jsou ovlivněny podmínkami při produkci mluveného jazyka a celkovou dynamikou interakce. Hoffmannová (2016: 49) připomíná časté užívání intenzifikátorů typu *děsný, strašný, hrozný, úžasný* a některých expresivních frazémů (např. *dát si majzla*).

---

<sup>39</sup> FSMČ (Čermák et al., 2007) je vytvořen na základě korpusu PMK, obsahujícího cca 675 000 slov. Podíváme-li se do novějšího a v současné době největšího mluveného korpusu ORAL (5,4 mil. slov), zjistíme, že tento seznam nejfrekventovanějších výrazů se nezměnil.

<sup>40</sup> Podle absolutní frekvence se ve *Frekvenčním slovníku češtiny* (Čermák & Křen, 2004) objevuje *mít* mezi prvními dvaceti slovy, *jako* a *já* mezi třiceti a *no* mimo sto nejfrekventovanějších.

<sup>41</sup> Idiomaticita v jazyce (*idiom principle*) vztáhl Sinclair (1991) k opačnému principu – volnému výběru jazykových jednotek (*open choice principle*): mluvčí má k dispozici množství předpřipravených (prefabrikovaných) frází, které může při produkce zvolit, příp. je upravit. Na druhou stranu každá pozice ve větě nabízí možnost obsadit ji, konkrétní možnosti obsazení jsou dány příslušným jazykem. Idiomaticita se zabývá řada lingvistických disciplín od lexikologie po jazykovou akvizici (přehled viz např. Erman & Warren, 2000), v mluveném jazyce se na ni zaměřila Aijmerová (1996).

Podobně jako na fonetické rovině jsou i na rovině lexikální nápadným rysem dialektismy a regionalismy. Tyto prostředky však podle Bogoczové et al. (2000: 144) „stojí na okraji lexikonu běžné komunikace“ a jsou vázány „na oblast nejintimnější rodinné a přátelské konverzace“.

#### **2.2.1.4 Syntaktická rovina**

Syntaxí mluveného jazyka se dlouhodobě zabývala O. Müllerová. Jejím pojmem „mluvená syntax“ odkazuje Hoffmannová (2016: 49) k charakteristice syntaktické roviny prostřednictvím výčtu syntaktických rysů, které se vyskytují nejčastěji v běžně mluveném jazyce<sup>42</sup> (výčet viz např. Müllerová, 1994, 2011).

Na syntaktickou rovinu mají více než regionální odlišnosti vliv situační podmínky vzniku textu. Jejich součástí je i čas, v němž se komunikace odehrává (viz Auer, 2009). Plynutí v čase zapříčiňuje lineární uspořádání textu, postupné a následné kladení syntaktických konstrukcí za sebou a nemožnost editovat již vyslovené. Syntaktické konstrukce jsou napojeny bez explicitního vyjádření jejich vztahu asyndeticky nebo stereotypně pomocí výrazů *a*, *tak* nebo *no* (Müllerová, 2011). Chloupek (2000) usouvztažňuje toto asociativní napojování s málo výraznou intonační konturou výpovědi; její konec je signalizován spíše vystřídáním mluvčího.

K dalším výrazným rysům patří podle Müllerové (1994, 2011) křížení syntaktických výpovědních perspektiv na úrovni větných i souvětných vztahů, paralelismus ve výstavbě syntaktických konstrukcí a jejich nedokončování, aditivnost (tj. dodávání, rozvíjení syntaktické struktury, přičleňování), subjektivní pořadí v tematicko-rematické výstavbě, neobvyklé slovosledné postavení některých slov nebo konstrukcí, opakování výrazů i konstrukcí (včetně tzv. falešného začátku, restartu), opravy a dvojí vyjádření téhož, eliptické konstrukce, anakoluty, vsuvky atd. Syntaktická stavba mluveného textu by se tak často dala označit za útržkovitou, neurovnanou, neuspořádanou, záleží však vždy na konkrétním textu.

K syntaktickým jevům bývají také řazeny různé způsoby reprodukce cizí a vlastní řeči, střídání slovesného času ve vypravování, citoslovce nebo onomatopoické zvuky ve funkci predikátu nebo jiného větného členu.

#### **2.2.2 Směry a přístupy zabývající se mluveným jazykem**

Mluvený jazyk ve sféře běžného dorozumívání je často zdrojem pro studium jazykového úzu, a to zejména v dialektologii, sociolingvistice, interakční lingvistice, analýze diskursu a

---

<sup>42</sup> Müllerová (2011: 154) uvádí, že „typické syntaktické mluvenostní jevy jsou i ve značné kumulaci nejcharakterističtější pro soukromé projevy mluvčích s nižším vzděláním v běžném denním dialogickém styku“.

konverzační analýze.<sup>43</sup> Uvedené směry se zabývají variabilitou mluveného jazyka podmíněnou především regionální příslušností mluvčích (dialektologie) a situačními charakteristikami (ostatní zmiňované směry); důraz je kladen zejména na neformálnost, neoficiálnost, soukromé prostředí.<sup>44</sup> Vlivem mimojazykových charakteristik na podobu mluveného i psaného jazyka se dále zabývá tzv. multidimenzionální analýza (Biber, 1988).

V této kapitole stručně představíme tři lingvistické směry, které jsou pro druhou část práce nejvíce inspirativní. Jedná se o sociolingvistiku, konverzační analýzu a interakční lingvistiku. Pro analytickou část této práce je výchozím bodem a metodologickou oporou multidimenzionální analýza. Její principy proto uvedeme v kontextu ostatních tří směrů v této kapitole. Kombinací těchto a dalších lingvistických směrů a metodologií vzniklo několik souhrnných příruček o mluveném jazyce, které představíme na závěr kapitoly.

### **2.2.2.1 Sociolingvistika**

Sociolingvistika je obecné označení různých přístupů zaměřujících se na vztah jazyka a společnosti. (k rozdělení sociolingvistiky viz Nekvapil, 2017) Pro naše úvahy o konceptuální mluvenosti jsou inspirativní zejména využití sféry běžného dorozumívání pro studium jazykových variací a metodologické postupy používané pro sběr takových dat.

Korelace jazykových aspektů se sociálními charakteristikami se zabývá variační sociolingvistika, jejímž čelným představitelem je americký lingvista W. Labov. Jedním z jeho přínosů na poli (socio)lingvistického výzkumu je metodologie sběru dat využívající vlastností běžného dorozumívání. Tato metoda, aplikovaná např. při studiu fonologických variací na ostrově Martha's Vineyard (Labov, 1963), spočívá ve vedení rozhovoru s informantem v co nejméně formálním duchu. Při dotazování se Labov zároveň snažil, aby informanti použili cílená slova (obsahující určité fonémy v různém fonologickém okolí) sami, bez jeho přímé žádosti, a v přirozeném kontextu. Informanti také neměli postřehnout, co je pravým předmětem výzkumu.

Fonologické variace, kterým se Labov věnoval i např. v New Yorku (2006 [1966]), jsou podmíněny osobou mluvčího, recipienta/adresáta i situace. Při výzkumu ve třech nákupních střediscích New Yorku se Labov zaměřil na vibrantu /r/: zda existuje souvislost mezi sociálním

---

<sup>43</sup> Záměrně zde neuvádíme stylistiku, ale směry blíže specifikující svůj předmět zájmu.

<sup>44</sup> Srov. s metodologií sběru dat pro obecné mluvené korpusy a nářeční korpusy (Benešová et al., 2015; Goláňová, 2015).

statusem a frekvencí rhotického /r/.<sup>45</sup> Při sběru dat použil dvě metody: nechal participanty číst seznamy slov a delší text a také s nimi vedl neformální rozhovor. Snažil se tedy obsáhnout jak situaci, v níž mluvčí dbá na svůj projev (zde zejména výslovnost), protože je požádán o splnění určitého zadání a ví, že je pozornost soustředěna na jazyk, tak na opačnou situaci, při níž nemusí plnit žádný úkol a v popředí je komunikace jako taková (příp. její obsahová složka). Výzkum ukázal, že frekvenci výskytu rhotického /r/ ovlivňuje jak sociální status, tak situace: nejvíce ho bylo zaznamenáno při čtení seznamu slov, nejméně při neformálním rozhovoru.

K sociolingvistickým výzkumům orientovaným na češtinu a inspirovaným Labovem patří např. studie J. Wilsona (2010) o jazykové akomodaci moravských mluvčích v Praze a J. Chromého (2017) o užívání protetického /v/ ve vybraných českých městech. Oba lingvisté sbírali data prostřednictvím rozhovorů: Wilson měl od každého informanta dva rozhovory,<sup>46</sup> z nichž první byl veden jako neformální dialog mezi dvěma vysokoškoláky a druhý jako řízený rozhovor, Chromý měl od každého informanta jeden přibližně hodinový rozhovor, který byl veden jako neformální nestrukturovaný dialog mezi dvěma dobrými známými (rodinnými příslušníky, kamarády). Obě práce se zvolenou metodou snažily zajistit takový jazykový materiál, v němž by byla míra pozornosti mluvčích vůči svému jazyku co nejmenší.

Výše již bylo nastíněno, že při sběru dat záleží také na tom, jakou skupinu lidí pro výzkum vybrat. Sociolingvisté vymezují různé typy skupin, při jejichž odlišení se kromě charakteristiky vztahu, sociálních rolí, socioekonomického statusu apod. bere v úvahu i účel komunikace a prostředí, v němž se interakce odehrává. Například jazyková komunita (*speech community*)<sup>47</sup> akcentuje oproti společenství praxe (*community of praxis*), zaměřené na společnou činnost, jazyk. Členové jazykové komunity používají jazykový kód k manifestaci sebe sama s důrazem na své členství v komunitě, která je pro ně zároveň základnou sociálních vztahů. Hymes (1972) připomíná výskyt pravidel (komunikace i interpretace), které komunita sdílí. Gumperz (1972) dodává, že k interakci členů komunity dochází pravidelně a frekventovaně. Příkladem jazykové komunity může být rodina, jejíž členové se pravidelně schází např. u společného jídla či oslav,

---

<sup>45</sup> V angličtině se rozlišují rhotická a nerhotická nářečí. Ve rhotických nářečích se /r/ vyslovuje ve všech pozicích (např. *more* [mɔr]). V nerhotických se výslovnost zachovává před samohláskami, zatímco po samohláskách se /r/ nevyslovuje.

<sup>46</sup> Oba rozhovory byly nahrávány ve stejný den. Nejprve 20minutový neformální dialog (*sociolinguistical interview*), který vedla Wilsonova přítelkyně Markéta, poté následoval strukturovaný rozhovor s J. Wilsonem, v němž informanti odpovídali na 28 otázek, rozdělených na dvě sady (*life-style questions, attitude questions*). Druhý rozhovor nebyl časově omezen a varioval od 17 do 49 minut.

<sup>47</sup> Lze přeložit i jako řečové společenství. Srov. pojem „diskurzivní“ komunita (*discourse community*) (Swales, 1990) založený na žánrech.

a pravidelně spolu komunikují. Každý člen rodiny zná obvyklý způsob vyjadřování jednotlivých členů, takže ví, jak má jejich slova interpretovat.

### **2.2.2.2 Interakční lingvistika a konverzační analýza**

Interakční lingvistika a konverzační analýza (dále jen KA) jsou spolu úzce spjaty; Havlík (2017) uvádí KA jako směr, z něhož se interakční lingvistika postupně vydělila. Oba směry se zaměřují na jazykové prostředky užívané v rozhovorech při vytváření jejich sekvenčního uspořádání. Oproti strukturalistické či generativní lingvistice je akcent kladen na samotný proces interakce mluvčích. Interpretaci užívaných jazykových prostředků a jejich vnímání osobou mluvčího se pak zabývá interakční sociolingvistika (Nekvapil, 2017). O interakční lingvistice jako o specifickém směru se hovoří zejména v souvislosti s analýzou syntaktické roviny mluveného jazyka.

Interakční lingvistika pracuje výhradně s přirozenými daty, což znamená, že analyzované interakce probíhaly v přirozeném prostředí. V tomto přístupu souzní s KA orientovanou především ve svých počátcích na přirozenou mluvenou komunikaci ve sféře běžného dorozumívání (Sacks et al., 1974). KA se podobně jako interakční lingvistika zabývá jazykovými prostředky užívanými mluvčími při produkci přirozených rozhovorů, mimoto se zajímá i o metody užívané mluvčími při sekvenčním uspořádání rozhovoru a interpretaci všech těchto aspektů (Nekvapil, 2017). KA je známa detailní analýzou jazykového chování ve vztahu ke komunikační situaci a jejích účastníků, z čehož plyne i její důraz na pečlivou transkripci rozhovoru.<sup>48</sup> Interakční lingvistika i KA se v současnosti zaměřují i na analýzu neverbální složky komunikace, sbírají tedy audiovizuální materiál.

Jeden z konceptů, s nimiž interakční lingvistika pracuje, se týká lineární a v reálném čase probíhající výstavby rozhovorů, tzv. online syntaxe. Za základní principy v online syntaxi považuje Auer (2009) a) projektivitu, tj. vytváření očekávání o tom, jak bude replika pokračovat, b) expanzi, tj. (zejména) syntaktické rozvíjení stávající struktury, a c) retrakci, tj. obměnění či oživení stávající struktury v oblasti významu. Auer také zdůrazňuje, že v mluveném jazyce se syntaktická rovina může kdykoli změnit, např. participací dvou mluvčích na replice či samoopravě jednoho mluvčího.

V souvislosti s KA je nutno zmínit práce O. Müllerové, J. Hoffmannové, S. Čmejrkové, v nichž se spojují postupy textové lingvistiky s KA a analýzou diskursu (např. Müllerová, 1994; Müllerová & Hoffmannová, 1994). Zmiňované autorky se kromě mluveného jazyka ve sféře

---

<sup>48</sup> Transkripční zásady otištěné v článku Sackse et al. (1974) jsou s obměnami a doplněními využívány dodnes.

běžného dorozumívání, konkrétně např. vzpomínkovému vyprávění (např. 2007), zabývaly i institucionální (Hoffmannová & Müllerová, 2000) či mediální komunikací (např. Čmejrková et al., 2013).

Příkladem spojení interakční lingvistiky, sociolingvistiky a korpusové lingvistiky je výzkum B. Clancyho shrnutý v knize *Investigating intimate discourse* (2016). Předmět studia – *intimate discourse* – můžeme na základě jeho definice na s. 171 ztotožnit se sférou běžného dorozumívání, konkrétně s mluvenými dialogy probíhajícími mezi komunikanty s blízkým vztahem (milostní partneři, rodina, přátelé) v soukromém, neformálním prostředí ve sdíleném fyzickém prostoru.<sup>49</sup> Clancy si pro analýzu vybral různé jazykové prostředky, které bývají uváděny jednak jako projevy určité charakteristiky spontánní komunikace (např. osobní zájmena ve vztahu k interpersonalitě), jednak jako frekventované v tomto typu mluvené komunikace (např. osobní zájmeno *já*). Závěrem shrnuje zkoumané charakteristiky do čtyř „schémat“ (*patterns*), které *intimate discourse* charakterizují oproti jiným komunikačním sférám (např. pedagogické). Jedná se o následující:

- 1) vysoká frekvence 3. zájmené osoby v singuláru
- 2) nízká frekvence 1. zájmené osoby v plurálu
- 3) existence vokativu a vyrovnaný podíl jeho užívání ve funkci interpersonální (tj. kontaktové) a organizační
- 4) vyšší frekvence tzv. *disjunctives* oproti tzv. *adjunctives*<sup>50</sup>

### 2.2.2.3 Multidimenzionální analýza

Multidimenzionální analýza (dále jen MDA) nepředstavuje na rozdíl od výše zmiňovaných směrů samostatné odvětví lingvistiky. Jedná se o metodologický přístup zaměřený na lingvistickou analýzu textů, jejímž cílem je odhalit, které mimojazykové faktory<sup>51</sup> mají největší vliv na odlišnost mezi různými texty a které tedy konstituují registry.<sup>52</sup> Zakladatel MDA D. Biber (1988: 9) nazývá tyto faktory dimenzemi, protože definují spíše kontinuum než diskrétní póly. Texty jsou umístěny v multidimenzionálním prostoru; z pohledu jedné dimenze se mohou překrývat, z pohledu jiné být na opačných koncích prostoru, proto je nutný

---

<sup>49</sup> Clancy pro svou analýzu použil korpus mluvené irské angličtiny LCIE (Limerick Corpus of Irish English), který obsahuje cca 1 mil. slov.

<sup>50</sup> Oba termíny označují slova či slovních spojení s obecným významem, které se vyskytují se na koncích vět a které neslouží k referenci o něčem specifickém, ale spíše vágním. *Disjunctives* jsou spojení začínající slovem *or* (např. *or whatever*), *adjunctives* začínající slovem *and* (např. *and stuff like that*).

<sup>51</sup> Faktory, které variabilitu mezi texty zapřičiňují, jsou v české stylistické tradici rozděleny na objektivní a subjektivní.

<sup>52</sup> V české stylistice tomuto pojmu nejvíce odpovídají pojmy styl objektivní nebo slohový postup. Ke vztahu těchto pojmů viz Hoffmannová (1997: 138n.) a Mrázková (2017).

komparativní pohled. Každá dimenze se dá definovat jazykovými prostředky (dále rysy), které se v ní uplatňují. MDA je založena na předpokladu, že dimenze jsou konstituovány významnými souvěsky<sup>53</sup> určitých rysů.

MDA se od tradiční stylistiky, orientující se taktéž na analýzu konkrétních textů a stylových útvarů (žánrů), liší důrazem na kvantitativní postup a cílem, „jímž není určení textových prototypů a vymezení jejich charakteristik, ale spíše identifikace dimenzí variability na základě využívání jazykových prostředků (dále rysů), u nichž funkční variabilitu lze předpokládat“ (Cvrček et al., 2018b: 294).

MDA probíhá v několika krocích. Prvním z nich je shromáždění dostatečně pestrého souboru jazykových dat (jazykového korpusu), na kterém bude jazyková variabilita zkoumána. Druhým krokem je sestavení souboru jazykových rysů, které můžeme považovat za nositele variability. Tyto rysy jsou posléze v korpusu identifikovány<sup>54</sup> a pomocí statistických metod je vyhodnocena jejich distribuce v textech. Následuje faktorová analýza, která na základě distribuce rysů vytvoří několikadimenzionální model variability v daném korpusu. Na ustavení každé dimenze se podílí frekvence, souvěskyt i distribuce určitých rysů. Poslední fází MDA je interpretace těchto dimenzí, tedy pojmenování faktoru způsobujícího variabilitu. Po tomto kroku přichází na řadu ustavení registrů, tj. skupin textů založených na sdílených jazykových charakteristikách v rámci jednotlivých dimenzí. (Pro více detailů o jednotlivých krocích viz Biber, 1988; Cvrček et al., v tisku.)

Vytvoření MDA bylo motivováno kromě empirické/kvantitativní identifikace faktorů také snahou o srovnání psaného a mluveného jazyka. Podle Bibera (1988) není mezi oběma jazykovými módy absolutní rozdíl, proto nelze hovořit o dvou odlišných jazycích. Mluvený a psaný jazyk se v Biberově analýze z r. 1988 nejvíce liší ve třech dimenzích (dimenze 1: *involved* vs. *informational*, dimenze 3: *explicit reference* vs. *situation-dependent reference*, dimenze 5: *abstract* vs. *non-abstract information*)<sup>55</sup>, jejichž kombinací lze nejlépe postihnout charakter mluveného a psaného jazyka. Mluvený jazyk je založen na zaujetí či angažovanosti produktora (*involved*), referuje k situaci, v níž probíhá, a obsah vyjádření v něm není abstraktní. V psaném jazyce hraje větší roli jeho informační obsah než zaujetí produktora, příznačná je explicitní a propracovaná struktura a obsahově se jedná o abstraktní, technické informace.

---

<sup>53</sup> Biber (1988) píše přímo o vzorcích či schématech (*pattern*), které souvěskyt rysů vytváří.

<sup>54</sup> Jazykový rys představuje teoretický koncept. Pro jeho vyhledání v textech je nutné ho operacionalizovat, tj. vytvořit způsob, jakým bude nalezen v textu. Konkrétněji v 3.1.2.

<sup>55</sup> Pojmenování dimenze je vždy víceslovné, pojmenován je kladný i záporný pól.

Konkrétní typy textů či žánrů se v rámci obou modů vyskytují na škále, jejíž póly Biber ztotožňuje s běžně mluveným dialogem tváří v tvář (pól mluveného jazyka) a akademickým výkladem (pól psaného jazyka).

MDA byla aplikována na řadě typologicky odlišných jazyků (např. angličtina, španělština, korejština). Aplikací na češtinu se zabýval tým z ÚČNK FF UK,<sup>56</sup> který navázal na nedokončený pokus V. Kodýtky z r. 2007. Podle Cvrčka et al. (2018b) se na variabilitě českých textů (včetně mluvených projevů) nejvíce podílejí následující čtyři dimenze: dimenze 1 (dynamický vs. statický), dimenze 2 (spontánní vs. připravený), dimenze 5 (vyšší vs. nižší míra explicitní adresnosti), dimenze 8 (postojovost vs. faktualnost).

Tyto dimenze ukazují rozdíly nejen obecně mezi mluveným a psaným jazykem, ale také mezi jednotlivými žánry či podtypy daného komunikačního modu. Zaměříme-li se pouze na rozdíly v mluveném jazyce, podchycují největší rozdíly první a druhá dimenze (obr. 2 a 3). Prostřednictvím první dimenze lze odlišit interaktivní, dialogický mluvený jazyk (na obrázku označen spo-int)<sup>57</sup> od neinteraktivního, předem připraveného jazyka (označení spo-nin): první z nich zastupuje dynamický způsob vyjádření, spočívající ve využití slovesných rysů, zatímco druhý prezentuje statické vyjadřování, využívající především jmenných rysů. Oba typy mluvené komunikace se ještě výrazněji odlišují v druhé dimenzi, v níž se interaktivní typ mluveného jazyka zcela vyděluje od ostatních žánrů. Jeho odlišností je právě vysoká míra spontánnosti. V míře spontánnosti se mu z psaných žánrů blíží soukromá korespondence (označena wri-pri) a interaktivní internetová komunikace (označena web-mul).<sup>58</sup> Pro všechny ostatní žánry včetně neinteraktivní mluvené komunikace je určující připravenost na rozdíl od spontánnosti.

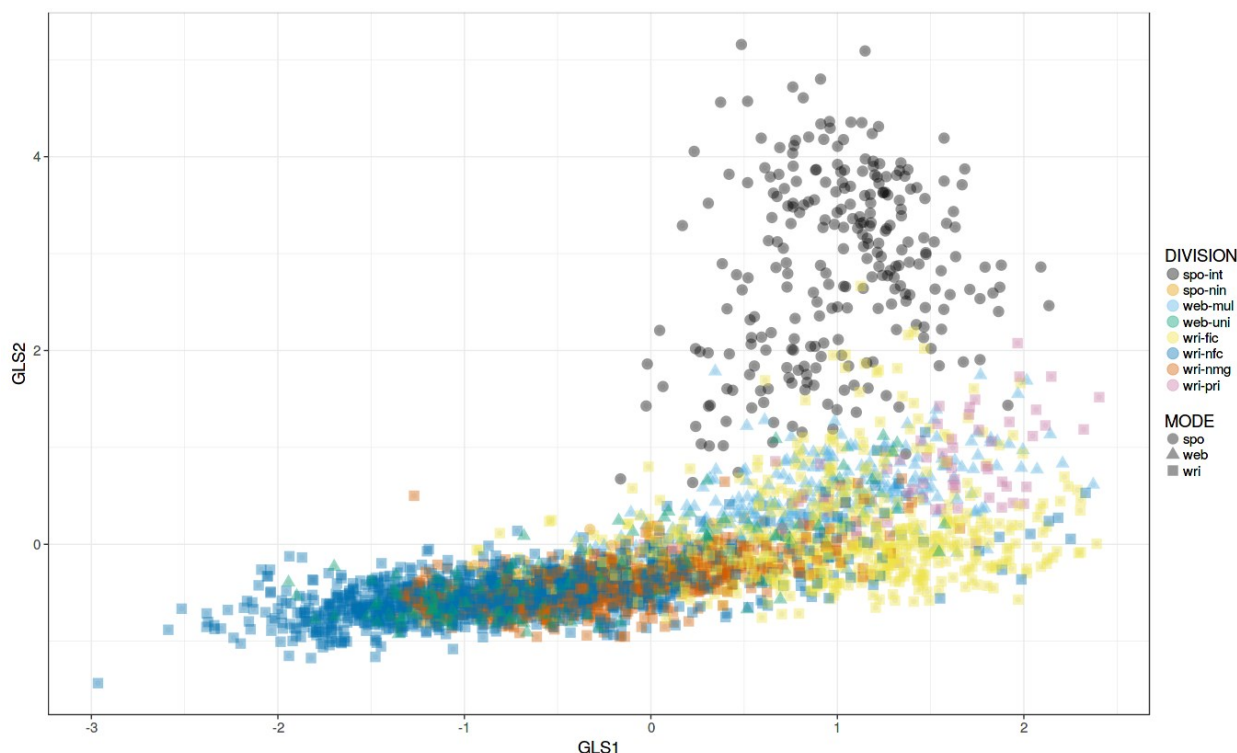
---

<sup>56</sup> Jeho členkou je autorka této práce.

<sup>57</sup> Pod tímto pojmem jsou sdružena data, jejichž společným jmenovatelem je nepřipravenost. Jedna část dat zahrnuje běžnou, neformální soukromou dialogickou komunikaci (z korpusu mluvené češtiny ORAL2013, jehož nástupcem je korpus Ortofon, použitý v této práci), druhou část tvoří elicitovaná řeč, konkrétně odpovědi na otázky kladené výzkumníkem (z mluvených korpusů PMK a BMK).

<sup>58</sup> Pod tímto označením se skrývají data z internetových diskusí, diskusních fór a veřejné statusy na sociální síti Facebook (viz 3.1.1).

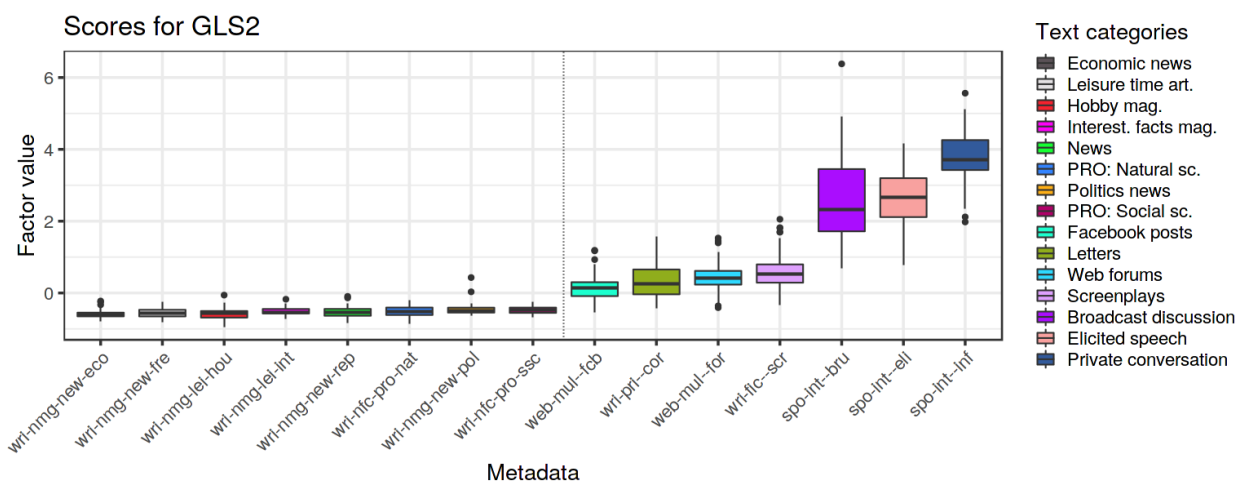




Obr. 2: Rozdělení textů na základě MDA podle dimenzí 1 (GLS1) a 2 (GLS2) (zkratka *spo* označuje mluvený jazyk, *wri* psaný, *web* webový; *int* interaktivní, *nin* neinteraktivní, *mul* vícesměrný, *uni* jednosměrný, *fic* beletrii, *nfc* odbornou literaturu, *nmg* publicistiku, *pri* soukromé dopisy)<sup>59</sup>

Lepší znázornění toho, jaké typy textů, resp. žánry, se ve druhé dimenzi nejvíce blíží pólu spontánnosti, ukazuje obr. 3. V pořadí jednotlivých typů textů se nejvýrazněji odlišují tři žánry interaktivních mluvených textů: neformální (*spo-int-inf*), elicitované (*spo-int-eli*) a mediální (*spo-int-bru*). Za nimi již následují psané texty: scénáře (*wri-fic-scr*), diskusní fóra (*web-mul-for*), soukromá korespondence (*wri-pri-cor*), facebookové příspěvky (*web-mul-fcb*). Na levé straně grafu jsou znázorněny texty jevící se jako připravené; jedná se především o různé typy publicistických textů (ekonomické zpravodajství, volnočasová publicistika atd.).

<sup>59</sup> Obrázek 2 je převzat z nástroje pro vizualizaci MDA v češtině, který byl vytvořen Mgr. Davidem Lukešem.



Obr. 3 Krabicový graf (boxplot) žánrů ve druhé dimenzi z korpusu Koditex<sup>60</sup>

### 2.2.3 Vybrané publikace o mluveném jazyce

Mluvený jazyk přitahuje pozornost lingvistů dlouhodobě. Vznikla již řada samostatných publikací věnujících se jeho specifičnosti, z nichž některé jsou přímo nazvány gramatikami. Takové gramatiky se snaží popsat mluvený jazyk zpravidla odlišně od gramatiky psaného jazyka, v níž se objevuje charakteristika jednotlivých jazykových rovin, jejich základních jednotek apod. V této části stručně představíme tři publikace o mluveném jazyce, které akcentují jeho odlišnost od psaného a snaží se vytvořit samostatný koncept pro jeho další zkoumání. Ve většině z nich najdeme důraz na analýzu bezprostředního okolí zkoumaných jednotek a zájem o omezený úsek textu, což jsou aspekty využívané v KA a interakční lingvistice.

#### 2.2.3.1 David Brazil (1995): A Grammar of Speech

Tato publikace bývá označována za inovativní, a to nejen kvůli mluvenému jazyku, ale především kvůli zaměření na účastníky komunikace a jejich potřeby. Brazil nazývá svůj přístup *purpose-driven* a staví ho do opozice k tradiční gramatice zaměřující se na abstraktní systém s centrálním postavením věty (*sentence-oriented grammar*).<sup>61</sup>

*Grammar of speech* je ovlivněna především KA (viz 2.2.2.2) a M. A. K. Hallidayem. Tento vliv je možno pozorovat v akcentu na interaktivní povahu mluveného jazyka, důraz na kontextovou vázanost významu a zdůraznění komunikačních potřeb mluvčích.

<sup>60</sup> Obrázek 3 je převzat z nástroje pro vizualizaci MDA v češtině, který byl vytvořen Mgr. Davidem Lukešem.

<sup>61</sup> Tato změna se promítá také do používaného jazyka, např. místo popisu produktů jazykového vyjádření (otázka, tvrzení) užívá Brazil raději sloves popisujících proces produkce (*ptát se, vyprávět*).

Brazilův přístup je založen na analýze konkrétních mluvených dat, z nichž abstrahuje závěry formulované jako rysy mluveného jazyka. Bere přitom v úvahu lineární uspořádání mluveného jazyka, tedy fakt, že slova jsou produkována v čase po sobě a nelze je po vyřčení vzít zpět. KA je inspirací také ve zdůraznění interaktivního pohledu na mluvený jazyk. Právě interaktivnost je podle Brazila nejcharakterističtějším rysem mluveného jazyka, který se promítá do všech jazykových rovin. Mluvčí v komunikaci vzájemně kooperují; pokud dojde k selhání, nastoupí mechanismy oprav.

Jednotlivé inkrementy, tj. úseky řeči, které Brazil sleduje, plní komunikační potřebu mluvčího v tom smyslu, že navazují na předchozí inkrementy a jejich téma. Inkrementy mohou plnit vícero komunikačních potřeb; často je lze vzhledem ke komunikační potřebě označit za neurčité, jejich interpretaci je nutno založit na kontextu. Každý inkrement probíhá z iniciálního do koncového stavu, přičemž do každé fáze se promítají situační charakteristiky včetně představ mluvčího o tom, jak by měl vypadat cílový stav. Brazil popisuje systém pravidel, která jsou aplikována přísně lineárně a která jsou výsledkem produkce jednoduchých řetězců.<sup>62</sup> Co se týká jedné jednotky či inkrementu, týká se i vztahu mezi jednotkami či inkrementy.

Na Brazilův přístup ke zkoumání a popisu mluveného jazyka navazují *pattern grammar* (Hunston & Francis, 2000) a *linear unit grammar* (Sinclair & Mauranen, 2006). V této kapitole se budeme věnovat pouze poslední jmenované, která se na rozdíl od *pattern grammar* nezaměřuje pouze na lexikální rovinu.

### **2.2.3.2 John Sinclair – Anna Mauranenová (2006): Linear Unit Grammar**

Také přístup nazvaný *linear unit grammar* (dále jen LUG) je založen na studiu neformální mluvené angličtiny. Jeho základní myšlenkou je rozdělení jakéhokoli textu do jednotek s využitím principu linearity, který je přítomný jak v mluveném, tak v psaném jazyce. LUG je podobně jako Brazilův přístup založen na analýze bezprostředních jednotek, zajímají ho syntagmata.

LUG vychází z předpokladu, že mezi lidmi panuje podobnost v tom, jak text při vnímání analyzují a rozkládají na menší jednotky (*chunky*). Hranice jednotek je samozřejmě variabilní a souvisí s možnými interpretacemi. Maximální velikost jednotky přitom není dána, autoři však uvádějí, že jen málo jich je větších než 5 slov. LUG vychází z lidské intuice a z již vytvořených konceptů, z nichž se snaží vytvořit koherentní teorii a analytický systém. Výsledkem je text

---

<sup>62</sup> Brazil píše, že určité slovní druhy vyhledávají jiné slovní druhy, např. gramatický člen nebo adjektivum vyhledávají substantivum.

rozčleněný do jednotek a doplněný poznámkami o aspektech interakce, které nebyly obsaženy ve vstupním materiálu. Takový výsledek pak může být zpracován tradiční gramatikou a opakováním analýzy lze dospět k totožnému výsledku.

Prvním krokem analýzy je rozdělení na provizorní jednotky podle intuice. Ty jsou následně klasifikovány do dvou typů: M (*message-incremental*), nebo O (*organisational*). Jednotky M se podílejí na rozvinutí tématu (a tím i zvětšování rozsahu sdílených vědomostí komunikantů), zatímco jednotky O mají funkci organizační (např. určují vztah mezi jednotkami M, napomáhají organizaci dialogu). Nezáleží na to, že v jednom kontextu je daná jednotka M a v jiném O, vždy se jedná o jednotku v určitém kontextu. V dalších krocích dochází k jemnější klasifikaci obou typů jednotek. Posledním krokem je sdružení jednotek M do vyššího významového celku (tzv. *linear unit of meaning*, LUM). Hranice LUM se více méně kryjí s jednotkami tradiční gramatiky – větami, souvětími. Jejich význam je považován za integrální součást způsobu vyjádření.

Autoři LUG testovali metodu na různých anglickojazyčných datech (kromě mluveného jazyka i na SMS, beletrii, publicistice), aby ověřili její univerzálnost. Všechny kroky analýzy s výjimkou prvního by také bylo možné automatizovat a po tomto jednotném zpracování porovnávat různé typy textů. Další aplikaci LUG vidí autoři v translatoologii a při výuce cizích jazyků.

### **2.2.3.3 Mathilde Hennigová (2006): Grammatik der gesprochenen Sprache in Theorie und Praxis**

Hennigová v této knize představuje vlastní teoretický přístup k mluvenému jazyku, který by umožnil uspokojivě popsat mluvený jazyk a jeho specifika. Nejedná se o již hotovou gramatiku mluveného jazyka nebo ucelenou teorii, ale spíše o příspěvky k budoucí teorii. Ta by měla čerpat z již zavedených konceptů a přístupů vybírajíc z nich ty, které jsou vzájemně kompatibilní a pro teorii výhodné; spojovat by měla teorii s empirií, gramatiku ve smyslu morfologie a syntaxe s pragmatikou, symbolickou rovinu jazyka s vnějším kontextem, tj. externími podmínkami. V podtextu celé knihy zároveň stojí otázka, co rozumět pod termínem gramatika mluveného jazyka.

V teoretické části knihy autorka představuje dva základní kameny, na kterých by založila novou teorii gramatiky mluveného jazyka. Prvním je rozvinutý model blízkosti a distance (Koch & Oesterreicher, 1985), ve druhém, tzv. víceúrovňovém modelu (*Mehrebenenmodell*), čerpá

z Coseriuova rozlišování systému a normy (Coseriu, 1975). Rozpracovaný model<sup>63</sup> Kocha a Oesterreichera je založen na existenci kontinua mezi póly blízkost a distance. Tvoří ho pět hierarchicky uspořádaných rovin s jasně stanovenými vztahy. První rovina představuje rozdíl mezi veřejnou a uzavřenou (soukromou) produkcí a recepcí, která je vlastně přeformulováním pólů blízkosti a distance. Na druhé až čtvrté rovině jsou rozlišovány tzv. parametry (mohli bychom je označit za kritéria), které blíže charakterizují prototyp mezi oběma póly. Jedná se o roli produktora a recipienta, čas, prostor, situaci, médium a kód (pro detailní popis viz Hennig, 2006: 80–84). Na poslední rovině jsou přítomny konkrétní jazykové prostředky. Z tohoto modelu by však autorka pro teorii v gramatice mluveného jazyka vybrala pouze ty prostředky, které jsou typické pro jazyk blízkosti.

Víceúrovňový model, u kterého autorka čerpá od E. Coseriu, představuje pokus o zodpovězení otázky po vztahu gramatiky mluveného a psaného jazyka. Coseriu (1971) v návaznosti na Saussura rozlišuje trichotomii: funkční systém (*langue*) – jeho konkrétní realizace (*parole*) – dočasná norma, jejíž prvky nejsou v opozici, ale vzájemně na sebe navazují. Podle Coseriu je jazyk jako systém (*Sprache*) odvozen z konkrétního užívání, tj. z mluveného jazyka (*Sprechen*). Hennigová (2006: 111n.) toto rozlišení usouvztažňuje s primárními a sekundárními rozdíly mezi mluveným a psaným jazykem. Primárními míní rozdíly na úrovni systému, např. výskyt kontaktních prostředků v mluveném jazyce, k nimž nemá psaný jazyk žádnou systémovou opozici; sekundárními rozdíly založené na konvenci, tj. ustálené pro užívání určitých prostředků v mluveném a psaném jazyce, např. použití perfekta v mluvené a préterita v psané němčině. Pro systém i normu mluveného i psaného jazyka by se dala vytvořit vlastní gramatika, přičemž mluvený jazyk Hennigová v tomto modelu slučuje s jazykem blízkosti a psaný s jazykem distance. Gramatika jazyka jako celku by měla slučovat všechny tyto roviny (odtud i název tohoto modelu).

Oba modely autorka posléze aplikuje na data, konkrétně na přepis rozhlasového dialogu a rozhovoru na pracovním úřadě. K praktické ukázce vybírá problematiku syntaktické jednotky v mluveném jazyce. Ve výzkumu mluveného jazyka často zavrhaná věta je zde po reformulaci její definice použita jako základní jednotka. Hennigová definuje větu v mluveném jazyce přítomností finitního slovesa s naplněným valenčním rámcem; její součástí jsou zároveň i slova, která by při přeformulování do věty psaného jazyka<sup>64</sup> vypadla. Kromě věty navrhuje

---

<sup>63</sup> Tento model (viz 2.1.3) byl aplikován také na diachronní data (viz Ágel & Hennig, 2006).

<sup>64</sup> Tento postup, tzv. standardizace mluveného jazyka, byl zvolen při vytváření korpusu PDTSC (Mikulová & Urešová, 2008).

Hennigová další jednotky: elipsu, anakolut, struktury na pomezí věty (např. přístavkovou konstrukci) a diskurzní markery (*Nähezeichen*).<sup>65</sup>

Ve všech zde představených gramatikách je vyzdvihován paradox primárnosti mluveného jazyka a jeho nedostatečná deskripce včetně chybějícího pojmového aparátu. Hennigová usouvztažňuje ono časté nazírání na mluvený jazyk prizmatem psaného jazyka s nutností transkribovat mluvený jazyk, což je podmínkou pro jeho další zkoumání.<sup>66</sup> Pro teorii, na níž by byla založena gramatika mluveného jazyka, nachází ve stávajících teoriích řadu aplikovatelných konceptů včetně použitelné terminologie.

### 2.3 Klíčové aspekty výzkumu psané elektronické komunikace

Výzkum elektronické komunikace patří v současné době k populárním tématům. Tento zájem, který lze vysvětlit jednoduše jako nadšení pro vše nové a dosud neprobádané, se v zahraniční, především anglickojazyčné lingvistice projevoval už od 80. let minulého století (např. Baron, 1984; Chafe & Danielewicz, 1987), zatímco v české lingvistice lze jeho počátky spatřovat spíše v průběhu 90. let (např. Čmejrková, 1997).<sup>67</sup> Obliba tohoto tématu však kromě novosti souvisí i se stále se zvyšujícím zapojením internetu a elektronických komunikačních zařízení do běžného života a s celosvětovým zvyšováním počítačové gramotnosti.

Na internetu můžeme nalézt texty z různých oborů a komunikačních sfér; některé přitom mají tradiční tištěnou podobu a na internetu jsou zveřejněny ve stejné podobě (např. vyhlášky a zákony), jiné jsou určeny pouze pro internet a v jiné podobě se ani nevyskytují (např. diskusní příspěvky pod novinovými články).<sup>68</sup> Zájem lingvistů směřoval spíše k posledně zmíněným, mezi nimiž se vyskytují především nepřipravené, spontánně laděné texty reagující na konkrétní podnět v konkrétním čase, které bychom zařadili do sféry běžného dorozumívání. Jejich jazyková stránka budila pozornost nejen kvůli řadě inovací (např. akronymy častých frází), ale i kvůli jisté podobnosti s mluveným jazykem (pomineme-li rozlišování spisovného a nespisovného jazyka, patřil by sem např. zápis slov podle výslovnosti: *vždyť* vs. *dyť*) a větším začleněním obrazové složky do textu (např. emotikony). Debata o tom, zda je tento typ

---

<sup>65</sup> Tento pojem užíváme ve shodě s Čermákovou et al. (2019). Sdružuje širokou skupinu výrazů, které plní řadu funkcí, např. se podílejí na určení vztahu a zajištění kontaktu mezi komunikanty, organizaci dialogu.

<sup>66</sup> Detailnost přepisu je velmi důležitá, zejména je nutné doplnit informace dané mimojazykovým kontextem. Čím větší nepodobnost s psaným jazykem objevujícím se v knihách či novinách, tím menší riziko, že výzkumník zapomene, že se jedná o mluvený jazyk transkribovaný.

<sup>67</sup> Vzhledem k tomu budeme v této kapitole čerpat i z cizojazyčných prací. V charakteristice podle jazykových rovin pak budeme uvádět jen ty rysy, které jsou aplikovatelné i na češtinu.

<sup>68</sup> Existuje samozřejmě i třetí možnost: texty určené pouze pro internet, které se později dočkaly zveřejnění v tradiční tištěné podobě. V tomto případě se často jedná o umělecké texty (např. blog *Deník ostravaka*).

elektronické komunikace bližší psanému, či mluvenému jazyku, se promítla do jeho charakteristiky označením psaná mluvenost, mluvená psanost apod. (více pojmenování viz Jílková, 2017).<sup>69</sup> Ačkoli již toto téma není v oblasti elektronické komunikace centrem zájmu<sup>70</sup> a např. Baronová (2010a: 14) ho považuje za uzavřené (s výsledkem, že elektronická komunikace je bližší mluvenému než psanému jazyku<sup>71</sup>), dotýká se širší problematiky žánrové pestrosti, jež je v elektronické komunikaci stále živá. Dokládají to četné práce o různých žánrech včetně jejich spojení do tzv. hybridních žánrů (např. blog: Biber & Egbert, 2016; chat: Jandová et al., 2006; IM: Baron, 2004, 2010b; e-mail: Baron, 1998) nebo o elektronické komunikaci zprostředkované různými nástroji v různých podmínkách (např. komunikace prostřednictvím mobilních telefonů: Deumert, 2014, elektronická komunikace na pracovišti: Scott & Timmerman, 2005).

Jazyk konkrétních textů je determinován nejen jejich (a)synchronností, ale i počtem adresátů (zpravidla se rozlišuje komunikace mezi jedním původcem a jedním adresátem vs. jedním původcem a vícero adresátů)<sup>72</sup>, konkrétní komunikační platformou, tj. prostředím, v němž probíhá, a mírou, s jakou je pisatel s touto platformou seznámen, veřejností/soukromostí<sup>73</sup> komunikace a dalšími faktory, které ovlivňují jakoukoli mezilidskou komunikaci.<sup>74</sup> V této kapitole se budeme věnovat jazykové charakteristice neformálních psaných elektronických textů, mimoto však výběrově zmíníme i nejazykové prostředky typické pro tuto komunikaci (k průběhu komunikace v dále analyzovaných typech textů viz 3.1.1). Pro takovou elektronickou komunikaci považují Thurlow et al. (2004: 124) za klíčové dva faktory: rychlost komunikace a neformálnost, z nichž plyne snaha o kreativní užití psaného jazyka s využitím co nejnižšího počtu znaků, resp. úhozů do klávesnice, a zároveň rezignace na některá gramatická pravidla či zvyklosti.

---

<sup>69</sup> Podobně je tomu i v angličtině, srov. Herringová (2011).

<sup>70</sup> Baronová (2010a: 28n.) nabízí stručné časové shrnutí základních výzkumných témat elektronické komunikace. Na konci 80. a na počátku 90. let 20. století stála v centru pozornosti otázka, zda je elektronická komunikace novou formou jazyka (zájem se soustředil např. na akronymy, zkratky, emotikony). V 90. letech se debata posunula k otázce psanosti a mluvenosti elektronické komunikace, která byla podle Baronové vyřešena ve prospěch mluveného jazyka. Na přelomu tisíciletí byl zájem věnován vlivu elektronické komunikace na psaný jazyk.

<sup>71</sup> K témuž názoru se kloní Herringová (2011). Charakter spíše mluveného (*conversation-like*) jazyka mají synchronní, ale i asynchronní elektronické texty.

<sup>72</sup> Občas se i v češtině vyskytují anglické termíny *one-to-one* (komunikace mezi jedním autorem a jedním adresátem) a *one-to-many* (jeden autor, vícero adresátů).

<sup>73</sup> Veřejné texty bývají cenzurovány správcem daného prostředí podle zveřejněných pravidel. Mazány jsou např. ty, které obsahují vulgární výrazy nebo mají rasistický obsah.

<sup>74</sup> Vliv konkrétního elektronického prostředí a situačního kontextu na jazyk elektronické komunikace ukazuje Herringová (2007).

### 2.3.1 Charakteristiky podle jazykových rovin

V této kapitole představíme jazykové prostředky charakteristické pro příslušnou jazykovou rovinu. Vzhledem k novosti tématu budeme poznatky čerpat nejen z českých, ale i ze zahraničních publikací; z posledně zmíněných budeme vybírat pouze ty, které podle našeho názoru nejsou vázány na konkrétní jazyk.<sup>75</sup> Nebudeme se přitom omezovat pouze na práce o žánrech analyzovaných v druhé části práce (tj. internetové diskuse, diskusní fóra, veřejné facebookové statusy), a to ze dvou důvodů: 1) k většině z nich byly publikovány spíše kusé poznámky nebo nebyly jazykově analyzovány vůbec, 2) v odborných publikacích panuje přesvědčení o typických jazykových prostředcích neformální elektronické komunikace, jež se vyskytují (v různé míře) téměř ve všech žánrech. K jazykovým rovinám přidáváme rovinu grafickou, pod kterou sdružujeme jevy projevující se ve vizuální percepci textů.

#### 2.3.1.1 Fonetická rovina<sup>76</sup>

Tuto rovinu zařazujeme do tohoto přehledu právě kvůli výše zmíněné debatě o blízkosti psané elektronické komunikace k mluvenému jazyku. Za typický rys elektronické komunikace je považována fonetizace pravopisu,<sup>77</sup> tj. zápis slov podle výslovnosti, zohledňující především redukce, např. *ňáký* místo *nějaký*, *dycky* místo *vždycky*, *de* místo *jde*,<sup>78</sup> emfatické dloužení, např. *jó*, *bóže*, ale i krácení vokálů, např. *prosim*, asimilace, např. *gdo*, a zápis protetického *v*. Nestandardní podoby slov se týká iterace (multiplikace) grafémů, zejména takových, které se dají prodloužit i v řeči (tj. vokálů, např. *ahoooj*, a souhlásek trvacích, např. *neviššš*, *mmmmňam*) připisovaná Čmejrkovou (1997: 244), Hladkou a Šmilauerem (2006), Schnitzerovou (2012) i Jílkovou (2017) nápodobě intonace, odrazu citového rozpoložení a nálady pisatele. Těmto důvodům je připisována i iterace interpunkčních znamének (konkrétně teček,<sup>79</sup> vykřičníků a otazníků) nebo jejich kombinace, např. *?!.*

---

<sup>75</sup> Ne všechny jevy uváděné jako typické pro elektronickou komunikaci jsou univerzální; naopak řada z těch, které pracují s kreativní typografií je vázána na jednotlivé jazyky (Herring, 2012). O konsenzu týkajícím se typičnosti zmiňovaných jevů pro elektronickou komunikaci svědčí např. souhrnné práce Hårdové af Segerstad (2002) či Jílkové (2016, 2017).

<sup>76</sup> Tento název byl zvolen kvůli zachování jednotnosti s oddílem 2.2.1.1. Zařazujeme sem rysy, které jsou spojeny s mediální mluveností, tedy které přenášejí např. výslovnost a suprasegmentální jevy do psané podoby. Naopak Herringová (2012) tuto rovinu vynechává s poukazem na mediální psanost; funkce vyjadřované fonicky přebírají typografická a ortografická rovina (srov. grafická rovina v této práci).

<sup>77</sup> Nestandardní podoba anglických slov, zahrnující různé typy zkracování a přiblížení reálné výslovnosti, se podle Kapidzice (2010) etablovala napříč různými žánry elektronické komunikace pouze u někalika málo, např. *u* pro *you*, *wanna* pro *want to*, *msg* pro *message*.

<sup>78</sup> V angličtině se redukce týkají např. stažených forem zájmena + slova *být*, tj. *I'm* místo *I am*. Baronová (2004: 413) v datech z IM zjistila, že pouze 65 % ze všech případů, kdy mohlo k této redukci dojít, bylo skutečně redukováno.

<sup>79</sup> Hoffmannová (2016: 58) popisuje tři tečky jako typickou součást e-mailů a SMS; jejich funkcí je dělení výpovědi nebo textových jednotek.



Vliv mluvenosti na psanou podobu textu můžeme pozorovat při dělení slov, resp. při zápisu více slov, která lze vyslovit v jednom mluveném taktu, do jednoho celku, např. *tyjo, nachvilku* (např. Svobodová, 2006). Na úrovni celého textu můžeme tomuto vlivu připisovat nečlenění textu do vět oddělených koncovou interpunkcí, nápodobu mluveného proudu řeči v jedné dlouhé „větě“ (někdy včetně absentující interpunkce). Tento rys by se dal považovat také za nápodobu lineárního průběhu mluveného jazyka. Lineárnost se však nejvíce projevuje v přítomnosti překlepů a odchylek v textové návaznosti, které dokládají, že pisatelé nevyužívají možnost editace textu (Čmejková, 1997). Výjimkou samozřejmě nejsou ani záměny grafémů způsobených rozdílným umístěním grafémů na různých klávesnicích (zejména *z* a *y*) (Jílková, 2017). Výskyt překlepů, odchylek od pravopisných pravidel a norem psaného textu jsou samozřejmě dány i dalšími důvody: nepozorností při psaní, mírou pisatelovy dovednosti psát na klávesnici, akcentem na rychlost komunikace, nechutí k opravám apod. (Hladká & Šmilauer, 2006: 155).

Často zmiňovaným rysem je dále výskyt krátkých, zpravidla jednoslabičných, nepřízvučných slov na začátku věty či např. repliky na chatu (např. Hoffmannová & Müllerová, 2007; Hoffmannová, 2016; Jílková, 2017). Tento jev pozorovali Jandová et al. (2006) v chatové komunikaci také u postfixu *-s*, např. *morrisek: S nějak dlouho nepsal?*

### 2.3.1.2 Morfologická rovina

Na rozdíl od mluveného jazyka se u elektronické komunikace (a jejích žánrů) slovním druhům dosud věnovala pozornost spíše okrajově. Krčmová (1995: 29) zjistila na materiálu e-mailů<sup>80</sup> vyšší frekvenci sysémantik oproti autosémantikům.

Více pozornosti se dostalo flektivní morfologii (i vzhledem k otázce spisovnosti a vlivu mluveného jazyka). Za jevy typické pro komunikaci na chatu označili Jandová et al. (2006) tyto: koncovku *-(a)/(e)ma* v I pl. podstatných a přídavných jmen, kondicionálový tvar 1. os. sg, a pl. slovesa *být bysem* a *bysme*, vynechávání koncového *-l* v minulém participiu, vynechávání koncového vokálu u přítomné tvary 3. os. pl. od sloves 5. třídy (např. *čekaj*),<sup>81</sup> náhradu pomocného slovesa v 1. os. sg. v minulém čase osobním zájmenem. Autorka si dále všimla nízké frekvence plurálu, způsobené vzájemným tykáním všech chatujících.

---

<sup>80</sup> Autorka porovnávala e-maily a mluvený projev od téhož člověka. Rozsah jejích dat nebyl ve studii specifikován.

<sup>81</sup> Pro německou elektronickou komunikaci je typický obdobný rys – vynechávání slovesné koncovky, např. *ich hab* místo *ich habe* (Schitzer, 2012).

### 2.3.1.3 Lexikální rovina

Lexikální rovina je v souvislosti s elektronickou komunikací reflektována českými lingvisty často s ohledem na anglicismy (např. Lišková, 2006; Mlčoch, 2005; Svobodová, 2006), které jsou jednak součástí elektronického prostředí, jednak se objevují přímo v elektronických textech (v původní i v počeštěné podobě). Pozornosti se obdobně jako u mluveného jazyka nevyhnuly ani vulgarismy, jejichž používání je ve veřejných elektronických textech regulováno či zcela zakázáno.<sup>82</sup> Uživatelé vulgarismů se v obraně před smazáním svých textů nebo svého účtu snaží o různé maskování, např. opisem daného výrazu, nahrazením pomocí tří teček nebo nahrazením jednoho grafému ve vulgarismu pomocí hvězdičky (např. *ku\*va*), což Hoffmannová (2006) nazývá grafickou eufemizací.

Dalším výrazným lexikálním rysem je zkracování, jež souvisí s potřebou vyjádřit se rychle a úsporně (zejména v textech omezených počtem znaků, např. SMS, tweet). Kromě ustálených zkratk a univerbizovaných výrazů se v elektronické komunikaci vyvinuly nové zkratky z často užívaných českých i anglických slov, např. *pls* pro *please*, *thx* pro *thanks*,<sup>83</sup> *ptž/pže* pro *protože*, *jj* pro *jo jo*. Princip zkracování vedl také k vytvoření akronym, tj. zkratk vzniklých z iniciálních písmen určité fráze. V české elektronické komunikaci se používají stejně jako u zkratk akronyma anglická i česká. Podle Komrskové (2015) i Jílkové (2017) je nejčastějším akronymem v českých textech *BTW* (z angl. *by the way*), mezi dalšími jsou např. *IMHO*, *LOL*, *OMG*, *WTF*. Ke zkracování jsou využívány též číslice, resp. jejich zvuková podoba, jíž se nahradí slovo nebo jeho část, např. *4 you* pro *for you*, *z5* pro *zpět*. (např. Jílková, 2017; Hladká et al., 2005; Tagliamonte & Denis, 2008; Varnhagen et al., 2010).

V chatu si Svobodová (2006) povšimla častého opakování slov a frází, které zdůvodňuje nepřipraveností a bezprostředností komunikace. Naopak zopakováním celých replik zvyšuje pisatel jejich naléhavost, snaží se přitáhnout pozornost či vzbudit emoce.

### 2.3.1.4 Syntaktická rovina

Na syntaktické rovině se zřejmě nejvýrazněji projevuje vliv konkrétního typu elektronické komunikace, zejména délky textu a míry interaktivity (např. synchronní texty obsahují spíše kratší věty).<sup>84</sup> V řadě rysů, projevujících se nejen na syntaktické, ale např. i na lexikální rovině,

---

<sup>82</sup> Bývá součástí pravidel komunikace v příslušném elektronickém prostředí. Tento soubor pravidel se někdy označuje jako netiketa. K užívání vulgarismů v diskusních fórech viz např. Gromnica (2017).

<sup>83</sup> Tato zkratka využívá anglické výslovnosti. Podobně je tomu např. u akronymu *CU* pro anglický pozdrav *see you* a zkratk užívajících číslici místo (části) slova.

<sup>84</sup> Např. Jandová et al. (2006) uvádí, že pro chat jsou typické krátké repliky vytvářené v krátkém sledu, mnoho z nich je tvořeno pouze zpětnou vazbou typu *jj*, *aha* nebo emotikonem a u řady z nich nejsou patrné hranice vět.

se elektronická komunikace neliší od mluveného jazyka ve sféře běžného dorozumívání. Týká se to např. elipsy a redukcí, které jsou podle Hoffmannové (2016: 57) umožněny těsnou vazbou na sdílenou situaci a společné znalosti účastníků. O elipse autorka uvažuje v souvislosti se zdvořilostními frázemi (*dobrou, máš se?*) a v iniciačních replikách dialogických textů (*OK, budu tam*).<sup>85</sup>

Neukončené a syntakticky nestrukturované repliky jsou typické pro synchronní elektronickou komunikaci, neboť jejich výskyt je ovlivněn její rychlostí, příp. nápodobou mluvené syntaxe (např. Jandová et al., 2006; Herring, 2012; Schnitzer, 2012). Syntaktickou nestrukturovanost v e-mailu usouvztažňuje Krčmová (1995) s asociativním řazením myšlenek a absencí členění textu do odstavců.

Za syntaktickou specifičnost považuje Herringová (2012) výskyt performativních výrazů v 3. os. sg. přítomnosti, tzv. *emotes*, které se zobrazují adresátovi zejména při chatování. Tato performativa vyjadřují akci, kterou pisatel vykonává, např. *Míša píše...*, *Tonda23 odešel z místnosti*.

### 2.3.1.5 Grafická rovina

Všeobecně se soudí, že vizuální kanál slouží ke kompenzaci neverbálních jazykových prostředků, zejména zvukových vlastností. Pomineme-li změny velikosti, barvy či stylu písma, které jsou spíše odrazem pisatelovy individuality, jeví se jako sémioticky významné a pro elektronickou komunikaci typické užívání kapitálek, iterace grafémů či interpunkce (viz výše) a emotikonů. Tyto jevy mají jednak přitáhnout ke sdělení pozornost, oživit ho, příp. zdůraznit (a zvýraznit) jeho části, např. užívání kapitálek v chatu je někdy interpretováno jako hlasitá mluva až křik, což může být podle Jandové et al. (2006) důvodem pro vyloučení pisatele z chatovací místnosti. Pro elektronickou komunikaci jako celek platí nedodržování zvyklostí v psaní velkých písmen, tj. absence velkých písmen u vlastních jmen nebo na začátku vět, aniž by to bylo chápáno jako silný prohřešek (např. Svobodová, 2006: 152).<sup>86</sup>

---

V korpusu IM zkoumaného Baronovou (2010b) činila průměrná délka odeslané zprávy 5,4 slov; 22 % všech zpráv bylo tvořeno jedním slovem. Délkou vět měřenou ve slovech se zabývala Krčmová (1995). V osobních e-mailech se délka úseků ohraničených jakoukoli interpunkcí pohybovala kolem 5 slov, v úsecích mezi koncovou interpunkcí se jednalo o 10–13 slov. Na korelaci mezi délkou textu a elektronickým přístrojem, na kterém byla napsána, poukázala Schnitzerová (2012) u facebookových statusů: při použití mobilu byly texty kratší než při psaní na počítači.

<sup>85</sup> Elipsa subjektového zájmena je podle Bibera a Conradové (2005: 194) typická pro anglickou elektronickou komunikaci v diskusních fórech. Výskyt tohoto jevu je ovlivněn blízkostí s běžným neformálním dialogem.

<sup>86</sup> Tato tendence se objevuje i v angličtině, např. v zápise zájmena *I* jako *i* (Tagliamonte & Denis, 2008).

Emotikony patří k nejvýraznějším prostředkům elektronické komunikace na grafické rovině, o čemž svědčí i jejich široká variabilita (např. Komrsková, 2014) a pronikání do psané komunikace, např. tištěných textů nebo ručně psaných dopisů (Jílková, 2016). Emotikony mohou plnit různé funkce v závislosti na typu (typografický, grafický, se znakem pro nos apod., viz např. Dresner & Herring 2010) a jeho pozici v textu (např. Skovholt et al., 2014; Komrsková, 2016), nejčastěji se jedná o expresi pocitů a psychického stavu pisatele, příp. chování spojeného s emocemi, nebo o strukturaci textu. Od emotikonů se odlišují obrázky složené z typografických symbolů, např. @}--‘,--- (Herring, 2012).

V elektronické komunikaci na chatu se oproti jiným psaným textům častěji vyskytují grafémy *f*, *g*, *q*, *x*, *y*, *w* (Svobodová, 2006). První dva se užívají na místě svých znělostních protějšků za účelem simulace výslovnosti, např. *f pohodě* nebo *fpohodě, gde*. Použití *q* a *x* je motivováno ozvláštňením psaného textu a redukcí souhláskových skupin: *-kv-* je nahrazeno grafémem *q* (např. *sqělý* místo *skvělý*, *qám* místo *k vám*), *-ks-* grafémem *x*<sup>87</sup> (*taxi* místo *tak (j)si*, *xicht* místo *ksicht*). Grafémy *y* a *w* jsou často pouhým nahrazením grafémů *i* a *v*, které je jedním z prostředků hry s jazykem, resp. s jeho psanou podobou.

K vnitřnímu členění elektronického textu uživatelé používají interpunkční znaménka mnohdy s výrazným rozdílem oproti pravidlům jejich užívání v tradičních psaných textech. Jandová et al. (2006) si na datech z chatu povšimli řídkého užívání čárek; ke strukturaci repliky slouží otazník, vykřičník (příp. jejich kombinace nebo iterace), tři tečky nebo emotikony. Tyto symboly se taktéž užívají jako koncová interpunkce, tečka samotná je spíše vynechávána. Její elize je pochopitelná: hranice repliky je v prostředí chatu dostatečně graficky výrazná, uživatel signalizuje konec repliky jejím odesláním ostatním komunikantům.

### 2.3.2 Směry a přístupy zabývající se elektronickou komunikací

Ve výzkumu elektronické komunikace se kříží různé lingvistické směry a přístupy s příbuznými humanitně či technicky orientovanými obory, např. pedagogikou (např. Paulus & Phipps, 2008), psychologíí (např. Anolli et al., 2002; Joinson et al., 2007), sociologií (např. Orton-Johnson & Prior, 2013), počítačovým zpracováním přirozeného jazyka (např. Erjavec et al., 2016). Většina výzkumů by se proto vzhledem k vlastnostem elektronických textů a prostředí dala označit za interdisciplinární. Pro naši práci jsou – s odhlédnutím od jazykových prostředků

---

<sup>87</sup> Grafém *x* má širší využití. Svobodová (2006) uvádí i případy nahrazení souhláskové skupiny *-kš-* (*xěfty* místo *kšěfty*) a hlásky *ch* (*xjo* místo *chjo*).

zmíněných v předchozí části – relevantní zejména sociolingvistické výzkumy týkající se vlastností autora elektronického textu a aplikace metod KA a MDA na elektronické texty.

Zřejmě nejvýznamnějším sociolingvistickým<sup>88</sup> tématem jsou genderové rozdíly (viz přehled takto orientovaných studií v Herring, 2000), přestože v některých žánrech elektronické komunikace není reálné pohlaví pisatele nikdy stoprocentně jisté. Zjištění o shodném, či rozdílném chování a vyjadřování obou pohlaví je relevantní např. pro lepší zacílení pracovních inzerátů, reklam a optimalizaci komunikačních systémů (např. Kim et al., 2007), zajištění vyšší efektivity při učení (např. Rovai & Baker, 2005). Otázka genderu je také zkoumána v souvislosti s textovým obsahem; např. Burgstahlerová a Doyleová (2005) analyzovaly obsah e-mailů psaných hendikepovanými adolescenty jejich vrstevníkům nebo vyučujícím. Studie ukázala, že muži se více zajímají o techniku a internet, zatímco ženy se raději zajímají o osobní záležitosti a posílají více e-mailů.

Zájem o gender se dále projevuje se zřetelem ke stereotypu připisujícímu ženám častější vyjadřování emocí a vyšší expresivitu. Jak ukazuje např. Foxová et al. (2007), důležité je kromě pohlaví autora i pohlaví adresáta. Tento autorský kolektiv sice potvrdil, že ženy se při komunikaci prostřednictvím IM vyjadřují expresivněji (měřeno podle počtu emotikonů, verbalizovaného smíchu, expresivních adjektiv a různých prostředků zdůrazňování v jednotlivých chatových replikách), ale je-li adresátem žena, je text nezávisle na pohlaví pisatele expresivnější a obsahuje méně slov. Ve výzkumu emotikonů, vnímaných primárně jako prostředky pro vyjadřování emocí, existuje k genderovým rozdílům mnoho studií vytvořených na různých datech (např. internetové diskuse: Wolf, 2000; blogy: Huffaker & Calvert, 2005; e-mail: Eimler et al., 2012; SMS: Tossell et al., 2012), z nichž se kromě adresáta ukazuje jako podstatné téma (emotikony jsou používány spíše v komunikaci zaměřené na mezilidské vztahy), např. Derks et al. (2007).

Aplikaci konverzačněanalytického přístupu ukazují např. Jandová et al. (2006) při analýze jednotlivých fází komunikace na chatu. Iniciální a finální repliky obsahující pozdravy, příp. oslovení,<sup>89</sup> se podobají neformálnímu mluvenému dialogu jak svou strukturou, tak lexikálním obsazením. Autoři se dále zaměřili na rozdíly při dobrovolném a nedobrovolném opuštění komunikace, což je chatové specifikum související s pravidly každé chatové místnosti, a

---

<sup>88</sup> Jinou ukázkou sociolingvistiky orientovaných výzkumů v oblasti elektronické komunikace představuje Androutsopoulos (2006).

<sup>89</sup> Podle Peréz-Sabaterové (2012) se pozdravy vyskytují v 80 % příspěvků zveřejněných na oficiálním Facebooku vybraných evropských a amerických univerzit.

tematické výstavbě dialogu, resp. prolínání různých tematických bloků skrz repliky několika zároveň chatujících lidí. Průběh komunikace na chat vykazuje podle autorů řadu podobností s mluveným dialogem.

Jinou aplikaci KA představuje tzv. mikroanalýza online dat (zkráceně MOOD z ang. *microanalysis of online data*), která se označuje také jako digitální KA (Giles et al., 2015). Styčnými body s tradiční KA je práce s přirozenými daty (tj. daty nevytvořenými pouze pro účely výzkumu) a zájem o totožné jevy, např. průběh interakce, řešení problému. Souhrn odborných publikací aplikujících KA na elektronickou komunikaci (Paulus et al., 2016) dokládá rozšířenost tohoto přístupu. Za téma specifické pro digitální KA bychom mohli považovat i komparaci mluveného jazyka s elektronickou komunikací, jehož vývoj je krátce shrnut v poslední zmíněné studii (Paulus et al., 2016: 5).

Výzkum jazykové variability prostřednictvím MDA umožnil srovnání některých žánrů elektronické komunikace s dalšími typy textů. Jazykovou stránku osobních e-mailů, diskusních fór a SMS ovlivňují podle Bibera a Conradové (2005) relativně malé rozdíly v podmínkách produkce, vztahu mezi komunikanty a účelu komunikace. Podobnost s mluveným jazykem se u těchto tři žánrů týká zejména frekvence sloves a adverbíí, délky klauzí a frekvence zájmen pro 1. a 2. osobu.

Výzkum MDA v češtině (Cvrček et al., 2018b) ukazuje odlišnosti mezi žánry elektronické komunikace lišící se mírou interaktivity, tj. zda umožňují a předpokládají dialog pisatele s adresátem, nebo nikoli. V tomto ohledu vykazuje elektronická komunikace podobnost s mluveným jazykem (obr. 3); interaktivní žánry, tj. internetové diskuse, diskusní fóra a facebookové příspěvky, se blíží interaktivnímu mluvenému jazyku, zatímco neinteraktivní blogy a elektronická encyklopedie Wikipedie tradičním psaným textům (různého typu podle detailnějšího zaměření). Nejbliže si je interaktivní mluvený jazyk a interaktivní elektronická komunikace v první a druhé dimenzi (obr. 2), které je označují za spíše dynamické a spontánní než statické a připravené. Pro interaktivní elektronickou komunikaci je dále charakteristická postojovost (v MDA prezentovaná osmou dimenzí), tj. zaměření na vyjádření subjektivního názoru a postoje k určitým faktům. Celkově lze říci, že interaktivní žánry elektronické komunikace stojí na pomezí mezi interaktivním mluveným jazykem a psaným jazykem beletrie, žurnalistiky a oborové literatury, což z pohledu jednotlivých dimenzí MDA znamená, že se více či méně překrývají, aniž by došlo k výraznému odlišení. (Více viz Cvrček et al., 2018b, v tisku.)

## 3. Analytická část

### 3.1 Data a metodologie

#### 3.1.1 Popis dat

Datovou základnu analytické části práce tvoří korpusy zpřístupněné Českým národním korpusem (dále jen ČNK) přes rozhraní KonText. Pro srovnání mluveného a psaného elektronického jazyka byly použity nejnovější korpusy, které shromažďují daný typ komunikace, tj. korpus Ortofon (Kopřivová et al., 2017b) pro mluvený jazyk, korpus Koditex-web (Zasina et al., 2018) pro psanou elektronickou komunikaci.

##### 3.1.1.1 Korpus Ortofon

Korpus Ortofon obsahuje 1 236 508 tokenů (bez interpunkce, hezitačních a citoslovených zvuků a komentářů jde o 1 014 786 tokenů, tj. textových slov), dále lemmatizaci a morfologické značkování. Korpus zahrnuje transkribované nahrávky z let 2012–2017, které zachycují nepřipravenou, spontánní mluvu používanou v neformálních situacích mluvčími, kteří se dobře znají.<sup>90</sup> Korpus je vyvážený s ohledem na čtyři sociolingvistické kategorie: pohlaví (žena vs. muž), věk (18–34 let vs. 35 let a více), vzdělání (základní a středoškolské vs. vysokoškolské), oblast pobytu mluvčího do 15 let (česko-moravská, jihočeská, severovýchodočeská, slezská, středočeská, středomoravská, východomoravská, západočeská, pohraničí české, pohraničí moravské a slezské); cílem bylo zachytit mluvený jazyk od co největšího počtu mluvčích z co největšího území ČR, aby byl omezen vliv idiolektu (více viz Komrsková et al., 2017).

Korpus Ortofon byl jako zdroj dat vybrán přesto, že není největším korpusem zachycujícím daný typ mluveného jazyka.<sup>91</sup> Prvním důvodem pro použití korpusu Ortofon je stejné časové rozpětí, ze kterého pochází nahrávky i srovnávaná elektronická komunikace. Dalším důvodem je dvouúrovňová transkripce těchto nahrávek (tj. základní ortografický přepis a zjednodušený fonetický přepis), která je pro vyhledávání a následné analyzování vybraných jazykových jevů výhodnější než jednoúrovňová transkripce v jiných mluvených korpusech ČNK. Více viz k transkripci Kopřivová et al. (2014).

Základní jednotkou v tomto korpusu není replika, ale segment, který je spíše technickou než lingvistickou jednotkou. Segment je tvořen maximálně 25 slovy, přičemž jeho hranice je vedena především v pauzách (př. 1). Souvislost segmentů s jakýmkoli lingvistickými předěly je

---

<sup>90</sup> Poloviční podíl všech slov pochází z hovoru nahraného doma (612 572 slov). Následuje hovor při návštěvě, neformální hovor v zaměstnání nebo o přestávce ve škole, v restauraci a při posezení na zahradě či chatě.

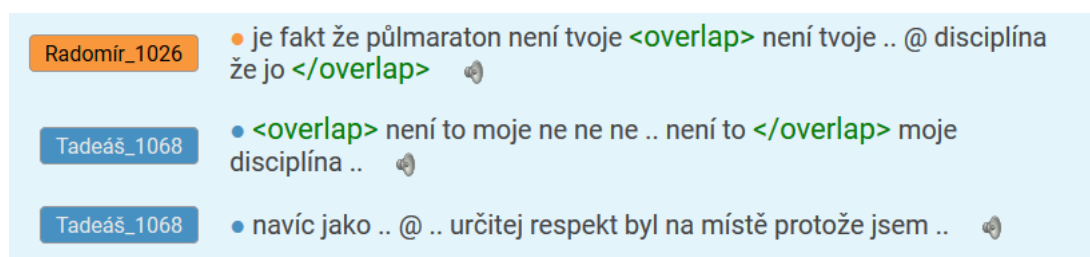
<sup>91</sup> Tím je korpus ORAL (Kopřivová et al., 2017a) s 5,3 mil. textových slov.

nesoustavná a nelze na ni spoléhat. Součástí přepisu je i značení překryvů<sup>92</sup>, což jsou situace, v nichž mluví více mluvčích zároveň (př. 1). V KonTextu je možné zobrazit dialog lineárně (tj. všechny segmenty za sebou jako v psaném textu) nebo s graficky oddělenými segmenty (obr. 4).

(1)

Radomír\_1026: je fakt že půlmaraton není tvoje [není tvoje .. @ disciplína že jo]

Tadeáš\_1068: [není to moje ne ne ne .. není to] moje disciplína .. navíc jako .. @ .. určitej respekt byl na místě protože jsem ..



Obr. 4 Zobrazení segmentů v KonTextu

### 3.1.1.2 Korpus Koditex-web

Pro analýzy psaného elektronického jazyka byla použita část korpusu Koditex (Zasina et al., 2018) zahrnující pouze data z internetu (odtud název Koditex-web), z nichž byly využity jenom tři žánry (viz dále).<sup>93</sup> Korpus Koditex-web obsahuje 1 003 517 textových slov (tj. bez interpunkce), je lemmatizován a morfologicky značkován.<sup>94</sup> Koditex-web je rozdělen do pěti přibližně stejně velkých subkorpusů (cca 200 000 textových slov) podle zdroje dat: Wikipedie, blogy, internetové diskuse,<sup>95</sup> internetová fóra,<sup>96</sup> Facebook.<sup>97</sup> První dva zdroje představují komunikaci jednosměrnou, probíhající od autora k adresátovi bez možnosti přímé interakce obou účastníků. Na produkci textu se podílí pouze jeden autor, dalo by se proto hovořit o textech monologických (Jílková, 2016). Další tři zdroje naopak zachycují komunikaci vícesměrnou; autor je zároveň adresátem, na budování komunikace se podílejí všichni.<sup>98</sup> Všechny texty

<sup>92</sup> V příkladech značen pomocí []. V KonTextu je překryv značen jako strukturní atribut <overlap> </overlap>.

<sup>93</sup> Korpus Koditex obsahuje texty ze tří modů: psaného (wri), mluveného (spo) a internetového (web). Každý je dále rozdělen na menší skupiny, více viz Cvrček et al. (2018b).

<sup>94</sup> Pro více informací viz <http://wiki.korpus.cz/doku.php/cnk:koditex>.

<sup>95</sup> Data pocházejí z roku 2015.

<sup>96</sup> Data pocházejí z roku 2015.

<sup>97</sup> Data pocházejí z let 2014–2015.

<sup>98</sup> Průběh komunikace v internetových diskusích, diskusních fórech a na Facebooku probíhá víceméně obdobně (srov. Biber & Conrad, 2009: 190n.; Jílková, 2016). Pro přidávání příspěvku je většinou nutné se přihlásit jako registrovaný uživatel (podmínky přidávání příspěvků závisí na konkrétním komunikačním prostředí). Uživatel může založit nový příspěvek nebo zareagovat na nějaký, který zveřejnil jiný uživatel před ním; tato odlišnost je zpravidla prezentovaná graficky (srov. lineární zobrazení v korpusovém manažeru KonText).



v korpusu Koditex-web patří mezi asynchronní. Vzhledem k zaměření této práce na porovnání textů charakterizovatelných jako dialogy, budeme se v následujícím popisu i analýze věnovat pouze vícesměrné elektronické komunikaci.<sup>99</sup>

Korpus Koditex byl sestavován ze vzorků textů tak, aby obsáhl co nejvíce textů od co největšího počtu autorů. Vzorky však musely mít 2000–5000 slov, což je mnohonásobně více než v běžném příspěvku jednoho autora v diskusním fóru, internetové diskusi nebo na Facebooku. Z toho důvodu byly všechny příspěvky od téhož autora (tj. od autora s týmž jménem) v rámci daného žánru elektronické komunikace sloučeny a z těchto agregátů byly do korpusu náhodně vybrány ty, které měly požadovanou délku (podrobněji viz Cvrček et al., 2018a). Tento přístup sice není vhodný pro sledování tematické návaznosti příspěvků, střídání jednotlivých autorů, ale umožňuje srovnávat individuální charakteristiky autorů. O autorovi je znám pouze jediný údaj – jméno (resp. přezdívka).<sup>100</sup>

Všechny elektronické texty byly do korpusu Koditex-web zahrnuty bez jakékoli editace při korpusovém zpracování. To v případě dále zkoumaných elektronických dat znamená, že obsahují všechny nedostatky, které může needitovaný text mít, např. špatně rozdělená slova, překlepy, chybnou interpunkci. Pravděpodobnost výskytu těchto jevů je např. u Wikipedie nižší, protože texty v ní procházejí procesem schvalování administrátorem (který pravděpodobně zahrnuje i korekturu). V některých kvantitativních analýzách (např. při porovnávání frekvence výskytů slova nebo slovního druhu) může tento fakt ovlivňovat výsledky. Další omezení se týká internetových diskusí a diskusních fór: v některých textech zůstaly zachovány technické údaje ohraničující začátek a konec jednotlivých příspěvků, ze kterých je také možné vyčíst jméno autora a adresáta (vyznačeného pomocí @) a podle daného zdroje i např. datum a čas zveřejnění (př. 2 a 3). V takových případech by bylo – byť v omezené míře – možné sledovat návaznost příspěvků a interakci komunikantů.

(2)

**Jana odpovídá : Zář 12 th , 2013 v 20.21** Leni , ale skutečně školila . Ještě když jsem byla u ní já a to jsem byla u ní mezi první desítkou tenkrát začínala s tím na Václaváku tak přímo na mně to ukazovala několika mladým lékařkám a bavily se o tom . Však dr . Vybulková je jedna z jejích žákyň :) **Lenka odpovídá : Zář 11 th , 2013 v 22.13** Ahoj Jani , pochybuji , že dr . Něm někoho ještě školí .

(fóra, forums\_65406)<sup>101</sup>

---

<sup>99</sup> Přesnější označení použitého subkorpusu by tedy vzhledem k jeho složení bylo Koditex-web-mul.

<sup>100</sup> U každého textu jsou další metatextové údaje, např. část dne, kdy došlo k zveřejnění, počet slov. Více viz <http://wiki.korpus.cz/doku.php/cnk:koditex>.

<sup>101</sup> U příkladů z korpusu Koditex-web uvádíme vždy žánr, z něhož příklad pochází (diskuse, fóra, Facebook). Přesné označení části textu, tzv. ID chunku, předkládáme pouze u delších příkladů.

(3)

@ jezzz píše : @ verusak.berusak Vidis , ja bych rekla , ze je dite hodne vysoke na svuj vek Ale to uz je slovickareni a do teto diskuze to nepatri Dite by v roce a pul melo zakladnim vecem rozumet . Pokud nerozumi , resila byc to s pediatrou , na neurologii ( nebo jinde , kde maji takove to prcky na starosti ) . Ditko muze treba spatne slyset a z toho pak vznikaji komplikace . Nebo to muze byt neco jineho , posoudit by to mel odbornik . @ verusak.berusak píše : @ stock @ jezzz holky , myslím , že takto to vůbec nebylo myšleno . Standardní = běžný . Nestandardní dítě reaguje / se chová / má jiné potřeby , než je obvyklé u naprosté většiny dětí

(fóra, forums\_11242)

### 3.1.2 Metodologie

#### 3.1.2.1 Výběr jazykových jevů pro analýzu

Jazykové jevy analyzované v následující části práce byly vybrány ze dvou zdrojů. Prvním je MDA provedená na češtině (Cvrček et al., 2018b), druhým dostupná literatura věnující se elektronické komunikaci.

Z MDA jsme se inspirovali druhou (rozdíl mezi spontánními vs. připravenými texty) a pátou dimenzí (vyšší vs. nižší míra explicitní adresnosti). Tyto dimenze jednak patří mezi čtveřici dimenzí, které postihují největší množství variability (viz 2.2.2.3), jednak cílí na neformálnost a dialogičnost, což jsou charakteristiky společné mluveným a psaným datům analyzovaným v této práci. Ačkoli nejsou tyto dvě charakteristiky explicitně zmiňovány v pojmenování jednotlivých dimenzí MDA, lze je konceptuálně spojit právě se spontánností, jež je demonstrována pozitivním pólem druhé dimenze. Mezi jejími projevy jsou v MDA označeny následující jazykové rysy:<sup>102</sup> kontaktné a výplňkové výrazy, ukazovací zájmena (bez slova *to*), citoslovce, expresivní částice,<sup>103</sup> pronoun non-dropping, ukazovací adverbia, změna *y > ej* v koncovce, opakování slov, úžení *é > í* v koncovce.<sup>104</sup> Z nich jsme pro účely této práce vybrali tyto rysy:<sup>105</sup> kontaktné výrazy, výplňková slova, pronoun non-dropping, opakování slov. První dva představují výčet slov sdružených do skupin podle převažující funkce, jež se odráží v jejím pojmenování<sup>106</sup>. Rys nazvaný pronoun non-dropping reprezentuje užívání zájmena ve funkci podmětu tam, kde pro porozumění plně postačuje pouze slovesný tvar (např. *já sem jednou*

---

<sup>102</sup> Z verze multidimenzionálního modelu označené jako 2017-12-05\_8f vypisujeme max. deset rysů s nejvyšším skóre.

<sup>103</sup> Jedná se o seznam vytvořený kolektivem pracujícím na MDA češtiny. Zařazeny sem byly např. *baže, bohudíky, fakticky, halt, kupodivu, nemlich, prý* a také výrazy, které jsou obsaženy i v seznamu kontaktných či výplňkových výrazů, např. *jo, ne, no*.

<sup>104</sup> Zdůvodnění jednotlivých rysů viz Cvrček et al. (v tisku).

<sup>105</sup> Jejich názvy přejímáme v neměnné podobě. Je-li název anglicky, ponecháváme ho v této podobě a neskloujeme ho.

<sup>106</sup> Více o odlišnostech mezi oběma skupinami viz 3.5.1.

*dělala ve třídě takovou anketu vs. dělala jsem ve třídě takovou anketu*). Čtvrtý rys je označením pro stejná slova stojící bezprostředně vedle sebe.

Také dialogičnost nefiguruje v pojmenování žádné dimenze MDA. Dialogičnosti ve smyslu komunikace v interakci, která implikuje vzájemné střídání komunikantů, se nejvíce blíží pátá dimenze se zaměřením na rozdílnou míru explicitní adresnosti. Kladný pól této dimenze, spojený s vyšší mírou adresnosti, tvoří následující rysy: otázky, resp. otazník,<sup>107</sup> 2. slovesná osoba, doplňovací otázky, zájmena pro 2. osobu, budoucí čas, vykřičník, vokativ, imperativ, přítomný čas, sloveso *být* ve tvaru sufixu (např. *dělals*). Z představených rysů jsme pro účely této práce vybrali následující tři: otázky, zájmena pro 2. slovesnou osobu, vokativ.

Druhým zdrojem pro výběr jazykových jevů vhodných pro analýzu byla odborná literatura. Řada rysů indikovaných MDA jako projev spontánnosti či explicitní adresnosti je v ní uvedena při charakteristice projevů konceptuální mluvenosti. U některých jazykových rysů, zejména rysů obecné povahy (např. slovní druhy, slovesné časy, otázky), však není v literatuře zmíněna žádná (potenciální) odlišnost mezi mluveným a psaným modem, proto jsme se rozhodli začlenit do analýzy i nějaký takový jazykový rys.

Při porovnání mluveného a psaného jazyka se v literatuře nejčastěji poukazuje na využití vizuální sémiotické soustavy pro neverbální komunikaci: zatímco při mluvení tváří v tvář můžeme sledovat např. směr pohledu mluvčího a jeho mimiku, při psaní si musíme vystačit s vizuální podobou textu. V elektronické komunikaci jsou za projev neverbální komunikace považovány emotikony. Ačkoli jich existuje velké množství s odlišnými významy, nejfrekventovanější (např. Park et al., 2014) z nich je emotikon :-) vyjadřující smích, úsměv, radost. Smích je v případě korpusu Ortofon zachycen na dvou anotačních vrstvách podle toho, zda doprovázel verbální projev, nebo ne. Ačkoli Ortofon zachycuje i další neverbální projevy, smích je zde nejčastějším, a proto jsme ho vybrali pro analýzu.

Celkem bylo pro analýzu vybráno 9 jazykových jevů (viz přehled níže), které jsou podle našeho názoru projevem neformálnosti a dialogičnosti, tj. charakteristik, které spojují mluvenou a psanou komunikaci analyzovanou v této práci. Každý jazykový jev je odděleně analyzován v rámci vzorku z příslušného mluveného či psaného webového korpusu. Poté jsou výsledky srovnávány: mezi mluveným a psaným modem a mezi jednotlivými žánry elektronické komunikace navzájem.

---

<sup>107</sup> Při transkripci mluveného jazyka jsou doplňovací i zjišťovací otázky opatřeny otazníkem.

Analyzované jazykové rysy jsme uspořádali do následujících skupin:

- 1) rysy spjaté s dialogičností
  - a. otázky
  - b. zájmena pro 2. osobu
  - c. nominální oslovení
- 2) rysy spjaté se spontánností
  - a. opakování slov
  - b. fragmenty slov, překlipy a přeřknutí
- 3) rysy související s naplněním pragmatických potřeb
  - a. diskurzní markery
  - b. pronoun non-dropping
  - c. iterace grafémů
- 4) neverbální komunikace
  - a. smích/emotikony

### **3.1.2.2 Metodologie analytické části**

Ke srovnávání mluveného jazyka s elektronickou komunikací s ohledem na výzkumné otázky (viz [1.1](#)) byly využity kvalitativní i kvantitativní metody. Vycházeli jsme z korpusově zpracovaných dat, tj. z dat opatřených lemmatizací a morfologickým značkováním dostupných prostřednictvím korpusového manažeru KonText. Prostřednictvím něj byly vytvářeny kvantitativní analýzy (např. frekvence slovních druhů a lemmat, frekvence slov v rámci textu či u jednoho mluvčího), jejichž podrobnosti jsou specifikovány u konkrétní analýzy. Vzhledem k rozdílné celkové velikosti korpusů je frekvence jevu vždy normalizována na milion slov (*instances per million*, dále jen ipm).

Kvalitativní analýzy byly zpracovány také pomocí KonTextu, který umožňuje různé způsoby třídění, filtrování a zobrazení dat. Pro každý žánr, resp. korpus, byl analyzovaný jev zpracováván samostatně podle stejných kroků. Nejprve byl vyhledán v příslušném korpusu, poté byl vytvořen náhodný vzorek o velikosti (zpravidla) 100 konkordančních řádků a ten ručně anotován (celkově se tedy jednalo o 300 konkordančních řádků z elektronické komunikace a 100 z mluveného jazyka). Práce se vzorky je pragmaticky podmíněným rozhodnutím, které je obvyklé při jakékoli analýze zahrnující ruční kvalitativní anotaci.

Jazykové rysy vybrané na základě MDA nejsou vždy identifikovány stejně jako v MDA. V ní byly jednotlivé rysy operacionalizovány nejčastěji jako komplexní dotaz CQL nebo výčet slov

(např. seznam kontaktních výrazů), přičemž cílem operacionalizace bylo najít v korpusu rys s co nejnižší chybovostí. V naší analýze však klademe důraz na kvalitativní analýzu, proto nepoužíváme u jazykových rysů operacionalizovaných v MDA jako výčet všechny položky, ale vybíráme z něj několik lexikálních jednotek, které vyhledáváme (a analyzujeme) samostatně. Teprve při interpretaci výsledků analýzy tyto jednotky sdružujeme. Také CQL dotaz upravujeme podle potřeby. (Více viz v metodologii jednotlivých případových studií.)

Při kvalitativní analýze jsme se inspirovali konverzační analýzou, zejména jejím důrazem na bezprostřední kontext a užívání jazyka v určité komunikační situaci v opozici k jiné. Anotace vzorků probíhala s ohledem na výzkumnou otázku týkající se daného jevu (viz případové studie), vždy však byl brán v úvahu kontext věty, segmentu, příp. jiné vyšší jednotky, v němž se analyzovaný jev nacházel. Jev byl primárně analyzován s ohledem na jednoho mluvčího/pisatele, protože nás také zajímá otázka systematického užívání jevu oproti občasným výskytům; to vše vztaženo k dostupným sociolingvistickým charakteristikám. U mluveného jazyka jsme se zaměřili i na fonetickou rovinu, tj. realizovanou výslovnost, mluvní takty, přítomnost pauz. Při analýze elektronických textů jsme se kvůli nepřítomnosti replik komunikačního partnera nemohli zaměřit ve stejné míře jako u mluveného jazyka na jevy, které jsou důsledkem interakce, nebo na charakteristiky analyzovaného jazykového prostředku, jež se právě v interakci projevují. Toto omezení bereme v potaz při každé komparaci mezi mluveným jazykem a elektronickou komunikací.

Při analýze se zaměřujeme především na otázku srovnatelnosti jazykového jevu v mluveném a psaném jazyce. Zajímá nás, zda se analyzovaný jev projevuje stejně či podobně v obou jazykových módech s ohledem na frekvenci, kontext, pozici v textu atd. Pokud ne, snažíme se najít jeho další realizace či varianty a odpovědět na otázku, zda je příčinou odlišnosti jiný jazykový modus. Zaměřujeme se také na srovnání mezi třemi žánry elektronických textů: diskusními fóry, internetovými diskusemi a veřejnými facebookovými statusy.

### **3.2 Frekvenční srovnání dat**

Cílem následujících srovnání je vytvořit si představu o datech, která budou posléze výběrově analyzována. Při srovnání mezi daty z mluveného korpusu Ortofon a ze psané elektronické komunikace v korpusu Koditex vycházíme ze značkování systémem Morphodita (Straková et al., 2014).<sup>108</sup> Komparace se týká především rozdílů mezi jazykovými módy a mezi třemi žánry

---

<sup>108</sup> Korpus Koditex byl mimoto značován ještě systémem vyvinutým v Ústavu teoretické a počítačové lingvistiky.

elektronické komunikace. Interpretace shod a rozdílů se týká každé z těchto čtyř skupin textů jako celku a je platná pouze pro použitá data, ačkoli jsou závěry formulovány obecně.

### 3.2.1 Frekvence slovních druhů

Největší rozdíly při srovnání slovních druhů<sup>109</sup> napříč daty se ukazují mezi mluveným a psaným jazykem (Čermáková & Kopřivová, 2018). Při počítání procentního zastoupení slovních druhů ze všech tokenů se ukazuje, že v mluveném jazyce<sup>110</sup> a v diskusních fórech dominují verba, zatímco ve zbylých žánrech psané elektronické komunikace substantiva. Rozdíly mezi oběma slovními druhy nejsou u psané elektronické komunikace tak velké jako při srovnání s mluveným jazykem (viz tab. 1), a to ani u diskusních fór, kde verba tvoří 17 % a substantiva 15 %. Největší rozdíly ze všech dat byly u internetových diskusí: verba zde pokrývají 19 %, substantiva 15 %.

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
podstatná jména	9,3 %	19,1 %	14,7 %	18,4 %
přídavná jména	2,8 %	6,6 %	4,7 %	6,4 %
zájmena	16,3 %	12 %	13 %	12,6 %
číslovky	1,9 %	1,6 %	2,1 %	1,4 %
slovesa	16,5 %	15,4 %	17,3 %	16,1 %
příslovce	10,6 %	7,6 %	8,6 %	7,3 %
předložky	4,7 %	7,4 %	7 %	7,3 %
spojky	9,5 %	8,5 %	8,8 %	8,6 %
částice	7,2 %	1,6 %	2,1 %	1,5 %
citoslovce	0,3 %	0,1 %	0,1 %	0,1 %
nerozpoznané	0,6 %	2,1 %	4 %	1,9 %
interpunkce	13,5 %	18 %	17,6 %	18,4 %
fragmenty	1,1 %	-	-	-
hezitace	2,2 %	-	-	-
komentáře	3,5 %	-	-	-

Tab. 1. Frekvence slovních druhů v datech

Rozdíl mezi oběma mody se projevuje i u zájmen, která jsou v mluveném jazyce druhým nejčastějším slovním druhem (16 %) a v psaném třetím. V první pětičce jsou ve všech datech ještě příslovce a spojky, které se mezi mluveným a psaným jazykem liší pouze v pořadí: v mluveném jazyce se vyskytuje více příslovců, v psaném více spojek. To poukazuje na vyšší zastoupení delších a komplikovanějších syntaktických celků v psaném jazyce.

<sup>109</sup> Mezi slovními druhy ponecháváme jako samostatné skupiny interpunkci a nerozpoznané výrazy.

<sup>110</sup> Mezi slovní druhy se v korpusu Ortofon počítají také tři speciální skupiny: hezitace (responzní a hezitací výrazy zapisované jednotnými značkami), fragmenty (tj. nedokončená slova), komentáře (tj. doprovodné zvukové charakteristiky). Tyto skupiny také zahrnujeme do přepočtu tokenů na procenta.

S ohledem na zmiňované slovní druhy (s výjimkou sloves) se mluvenému jazyku nejvíce podobají diskusní fóra. Charakteristické je nižší zastoupení podstatných a přídavných jmen na jedné straně a vyšší zastoupení zájmen a příslovcí na straně druhé.

Pořadí zbývajících slovních druhů se mezi psanými žánry liší minimálně, tj. pouze pořadí číslovek a částic je v datech z Facebooku opačné. Oproti tomu v mluveném jazyce jsou částice výrazně frekventovanější (7 % vs. 1,5–2 % v psaných elektronických žánrech). Minimální rozdíly mezi oběma mody vykazují číslovky a citoslovce, jež zároveň patří k nejméně častým slovním druhům.

Samostatnou skupinu tvoří interpunkce. V psaných datech zaujímá stejný podíl jako nejfrekventovanější slovní druh (17–18 %). V mluveném korpusu tvoří interpunkce 13 %, je jí tedy méně než sloves a zájmen. Příčiny tohoto rozdílu mohou souviset s faktem, že se s ní zachází odlišně. V prepisech mluveného jazyka je užívána pauzová interpunkce, zaměřená na značení různě dlouhých pauz, která nezahrnuje všechna interpunkční znaménka, např. čárky, dvojtečky, vykřičníku, uvozovek, závorek (viz Kopřivová et al., 2014). Veškerá interpunkce v mluveném korpusu je výsledkem interpretace při prepisu mluveného slova. V elektronické komunikaci je užívána větná interpunkce, ale její užívání variuje v závislosti na autorovi. Některé elektronické texty jsou psány téměř bez interpunkce, jiné používají interpunkci ve shodě s pravidly českého pravopisu. Vzhledem ke koncepčním odlišnostem interpunkce v mluvených a psaných elektronických datech se nesnažíme interpretovat rozdíl v jejím podílu.

### 3.2.2 Frekvence lemmat

Pro lepší představu o analyzovaných datech je vhodné srovnat nejčastější lemmata, příp. slovní tvary, které jsou v lemmatu nejprominentněji zastoupeny.<sup>111</sup> V této části stručně okomentujeme 20 nejfrekventovanějších lemmat (bez interpunkce a tří speciálních skupin značených pouze v korpusu Ortofon); některým z nich se budeme věnovat důkladněji.

Největší rozdíly jsou podobně jako u slovních druhů mezi mluveným jazykem a elektronickou komunikací. Nejfrekventovanější lemma v mluveném jazyce představuje zájmeno *ten*, v němž 60,5 % připadá na tvar *to*. V psaných žánrech je toto lemma druhé nejfrekventovanější (a tvar *to* zahrnuje 63,5–65 % lemmatu). Vysoká frekvence lemmatu *ten* a tvaru *to* souvisí jednak s neurčitostí či vágností, která je vlastní spontánním, předem nepřipraveným textům, v nichž není nutné (a v mluveném jazyce ani technicky možné) precizně formulovat, jednak s potřebou

---

<sup>111</sup> Srov. komentář k padesáti nefrekventovanějším slovům v současné mluvené a psané češtině z korpusů Oral a SYN2015 (Čermáková & Kopřivová, 2018).

provázat slova a informace vnitro- i vnětextově. Tvaru *to* je možné přisoudit řadu pragmatických funkcí, které lze pozorovat především v mluveném jazyce, neboť vznikají nezáměrně v průběhu produkce. Jedná se např. o snahu vyplnit potenciální pauzu v řeči při neschopnosti vybavit si konkrétní slovo. Tento důvod může být podle Klimešové et al. (2015) motivací pro užívání reduplikovaných a prefigovaných (např. *vyto*) tvarů zájmena *to*.

Nejfrekventovanějším lemmatem v psaných žánrech je sloveso *být*, které je v mluveném jazyce na druhém místě. Za pozornost stojí zejména kolokace *to je*, která je nejsilnější kolokací tvaru *to* v mluveném i psaném jazyce. Třetím nejfrekventovanějším lemmatem ve všech analyzovaných žánrech je spojka *a*. Mezi nejčastějšími dvaceti lemmaty se ve všech žánrech shodně objevují zájmena *já, co, on, se*, sloveso  *mít*, předložky *v, na*, spojky *ale, tak, že*.

Největší odlišnosti v rámci frekvenční špičky se projevují mezi mluveným a psaným jazykem jako celkem (ne mezi psanými žánry). V mluveném jazyce se vyskytují lemmata *no, jo, tam, hmm*,<sup>112</sup> *vědět, ne*. Z tohoto výčtu je patrná interakční povaha analyzovaného mluveného jazyka – lemmata *no, jo, ne, hmm* a některé tvary lemmatu *vědět* slouží jak k zodpovězení otázky, tak k vyjádření pragmatických funkcí souvisejících např. se změnou tematiky, předáním slova, členění repliky (Čermáková et al., 2019). Lemma *tam* je užíváno podobně jako *to* vnitro- i vnětextově k vzájemnému usouvztažnění informací již nebo následně zmiňovaných či těch, které jsou známy anebo přítomny mimojazykově. V psaných žánrech se oproti mluvenému jazyku objevují lemmata sloužící ke specifikaci vztahů mezi slovy a větami (předložky *do, o, s, z, za*, spojka *i*, vztažné zájmeno *který*).

Srovnání s psaným korpusem SYN2015 (Křen et al., 2015) ukazuje, že 11 z těchto lemmat (*a, ale, být, já, mít, na, on, ten, to, v, že*) patří k nejčastějším 20 v psaném jazyce obecně. Další čtyři lemmata (*i, o, s, z*) se pak objevují v nejčastějších 20 v psaném jazyce z korpusu SYN2015 a Koditex-web.

### 3.3 Rysy spjaté s dialogičností

Dialogičnost lze podle Müllerové (1989) nazírat ze dvou úhlů pohledu, a to jednak z hlediska výměny komunikačních rolí (mluvčí/pisatel vs. adresát), jednak jako míru explicitní adresnosti. Pro takto definovanou dialogičnost bychom mohli používat i pojem interaktivnost, který upřednostňuje procesuální pohled na komunikaci. Brazil (1995) přitom interaktivnost označuje za jednu z nejtypičtějších vlastností mluvené komunikace. Pro účely porovnání mluveného

---

<sup>112</sup> Jedná se o unifikovaný zápis všech kladných responzních reakcí, jejichž slovní zápis nemá konvencionalizovanou podobu.



a psaného elektronického jazyka jsme vybrali tři jazykové jevy (otázky, zájmena pro 2. osobu, nominální oslovení), jejichž prostřednictvím se mluvčí/pisatel aktivně obrací na adresáta.

### 3.3.1 Otázky

Otázky jsou podle Meyera (2017) takovým typem vyjádření, kterým dává mluvčí najevo absenci nebo nedostatek informace a signalizuje adresátovi svůj vlastní záměr či pobídku tuto informační mezeru zaplnit. Pojem otázka bývá v české lingvistice často používán synonymně s pojmem tázací věta (více viz Meyer, 2017), která je v psaném jazyce odlišena od jiných typů vět otazníkem. Užívání otazníku v českých psaných textech podléhá Pravidlům českého pravopisu (1999: 64), v nichž se uvádí, že otazník se píše „na konci tázacích výpovědí, ať mají formu věty, souvětí nebo samostatného výrazu nevětného“. V této studii pojmáme otázku pouze z hlediska formy jako jakoukoli syntaktickou strukturu zakončenou otazníkem. Stranou zůstávají nepřímé otázky, které také pobízejí k zaplnění informační mezery. Speciálním problémem jsou pak tzv. řečnické otázky.

Typologie otázek je různorodá, tradičně se v češtině rozlišují otázky zjišťovací a doplňovací, (např. Karlík et al., 2003; Štícha, 2013), k nimž se dále pojí např. otázky řečnické, rozvažovací, podivové. Zjišťovací otázky se poznají podle významu platnosti, nebo neplatnosti vyjádřené propozice (na tyto otázky se dá odpovědět ano/ne), doplňovací otázky podle přítomnosti tázacího výrazu, např. zájmenného příslovce *kdo*. V mluveném jazyce je pro naznačení otázky používána intonace podle toho, o jaký typ otázky se jedná. Otázky doplňovací jsou realizovány klesavým (ukončujícím) melodémem, otázky zjišťovací pak melodémem (ukončujícím) stoupavým (Palková, 2017).<sup>113</sup>

V této studii se otázkám věnujeme z pohledu dialogičnosti. Podle jejich přítomnosti a frekvence v textech budeme porovnávat míru adresnosti analyzovaných textů. Zajímat nás bude také to, zda je otazník dobrým signálem pro vyhledání otázek, příp. pro jaké další funkce je využit.

### Metodologie

Pro vyhledání otázek jsme po vzoru MDA (Cvrček et al., v tisku) využili jejich grafické značení pomocí otazníku. Vyhledání<sup>114</sup> bylo jednotné pro psaný i mluvený jazyk, protože transkripty

<sup>113</sup> Vzhledem ke stejné gramatické stavbě oznamovací věty a zjišťovací otázky je prostředkem pro jejich odlišení použití ukončujícího stoupavého melodému. Zvuková stavba kadencí, tzv. antikadence, je v tomto melodému v porovnání s ostatními nejstabilnější a nejvyhraněnější (Daneš, 1949; Palková, 2017).

<sup>114</sup> Použit byl následující dotaz:

```
[word!="[\?!\.]*\?[\?!\.]*"] [word="[\?!\.]*\?[\?!\.]*"]+[word!="[\?!\.]*\?[\?!\.]*"]
```

nahrávek v mluveném korpusu Ortofon užívají otazník v souladu s psanými texty, tj. pro značení otázky.

Samostatně byly vyhledány doplňovací otázky.<sup>115</sup> Tabulka 2 ukazuje, že tento typ tvoří v žánrech psané elektronické komunikace přibližně třetinu všech otázek, zatímco v mluveném jazyce cca pětinu. Tento rozdíl se pokusíme ozřejmit v následující analýze založené na seznamu slov, po kterých se v mluvených i psaných datech objevuje otazník. U kvalitativních analýz zaměřených na funkce slov v kombinaci s otazníkem pracujeme max. se 100 výskyty příslušného výrazu (v případě, že je absolutní frekvence vyšší než 100, pracujeme s náhodným vzorkem).

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
všechny otázky	14 839 (12 000,73 ipm)	1 494 (6 202,93 ipm)	1 247 (5 141,7 ipm)	1 568 (6 458,26 ipm)
doplňovací otázky	3 230 (2 612,19 ipm)	437 (1 814,38 ipm)	344 (1 418,4 ipm)	478 (1 968,78 ipm)

Tab. 2. Frekvenční přehled otázek v analyzovaných datech

## Analýza

V mluvených textech se oproti psané elektronické komunikaci vyskytuje otazník výrazně častěji; podíl doplňovacích otázek je v nich přitom menší. Vzhledem ke zkušenostem s transkripcí pro mluvený korpus Ortofon a dalším analyzovaným jevům víme, že otazník se užívá také u diskurzních markerů vyslovených se stoupavou intonací (pro různé funkce, např. kontaktovou, pro vyjádření překvapení). Porovnání seznamu nejčastějších dvaceti slov následovaných otazníkem ukazuje odlišnosti mezi mluveným a psaným elektronickým jazykem především ve druhé desítce: v psaných textech se objevují plnovýznamová slova, v mluvených nikoli (tab. 3). V mluveném jazyce se ve spojení s otazníkem častěji vyskytují tázací slova, z nichž je pro všechny žánry společný výskyt slov *co* a *proč*. Kromě nich se jedná o slova *ne*, *to*, *že*. Společně v mluveném jazyce a aspoň v jednom psaném žánru se dále vyskytují tři další slova: *bylo*, *je*, *jo*.

<sup>115</sup> Pomocí následujícího dotazu:

```
[lemma="(kdo|co|jaký|který|čí|kde|kdy|odkdy|dokdy|kam|kudy|odkud|proč|jak|kolik(rát)|nač)(s|pak)?" ] [pos!="[ZJ]" ]+[pos="Z"&word!="\?" ]*[word="\?" ] within <s/>.
```

Pro vyhledávání v mluveném korpusu byla struktura <s/> nahrazena strukturou <sp/>.

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
1	jo	ne	ne	ne
2	ne	že	co	že
3	co	co	že	co
4	že	proč	jo	proč
5	víš	jo	proč	je
6	je	EU	je	to
7	to	to	vid'	což
8	vid'	stát	to	EU
9	fakt	vážně	myslíš	lidí
10	cože	nevadí	není	tím
11	jako	lepší	tak	myslíte
12	kde	pravda	života	názor
13	proč	bude	ty	nic
14	nebo	bylo	dítě	politiky
15	teda	platí	jde	Sýrii
16	no	lidí	pořádku	NATO
17	NP	společnosti	dobře	sebou
18	tam	ulici	děti	asi
19	kdo	děti	dovolenou	dělat
20	bylo	jde	problém	vy

Tab. 3. Seznam dvaceti nejčastějších slov následovaných otazníkem

Vzhledem k zaměření této části práce na dialogičnost jsme se podrobněji podívali na funkce slov *co*, *jo*,<sup>116</sup> *ne*, *že* v kombinaci s otazníkem. Všechna z nich mohou sloužit jako tzv. tázací dovětky, kterými se mluvčí/pisatel obrací na adresáta, vyzývá ho jimi k reakci, proto bychom je označili za diskurzní markery (viz tab. 4 a př. 4).<sup>117</sup>

(4)

- To přeci není nebezpečí , **ne** ? (diskuse)
- O všechno je třeba se řádně starat , **že** ? (fóra)
- Těžká závist vědět , že budeš muset makat jak socka **co** ? (Facebook)
- nech\* nespěchejte na mě **jo** ? musíte pomalu .. (Ortofon)

<sup>116</sup> U tohoto výrazu sice už byly jeho funkce analyzovány (viz 3.5.1.1), nejednalo se však o vyhledání kombinace s otazníkem.

<sup>117</sup> Tato funkce již byla nalezena u diskurzních markerů *jo*, *fakt*, *víš* a *že jo* (viz 3.5.1), proto ji zde detailněji nerozebíráme. Kromě ní jsme u slov *jo* a *ne* identifikovali tyto dvě funkce: nahrazení kladné/záporné věty, vyjádření překvapení (především u slova *jo* tvořícího samostatný segment).

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
co?	18 %	36 %	40 %	35 %
jo?	63 %	88 %	87 %	-
ne?	85 %	61 %	84 %	80 %
že?	100 %	100 %	100 %	100 %

Tab. 4. Zastoupení funkce výzvy adresáta k reakci

Zatímco u spojení *jo?*, *ne?*, *že?* tato diskurzivněmarkerová funkce převažuje v mluvených i psaných textech, u spojení *co?* se v psaných textech vyskytuje max. ve 40 % případů a v mluveném jazyce pouze v 16 %. Tento rozdíl mezi slovy souvisí s jejich lexikálním významem: výraz *co* je sám o sobě otázkou a v tomto významu je ve spojení s otazníkem používán především,<sup>118</sup> zatímco ostatní slova primárně jako tázací výrazy nefungují. Navíc je u výrazu *co* možno rozlišit různé typy zjišťovacích otázek: zjišťování informací v pravém slova smyslu (např. *a to je co ?*), ve smyslu žádosti o zopakování předchozí repliky/informace (zejména v mluveném jazyce, viz př. 5), ve smyslu otázky na doplnění konkrétního slova (např. *na to se ti víš co ? ? ?*, *Ale oknem se co ? Oknem se neleze . . . !*).

(5)

Štěpán\_175: a proč nejseš zapnutá ?

Žaneta\_176: **co ?**

Štěpán\_175: proč nejseš zapnutá

Rozdíl v počtu doplňovacích otázek v mluveném a psaném jazyce (tab. 2) podle našeho názoru souvisí s počtem funkcí uplatňovaných tázacími výrazy. Jako příklad vezměme výraz *co*, který je jak v mluveném, tak psaném elektronickém jazyce nejčastějším tázacím slovem. Výše jsme naznačili jeho tři funkce: zjištění informace, výzva k reakci, žádost o zopakování. Právě posledně jmenovaná funkce se objevila pouze v mluveném jazyce (v 40 % vzorku), a to buď jako prostá otázka *co?* (př. 5), nebo doplněná o zopakování části repliky, na niž se mluvčí táže (př. 6). V psaném jazyce takové opakování vlastně není potřeba, protože k jakémukoli místu textu se lze kdykoli vrátit;<sup>119</sup> může být však užito pro naznačení nepochopení, nesouhlasu nebo ironie. V analyzovaných psaných datech se *co?* v tomto smyslu neobjevilo, ale dá se o něm uvažovat např. v příkladu z internetových diskusí *Co to plácáte ? Komu co bral ?*

<sup>118</sup> V SSJČ je *co* ve významu tázacího zájmena uvedeno na prvním místě, za ním následuje funkce vztažného zájmena. Slovník označuje funkci pojmenovanou v této práci jako výzvu k reakci za citoslovečnou a v exemplifikaci je slovo *co* uvedeno vždy na konci věty s otazníkem.

<sup>119</sup> Abstrahujeme od situací, ve kterých nelze potřebný text zobrazit.

(6)<sup>120</sup>

Oldřich\_933: za telefon jsem tam 2 nechal

Adam\_929: za telefon **co** ?

Oldřich\_933: za telefon tam strašně moc nechal .

U otázek nás dále zajímalo, jak často se v nich vyskytuje sloveso a v jaké osobě. Vzhledem k tomu, že vyhledávání otázek pouze pomocí otazníku vnáší do výsledků i neotázkové využití např. pro diskurzní markery, brali jsme v úvahu pouze doplňovací otázky. Sloveso bylo v mluvených a psaných doplňovacích otázkách užíváno v podobné míře: v 83 % vs. v 82–87 %. V užívání 2. slovesné osoby byla k mluvenému jazyku (31 %) nejbližší diskusní fóra (29 %); ve zbylých elektronických žánrech byla tato osoba využita méně často: v internetových diskusích v 20 %, na Facebooku v 16 %. Zastoupení plurálových forem bylo častější v psaném jazyce (viz tab. 5 na s. 63 či tab. 25 na s. 110).

Doplňovací otázky bez slovesa byly obecně kratší než otázky se slovesem. Jednalo se struktury vztahující se (zejména) k předchozímu kontextu, který chce mluvčí/pisatel dodatečně rozvinout, doplnit apod. (př. 7, 8). V psaném elektronické komunikaci byly tyto otázky nejčastěji uvozeny konektivními výrazy *a*, *tak(že)*, v mluvené komunikaci ještě výrazem *no* a spojením *no a*. Doplňovací otázky uvozené výrazem *co* byly v internetových diskusích a diskusních fórech častěji než ve zbylých žánrech využity pro získávání názoru adresáta.

(7)

- Zjevně nemůžeme zavřít 2 miliony lidí do kriminálu , **takže co s tím ?** (diskuse)

- přítel se bojí já jsem minulý rok měla zamlklý potrat . . a on nechce nějak nic riskovat no . . i když u mě by chuť byla už se vrátila . . **a co vy ?** (fóra)

- Nikdo se ani nepokusil tam jít žádat . Nikdo z nich . **proč asi ? ? ?** (Facebook)

(8)

Karel\_98: kdo ví jestli ty babky to to jestli to tam ještě žije

Jiřina\_99: **který babky ?**

Karel\_98: no ta máma vod vod @@

## Shrnutí

V této studii jsme porovnávali výskyt otázek v mluveném a psaném elektronickém jazyce. Pojetí otázky je zde determinováno možnostmi korpusového vyhledávání: otázky jsme v datech

---

<sup>120</sup> Číslicí je v transkriptech korpusu Ortofon označen počet nesrozumitelných slov.

identifikovali prostřednictvím otazníku. V analýze se však ukázalo jeho širší využití, nejedná se tedy o prostředek, který by byl jednoznačným signálem pro identifikaci otázek.

Otazník se v obou typech dat vyskytoval i u slov (a delší slovní spojení, ale ta analyzována nebyla), která sice nejsou tázacími výrazy, ale fungují (v mluveném jazyce s náležitou intonací) jako výzva k reakci. Mluvčí/pisatel jimi zároveň posiluje kontakt s adresátem či udržuje jeho pozornost. Jednalo se přitom o slova, která byla v mluveném i psaném jazyce mezi nejčastějšími deseti slovy následovanými otazníkem. Poměr této funkce oproti jiným souvisí s lexikálním významem daného slova: u spojky *že* sloužila všechna spojení *že?* jako výzva, zatímco u slov *jo* a *ne*, která mohou zastupovat kladnou či zápornou větu, na jejíž obsah se mluvčí táže, byla funkce výzvy zastoupena méně (cca v 80 % výskytů). Nejzajímavější bylo v tomto kontextu slovo *co*, které je samo o sobě tázacím výrazem (v doplňovacích otázkách dokonce nejčastějším). Jeho využití pro vyzvání adresáta k reakci se v psaných textech vyskytovalo v 35–40 % výskytů, v mluveném jazyce v 18 % výskytů. Otazníkem se tedy značí spíše stoupavá intonace, s níž jsou obecně otázky asociovány, ale která je realizována ve zjišťovacích otázkách (Palková, 2017). U doplňovacích otázek zakončených otazníkem se přibližně v 80 % případů vyskytuje sloveso. Pokud se v nich nevyskytuje sloveso, jedná se o kratší otázky, které jsou silně kontextově zapojeny.

Mnohonásobně vyšší frekvence otazníku v mluvených datech může poukazovat na vyšší dialogičnost tohoto typu dat, v němž je častěji než v psaném textu použita otázka za účelem zopakování přeslechnutého či nerozpoznaného slova. Lexémy, které jsou uplatněny pro tuto funkci (např. *fakt*, *co(že)*), se nevyskytují v psaných elektronických textech mezi nejčastějšími, jako je tomu v mluvených datech. Výrazný rozdíl v zastoupení otazníku může také souviset s editací dat: přepisy mluvených dat podléhají pravidlům, jejichž dodržování kontroloval omezený počet lidí, zatímco psané elektronické texty nebyly nijak sjednocovány. To by se mohlo týkat především diskurzivních markerů zajišťující vyzvání adresáta k reakci (viz 3.5.1). Počet doplňovacích otázek byl sice v mluvených datech také vyšší než v psaných, ale jednalo se o menší rozdíl než u zastoupení otazníku samého. Z žánrů elektronické komunikace bylo nejvíce doplňovacích otázek nalezeno na Facebooku, dále v internetových diskusích a nejméně v diskusních fórech.

### 3.3.2 Zájmena pro 2. osobu

Zájmena vyjadřující 2. slovesnou osobu představují v komunikaci orientaci na adresáta.<sup>121</sup> Rozdíl mezi singulárem a plurálem, který primárně vyjadřuje rozdíl mezi jedincem a skupinou,

---

<sup>121</sup> 2. slovesná osoba je v porovnání s 1. a 3. slovesnou osobou nejméně častá (viz tab. 25 na s. 110).

se u 2. osoby proměňuje v důsledku vykání, které je v češtině projevem úcty k adresátovi, ale i naznačením vzájemného vztahu ne-blízkosti, formálnosti. Při vykání jedné osobě se v plurál mění pouze tvary finitního slovesa, u složených slovesných tvarů se to týká pomocného slovesa, zbylé tvary zůstávají v singuláru. Využití zájmen pro 2. os. sg. se týká studie rysu pronoun non-dropping (3.5.2). Na rozdíl od ní se v této studii zaměřujeme na výskyt zájmen 2. osoby ve všech tvarech včetně zájmen přivlastňovacích. Zajímat nás budou potenciální rozdíly v jejich užívání v mluveném a psaném jazyce.

## Analýza

Při vyhledání jsme brali v úvahu zájmena *ty*, *vy*, *tvůj*, *váš* ve všech tvarech včetně tvarů nářečních, např. *ťá*, a tvarů bez diakritiky, např. *vas*. Ne všechny tvary byly při lemmatizaci správně označeny, proto jsme při vyhledávání použili výpis všech možných tvarů, jednoznačně patřících pod daná lemmata.<sup>122</sup> Jejich frekvenční přehled ukazuje následující tabulka:

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
ty	3 003 (2 428,61 ipm)	282 (1 424,61 ipm)	742 (3 708,07 ipm)	247 (1 238,60 ipm)
vy	1 005 (812,77 ipm)	584 (2 975,53 ipm)	184 (919,52 ipm)	633 (3 174,23 ipm)
tvůj	205 (165,78 ipm)	68 (343,52 ipm)	150 (749,61 ipm)	32 (160,46 ipm)
váš	121 (97,86 ipm)	176 (889,12 ipm)	45 (224,88 ipm)	159 (797,32 ipm)
celkem	4 334 (3 505,03 ipm)	1 115 (4 629,36 ipm)	1 121 (4 622,17 ipm)	1 071 (4 411,22 ipm)

Tab. 5. Frekvence zájmen pro 2. osobu v analyzovaných datech

Tabulka 5 ukazuje rozdíl mezi mluveným a psaným jazykem v používání zájmen pro singulár a plurál. Zájmena pro singulár jsou častější v mluveném jazyce a na fórech. Přestože jsou v tomto ohledu fóra podobnější mluvenému jazyku, celková frekvence plurálových zájmen je ve fórech vyšší než v mluveném jazyce. Výrazně častější užívání zájmen *vy* a jeho přivlastňovací podoby *váš* v psané elektronické komunikaci oproti mluvenému jazyku není překvapivé, uvážíme-li, že v analyzovaných žánrech elektronické komunikace jsou zahrnuty pouze veřejné příspěvky, ve kterých se může pisatel vztahovat k velikostně neurčené skupině adresátů, tedy ke všem potenciálním čtenářům, a ve kterých se komunikanti často neznají, takže

<sup>122</sup> Použit byl následující dotaz:

```
[word="(?)t[ě]|ťa|tebe|teb[óu]|tebou|tob[ěe]|tys|v[aá]m[ai]|v[aá]m|v[aá]s|vy|tv[áé]|tv[éý]ho|tvejch|tvejma|tvejm|tv[éý]m|tv[éý]mu|tv[íý]|tvým[ai]|tvoj[aeíí]|tvoj[íí]ch|tvoj[íí]m|tvou|tv[uú]j|tvých|tv[yý]|vašeho|vašem|vašemu|vaš[aeu]|vaš[íí]ch|vašim[ai]|vaš[íí]|vaš[íí]m|váš"] | [word="(?)t[ííyý]" & lemma="(?)ty"].
```

si vykají. Příklad 9 ukazuje, použití zájmena *vy* v jednom úseku z diskusního příspěvku pro vykání a zároveň pro označení skupiny, do níž je adresát zařazen. Z posledního výskytu zájmena není jasné, zda jej pisatel vztahuje k jednomu adresátu, nebo ke skupině.

(9)

**Vy** jste pomatený sluníčkář , který vůbec neví , co píše naštěstí je **vás** opravdu málo , odhadoval bych to tak na 4 % , a **vy** budete první , až se sem všichni ti co se sem valí obdařeni touto kulturou  
(diskuse)

Dále jsme se zaměřili na rozdíly v používání analyzovaných zájmen. Vycházeli jsme přitom z kolokační analýzy<sup>123</sup> kontextu příslušného zájmena v rámci každého žánru. Největší rozdíly mezi žánry byly u zájmena *ty*. V mluveném jazyce u něj byly identifikovány kolokace, které se neobjevily v psaných elektronických žánrech ani u ostatních zájmen, např. *ty jo, ty vole, počkej + ty* (např. *počkej a ty chceš, počkej já ti ukážu*). Kolokace *prosím tě/vás* a *říkat/říct + ty* (např. *jak jsem ti říkal, já ti řeknu*) se vyskytovala také jen v mluveném jazyce. V příspěvcích z diskusních fór a Facebooku se zájmena *ty* a *vy* objevila jako součást přání, např. *přeji ti/vám, moc ti držím palce*, a vyznání sympatií, např. *mám tě ráda*. Zájmeno *vy* kolokuje častěji než *ty* s předložkami, např. *u vás, s vámi*, v psaném jazyce se navíc objevovalo *podle vás*. I v předložkových kolokacích se projevuje rozdíl mezi jazykovými mody: v mluveném jazyce byla kolokace *u vás* používána jako místní určení, v psaném jazyce ve smyslu „podle vašeho názoru“. Ve všech žánrech se zájmena *ty* a *vy* objevila v otázkách; z nejčastějších tázacích výrazů se zde vyskytovaly *co, jak, proč*. Samozřejmě se jako vůbec nejčastější kolokace objevila slovesa *být* v různých časech a způsobech a *mít* v přítomnosti. V internetových diskusích stojí za zmínku srovnávací konstrukce *jako jsi ty* doprovázená emotikonem :-).

Přivlastňovací zájmena *tvůj, váš* tvořila ve všech žánrech elektronické komunikace kolokace se slovy *věc, názor, problém*, v diskusích dále *řešení* a na Facebooku *jednání, volba*. Takové kolokace se vztahují k výměně názorů (nejen mezi komunikanty navzájem, ale i mezi diskutujícími a realitou). V mluveném jazyce kolokují přivlastňovací zájmena výrazně sama se sebou v důsledku opakování slov nebo přivlastnění různých substantiv, např. *to tvoje @ tvoje drobky od tvého stolu*. Zájmeno *váš* se v mluvených datech dále vyskytovalo v kolokaci *ten váš*, která slouží jak k vytčení rozdílů mezi mým/naším a tvým/vaším (např. o obraze: *já jsem tam na ten váš taky koukala*), tak jako odkaz ke skutečnosti, která je komunikantům známá, ačkoli se v dialogu tematizuje poprvé (např. *jak je starej vlastně ten váš spolek*).

---

<sup>123</sup> Použili jsme automatické hledání kolokací v rozsahu -3 +3.



## Shrnutí

Při analýze zájmen pro 2. osobu jsme shledali frekvenční rozdíl v užívání singulárových forem, častějších v mluveném jazyce, a plurálových, užívaných výrazně více v psané elektronické komunikaci. Tento rozdíl souvisí s mírou známosti komunikantů a počtem adresátů: v mluvených datech jsou zachyceny dialogy probíhající tváří v tvář mezi tykajícími si komunikanty nejčastěji v počtu dvou osob, zatímco v psaných datech se jedná o blíže neurčený počet komunikujících osob, které si vykají. Ze psaných žánrů se mluvenému jazyku nejvíce s ohledem na frekvenci singulárových zájmen blížila diskusní fóra. Tato blízkost se však neukázala při analýze kolokací příslušných zájmen. Pomineme-li slovesa *být* a *mít*, vyskytující se v psaných i mluvených datech jako nejsilnější kolokáty, objevovaly se v mluveném jazyce kolokáty jiné než v psaných žánrech. V psaných žánrech se díky kolokační analýze ukázal jejich diskusní charakter.

### 3.3.3 Nominální oslovení

Oslovení představuje zvláštní způsob jazykového chování, při kterém se mluvčí/pisatel explicitně vztahuje k osobě adresáta. Podle Bergra (2002) se jedná o zvláštní jazykový fenomén indikující sociální vztah mluvčího a adresáta, který nepatří do zdvořilosti, protože ta je ovlivňována pragmatickými aspekty komunikační situace. Nekula (2010) spojuje oslovení se začátkem a koncem komunikační události nebo sekvence, ve kterých se objevují pozdravy a další typizované prostředky představování. Dále upozorňuje, že způsob oslovení není v rámci jedné komunikační události mezi týmiž mluvčími konstantní, ale mění se s ohledem na téma nebo komunikační cíl mluvčího. Při rozlišování funkcí oslovení (vokativu) v mluveném jazyce dospěl Clancy (2016: 115) k závěru, že nejčastěji je oslovení využito ve funkcích vztahujících se k tématu hovoru, např. nastolení, změna, uzavření tématu, a v interpersonálních funkcích, např. vytvoření a udržení sociálních vazeb mezi komunikanty.

V češtině je základním jazykovým prostředkem nominálního oslovení vokativ, např. *Jano, pane Nováku*.<sup>124</sup> Pozice vokativu je v českém pádovém systému specifická: na jedné straně stojí jeho zvláštní sémantika<sup>125</sup> a s ní související vyčlenění ze syntaktické struktury věty, na druhé straně existují pro tvoření vokativu speciální vokalické sufixy, které se uplatňují pouze u substantiv (Štícha, 2013: 306). Mimoto se v češtině za vokativní považují redukované formy jmen končící na *-i* nebo *-u*, např. *mami, Veru*. Vokativ je podle MSČ (Cvrček et al., 2015: 174)

<sup>124</sup> Při kombinaci titulu a příjmení se zejména v neformálních mluvených situacích užívá příjmení v nominativu, např. *pane Novák*.

<sup>125</sup> Např. Šmilauer (1966: 107) nepovažuje vokativ za pád stejného typu jako ostatní. Vokativní formy slov jsou pro něj „citoslovce tvořená ze substantiv“.

nejméně frekventovaným pádem v psaných textech (nejčastěji se objevuje v beletrii). Při srovnání vokativu singuláru u proprií se ukazuje výrazný nárůst v mluvené neformální komunikaci (15 %) oproti beletrii (5 %).

V této studii se budeme zabývat přímým nominálním oslovením, tedy oslovením pomocí substantiva ve vokativu včetně tzv. redukované formy a jeho kombinace s pronominálním oslovením.<sup>126</sup> V psané elektronické komunikaci se za oslovení dá považovat i specifický způsob adresování zprávy pomocí zavináče a uživatelského jména příslušného adresáta v nominativu, který se v analyzovaných datech objevuje v diskusních fórech (tučné označení v př. 10). Tento způsob budeme při kvantitativním srovnání výskytu oslovení uvádět samostatně vzhledem k tomu, že se jedná spíše o pozůstatky po čištění dat při korpusovém zpracování (viz 3.1.1). Kromě frekvenčního porovnání nás bude zajímat funkce oslovení vzhledem k jeho umístění v textu.

(10)

@ Neff píše : @ mireceklbecek dobre , ale par let je pro tebe co ?

(fóra)

## Metodologie

Pro vyhledávání tohoto jevu bylo nutné vzít ohled na skutečnost, že při automatickém morfologickém značkování nejsou formy redukovaného vokativu označeny jako vokativ,<sup>127</sup> ale jako nominativ. Tyto redukované formy jsme proto vyhledávali odděleně stejně jako případy oslovení obsahující zavináč. Oba tyto frekvenční údaje uvádíme v tabulce 6 zvlášť, protože se jedná o zkrácený údaj už z principu: redukované formy jsme vyhledávali na základě seznamu<sup>128</sup> vytvořeného manuálně (nečiní si proto žádné nároky na úplnost), tyto formy však mohou být použity v jakémkoli jiném pádě, tedy jako nesklonný výraz (př. 11), číslo v tabulce ukazuje pouze počet výskytů považovaných za vokativní.<sup>129</sup> U oslovení se zavináčem byla situace odlišná: po vyhledání všech zavináčů jsme ručně vytrídili případy, ve kterých se nejednalo o pisatele daného příspěvku, tj. jako oslovení jsme chápali všechny případy, ve

---

<sup>126</sup> Vynecháváme případy substantivizovaného adjektiva kombinovaného s přivlastňovacím zájmenem kvůli nízké frekvenci. Vyhledání tohoto jevu s lemmatem *můj* ukázalo v korpusu Ortofon pouze jeden výskyt (*můj milej zlatej boubelatej*), v korpusu Koditex-web žádný.

<sup>127</sup> Pro vyhledání vokativu byl použit následující dotaz: [tag="N...5.\*"].

<sup>128</sup> Nejprve jsme v korpusu Ortofon vyhledali substantiva končící v nominativu nebo vokativu singuláru na *-i* a *-u*. Ze seznamu o 67 slovních tvarech jsme poté ručně vytrídili ty, které v této formě mohou plnit roli oslovení. Tento seznam byl poté aplikován na data z psané elektronické komunikace.

<sup>129</sup> Z celkového počtu nalezených redukovaných forem bylo v mluveném jazyce odstraněno 42 případů a v psaném elektronickém jazyce 14 případů.

kterých po přezdínce (uvozené zavináčem) není uvedeno sloveso indikující autorství (v př. 10 sloveso *píše*).

(11)

- a to naše **mami** nepije .. (Ortofon)
- Na nákup jedeme zítra , ráno ještě k **babi** do bph (fóra)
- ale ona beze mě nebyla tak nějak nikdy . . max u **babi** která bydlí pod náma v bytovce (Facebook)

	Ortofon	Koditex		
		diskuse	fóra	Facebook
značkový vokativ	2 615 (2 114,83 ipm)	205 (851,14 ipm)	177 (729,82 ipm)	192 (790,81 ipm)
redukováná forma	167 (136,57 ipm)	0	40 (164,93 ipm)	3 (12,36 ipm)
oslovení se @	-	0	647 (2 667,74 ipm)	0

Tab. 6. Frekvence analyzovaných forem nominálního oslovení

Tabulka 6 ukazuje výrazně vyšší využití substantiv ve vokativu než v psané elektronické komunikaci. Naopak frekvence redukovaných forem oslovení je v mluveném jazyce mírně nižší než v diskusních fórech. Číselné údaje zahrnují apelativa a propria.

### Analýza

Obě dvě prvně zmíněné formy oslovení se v mluveném i psaném jazyce vyskytují nejčastěji uprostřed segmentu/věty. V mluveném jazyce se oslovení objevují častěji na konci než na začátku segmentu, vyloučíme-li však z toho výraz *vole*, obě pozice se vyrovnají. V psané elektronické komunikaci je druhou nejčastější pozicí oslovení začátek věty, tato pozice je zároveň jedinou, na níž bylo identifikováno oslovení začínající symbolem @.

Rozdíl mezi mluveným a psaným jazykem ve vyhledaných vokativech je v jiném zastoupení oslovovaných entit. V mluveném jazyce převažují oslovení rodinných příslušníků (např. *dcero, dědo, mamko, teto, tatínku*), milostná oslovení (např. *broučku, kočička, kocoure, láska, miláčku, myšáčku, prdelko*) a oslovení mezi kamarády (např. *hochu, chlape, kámo*); tyto tři skupiny jsou v psané elektronické komunikaci zastoupeny v menším rozsahu, tedy i s nižší frekvencí (např. u milostných oslovení jen *kočko, kočička, láska, miláčku*). Vokativ od slova *člověk* (v datech zachycený jako *člověče, čoveče, čéče*), který funguje nejen jako oslovení konkrétního adresáta,

ale i jako diskurzní marker<sup>130</sup> (srov. př. 5), se také vyskytoval v obou typech dat s vyšší frekvencí v mluveném jazyce.

V mluveném i psaném jazyce byly dále nalezeny dehonestující výrazy, které můžeme označit za nadávky: většinou se jednalo o výrazy degradující inteligenci (např. *blbe, blbečku, blbče, blboune, debile, huso, náno, troubo, tupče*). Rozsah této skupiny není možné jednoznačně určit, protože kromě jednoznačně dehonestujících výrazů se silným negativním nábojem (např. *hajzle, sračko, zmrde*) se v textech objevují i výrazy méně negativní (např. *chudáčku, káčo, kecalisto*) a výrazy, jejichž negativitu je možné vyčíst pouze z kontextu<sup>131</sup> (např. *To tomu ty holčičko Aličko ještě nerozumíš; Já jsem pouze citoval milá dámo; dej mně pár milionů doláčů , a já ti statisticky nevyvratitelně dokážu , že je to jenom koincidence , ty statistiko; jo to určitě , jestli je někde pohledávka nedostupná tak v rusku , ty platební morálko - když ti nezaplátí předem tak neuvidíš nic*). Blízko ke skupině dehonestujících výrazů mají slova, jež jsou formálně vokativem, ale mohou být užita kromě oslovení a nadávky i jako vyjádření překvapení. V analyzovaných datech se jedná o výrazy *vole*,<sup>132</sup> *pičo* a *(ty) krávo*, u posledně zmíněného výrazu se liší použití v mluveném a psaném jazyce (př. 12, 13). Naopak vokativy *(ty) krásu*, *(ty) vado* nejsou nikdy oslovením, zato vždy vyjádřením údivu.<sup>133</sup>

(12)

Novotná , **ty krávo** , přestaň to tady spamovat , nebo tady ty diskuse také zruší , stejně jako si to dokázala na Aktuálně !!!! (diskuse)

(13)

Patrik\_967: ty jo byl jsem ve Steak house . v Americe +

Lubomír\_959: no

Patrik\_967: + **ty krávo** .. hochu to jsem nežral nikdy ..

V mluveném a psaném jazyce se vyskytovala společenská oslovení *pane, paní, slečno*, u nichž v psané elektronické komunikaci většinou následovalo příjmení (viz dále).<sup>134</sup> Oslovení titulem nebo funkcí bylo výrazně častější v psaném jazyce (např. *autore, doktore, mistře, novináři*,

<sup>130</sup> V článku Klimešové et al. (2017) je tato funkce pouze naznačena (pod pojmem *lexical filler*), protože se nejedná o detailní textovou analýzu.

<sup>131</sup> Dalo by se zde uvažovat o ironii.

<sup>132</sup> Viz přehled funkcí tohoto výrazu v Čermákové et al. (2019: 275n.).

<sup>133</sup> U nich se ukazuje nejednotnost v morfologickém značkování. Domníváme se, že tyto dva výrazy jsou na stejné úrovni jako *(ty) brďo*, které je v Ortofonu i Koditexu značkováno jak citoslovce.

<sup>134</sup> Oslovení pomocí kombinace *pane* a příjmení/titul (např. *pane Kolář* místo *pane Koláři*), které není ve vokativu, bývá zpravidla pocíťováno jako neformální až familiární. V korpusu Koditex-web objevilo 8× z 80 takových kombinací. V korpusu Ortofon toto nebylo možné toto porovnání kvůli anonymizovaným příjmením udělat.

*premiére, prezidente, redaktore*). V něm se objevovalo i oslovení většího seskupení osob (např. *armádo, redakce, vládo*), v případě Evropské unie prezentované *proprii Evropo a Brusele*. V psané elektronické komunikaci se zejména mezi hapaxy, tj. tvary slov s jediným výskytem, objevovala apelativa označující (a dehonestující) politické smýšlení adresáta, např. *klerofašisto, muslimko, ruský propagandisto, sluníčkáři, soudružko, vlastenče*.

Nominální vokativy byly v datech nalezeny nejen jednotlivě, ale i v kombinaci s jiným substantivem, adjektivem nebo zájmenem ve stejném pádě. Kombinace dvou substantiv byla častější v psaném jazyce; vyskytovala se v něm zejména oslovení jménem/přezdívkou adresáta doprovázená titulem *pane/paní* (např. *paní Jitko, pane Kresto, pane Sýkoro Marku*). V mluveném jazyce se spojení s těmito tituly týkala pouze apelativ (např. *pane doktore, pane inženýre, paní ředitelko*). Opačný slovosled, tj. vokativ a po něm jeho specifikace (např. *lide obecný*), se vyskytoval ojediněle. Za kombinaci jsme nepovažovali případy se slovy *vole, člověče* (včetně variant) (př. 14).

(14)

- to bylo slyšet ne na tom videu ne + [...] + úplně vole co je vole co je **vole sračko vole** . takhle normálně tam před ním tančil ..  
(Ortofon)

- **Láňo čěče** , já bych se bejt Váma soustředil hlavně na ten americkéj imperialismus

(diskuse)

Z adjektiv se u oslovení v mluveném i psaném jazyce nejčastěji vyskytovala lemmata *vážený a milý*. Ze zájmen se s nominálním vokativem nejčastěji kombinovaly lexémy *ten* a *ty*. V mluveném jazyce se zájmeno *ty* převážně kombinovalo se slovy *kráso, krávo, píčo, vole*; tato spojení nebyla v psané elektronické komunikaci stejně frekventovaná jako v mluveném jazyce.

Vokativy se v datech vyskytovaly také jako součásti ustálených frází, které nepovažujeme za oslovení. V psaném jazyce se jednalo o vokativ *světe* ve spojení *světe div se, ševče* ve spojení *ševče drž se svého kopyta* a *člověče* ve zmínce o hře *člověče nezlob se*. Vokativ *bože* se v mluveném i psaném jazyce vyskytoval ve frázích *nedej bože* a *pane bože* (s různým psaním velkých písmen).

V rámci redukované formy bylo v mluveném i psaném jazyce nalezeno 23 různých oslovení:<sup>135</sup> *Ali, Andy, babi, Eli, Evi, Hani, Heli, Jani, Leni, Luci, Magdi, mami, Marti, Maru, Oli, Pepi,*

---

<sup>135</sup> Slova zvýrazněná tučně byla nalezena pouze v mluveném jazyce. Podtržené slovo bylo nalezeno pouze v psaném jazyce.

*Peti, sestři, Soni, tati, teti, Tery, Tomi, Vali, Veru*. Kromě slov *Pepi, tati* a *Tomi* se jedná o oslovení žen.

U nominálního oslovení byla v obou typech data velmi častá propria. V psaném jazyce se vyskytovala řada apelativ v přezdívce pisatele (např. *kudla, pastelka, suchar, trefa, žirafka*), která nebyla vždy psána s počátečním velkým písmenem. Tyto případy jsme při porovnání také započítali mezi propria.<sup>136</sup> Při porovnání apelativ a proprií se ukazuje velký rozdíl mezi rozsahem apelativních a propriálních seznamů a ve frekvenci, s jakou jsou tyto seznamy zastoupeny v datech. Propria (včetně redukovaných forem) se v mluveném jazyce podílí 52 % a v psaném elektronickém jazyce 64 % na seznamu všech vokativů. Propria však nebyla ani v jednom typu dat výrazně frekventovanější než apelativa: v mluveném jazyce se podílela pouze 15 % na celkové frekvenci vokativu, zatímco v psané elektronické komunikaci 53 %. Tento rozdíl je ovlivněn komunikačními podmínkami srovnávaných jazykových modů: v mluveném jazyce jsou všichni mluvčí fyzicky přítomni na jednom místě, takže „oslovení“ může probíhat i neverbálně, naopak v psané elektronické komunikaci je nutno verbálně označit toho, ke komu je sdělení směřováno.

Mezi žánry elektronické komunikace, které vykazovaly přibližně podobnou frekvenci vokativního nominálního oslovení, se v oblasti nominálního oslovení nejvýrazněji odlišovala diskusní fóra, v nichž se oproti ostatním žánrům nejčastěji objevovaly apelativní přezdívky a oslovení žen (v 66 %). Mimoto se na fórech nevyskytovala politicky motivovaná oslovení.

## Shrnutí

Tato studie ukázala, že nominální oslovení ve formě vokativu se nejvíce uplatňuje v mluveném jazyce. Pokud bychom vyloučili z vokativů výraz *vole*, byla by frekvence tohoto pádu ve všech analyzovaných žánrech podobná. U řady substantiv značkových jako vokativ se při jejich třídění do funkčnėsémantických skupin ukázalo, že je jejich status oslovení z našeho pohledu problematický, nejednoznačný (kromě zmíněného *vole* i další výrazy vyjadřující překvapení), příp. že se uplatňuje pouze v některých žánrech (např. výraz *krávo*). Rozdíly v oslovení mezi mluveným a psaným elektronickým jazykem se dále projevíly v zastoupení funkčnėsémantických skupin a v preferované pozici v rámci segmentu/věty.

---

<sup>136</sup> Psaní velkých písmen je v elektronických textech značně nesystematické. Pokud se apelativní přezdívka vyskytovala aspoň jednou s počátečním velkým písmenem, považovali jsme všechny její výskyty za přezdívku. Apelativa v iniciální pozici ve větě (tedy s velkým písmenem), která by mohla být považována za přezdívku, jsme počítali mezi propria (např. *Zakladatelko*, *hlavu vzhůru*, *zniš jako rozumná holka*). Propria psaná s počátečním malým písmenem jsme také počítali jako propria (např. *Bez urážky*, *juro*, *ale to je směšné* .).

### 3.3.4 Shrnutí kapitoly

V této části jsme analyzovali tři jazykové jevy, jejichž společnou vlastností je adresnost. Vzhledem k povaze analyzovaných dat jsme se nezabývali tím, jakým způsobem a zda vůbec je na tyto prostředky reagováno.

Na míru adresnosti určitého textu lze usuzovat z frekvence jevu, uvážit však je potřeba také to, zda (a nakolik) je daný jev spolehlivým ukazatelem adresnosti; tedy zda je používán pouze pro obrácení se na adresáta, nebo i pro jiné účely. Pokud se zaměříme na první část této problematiky, tj. frekvenční porovnání, můžeme konstatovat, že adresnost je vzhledem ke všem třem jazykovým rysům nejvyšší v mluveném jazyce. V žánrech psané elektronické komunikace se pohybuje (s malými rozdíly) v nižší frekvenční hladině než v mluveném jazyce. Nejedná se o překvapivé pozorování, protože dané psané žánry nejsou z hlediska synchronnosti stejné jako analyzovaná mluvená komunikace (viz [2.1.6](#) a [3.1.1.2](#)).

Co se týče posouzení jednotlivých jevů jako prostředků adresnosti, lze říci, že dva rysy – zejména 2. osoby a nominální oslovení – adresnost vyjadřují ve většině případů, třetí rys – otázky – ji vyjadřuje stoprocentně. U zájmen pro 2. osobu i nominálního oslovení byly do analýzy zahrnuty i výskyty, které jsou již úplně nebo částečně užívané v jiné funkci, např. kolokace *ty jo* nebo oslovení *ty vole*. Analýzy u těchto dvou rysů spíše než na neadresnost některého slova či slovního spojení poukázaly na limity operacionalizace. Pro její zlepšení by bylo nutné vydělit a zvlášť analyzovat slova či kolokace, jejichž současná převažující funkce se liší od původní. U otázek se způsob jejich operacionalizace nakonec ukázal pro naši analýzu jako výhodný: identifikace jevu v datech pomocí otazníku na konci segmentu/věty nevedla pouze k vyhledání otázek, ale i k nalezení případů, ve kterých je otazník připojen ke slovu užitému ve funkci výzvy adresáta k reakci. Za explicitní prostředek adresnosti tedy lze považovat i samotný symbol otazníku.

### 3.4 Rysy spjaté se spontánností

Pro označení jazykových jevů, jejichž výskyt je projevem jazykově nevynuceným, tj. bez vnětetových příčin, používáme pojem spontánnost. V různých jazykových popisech (viz např. výčet rysů konceptuální mluvenosti na s. 15–16) je k němu synonymně často užíván pojem nepřipravenost. Vzhledem k návaznosti této práce na MDA v češtině (Cvrček et al., 2018b), v níž je ve druhé dimenzi kladena spontánnost do protikladu k připravenosti, neuvádíme zde pojem nepřipravenost a ani necítíme nutnost definovat tento pojem v kontrastu ke spontánnosti.

Pojem spontánnost zde užíváme právě ze snahy vyhnout se otázce, co míníme onou „přípravou“ v pojmu (ne)připravenost.

Za jeden z projevů spontánnosti v textech lze považovat linearitu, která podle Müllerové (2011) souvisí s plynutím projevu v čase. Za lineární lze např. považovat kladení slov, slovních spojení a dalších syntaktických struktur za sebe, na první pohled nehierarchicky a s minimálním vyjádřením syntaktických vztahů; v případě návratu k předchozímu vyjádření (nebo jeho opravy, doplnění, rozvinutí) se tento návrat děje neočekávaně, s přerušením dosavadní syntaktické struktury. Opakování slov a výskyt překlepů, přeřeknutí a slovních fragmentů lze nazírat jako signály míry spontánnosti. V této práci však porovnáváme psanou a mluvenou komunikaci, která byla při srovnání různých dalších psaných a mluvených žánrů a situačních typů označena v MDA za blízkou. V následujících dvou studiích se proto pokusíme postihnout, zda je tato blízkost demonstrována vybranými dvěma jazykovými rysy.

### 3.4.1 Opakování slov

Opakování slov představuje další jazykový jev charakteristický pro neformální, spontánní jazyk mluvený (např. Leech, 2000; Müllerová, 1994, 2011) i psaný elektronický (např. Svobodová, 2006). Opakování identických slov, ale i celých frází s různou mírou obměny je v neformální, spontánní komunikaci ovlivněno podmínkami produkce, zejména časovým tlakem, rušným prostředím<sup>137</sup> a krátkodobou pamětí. Na obecné rovině lze opakování usouvztažnit s rutinizovaným užíváním gramatických a lexikálních prvků v případech, kdy to vyžadují společenské normy (např. pozdravy, poděkování), anebo kdy se jedná o ustálený vzorec chování, jehož plnění však není vyžadováno žádnou normou (např. rozloučení v telefonním dialogu) (viz Aijmer, 1996).

V této studii se zabýváme opakováním totožných slov bezprostředně za sebou v rámci jedné věty či segmentu.<sup>138</sup> Tento jev nazývá Karlík (2017) repeticí a rozlišuje případy záměrné a nezáměrné. Záměrnému opakování přisuzuje tři funkce: zesilovací, zdůrazňovací a navazovací. Nezáměrné opakování je podle něj způsobeno zejména nepřipraveností mluvčího, příp. jeho neschopností pohotově reagovat, nebo afektem. Odlišení záměrného a nezáměrného opakování

---

<sup>137</sup> V psané elektronické komunikaci označujeme za rušné takové prostředí, ve kterém jsou jednotlivé texty uspořádané nepřehledně, např. není zřetelná návaznost jednotlivých příspěvků (kdo na koho reaguje), nebo ve kterém probíhá zároveň více dialogů (např. v elektronickém chatu).

<sup>138</sup> Opakování představuje jazykový jev, který je vlastní typologicky různým jazykům a týká se jazykových jednotek z různých jazykových rovin, např. vět a frází, slov, morfémů, hlásek. S jazykovou rovinou souvisí i míra lexikalizace tohoto jevu (podrobněji viz Fischer, 2011: 57).



je tedy dáno tím, zda jsme schopni z textu (a informací o něm) rekonstruovat důvod, který motivoval mluvčího/pisatele k zopakování konkrétních slov.

Vedle toho, že se snažíme na základě kontextu a metainformací o analyzovaném textu odhalit možnou motivaci, se zabýváme analýzou lexikálního obsazení repetice a rozdíly mezi mluveným a psaným modelem. Podle Fischerové (2011) je opakování častým jevem v mluveném jazyce; pokud se vyskytuje i v psaném, častěji se jedná o lexikalizované opakování (např. citoslovce *řízy řízy*). Vyšší frekvenci tohoto jevu stejně jako větší lexikální bohatost, tj. více různých lexémů, tedy očekáváme v datech z mluveného jazyka. Ačkoli se naše mluvená i psaná data dají označit za neformální a spontánní, patří analyzované elektronické texty mezi asynchronní žánry elektronické komunikace; pisatel není v tak velkém časovém tlaku jako v případě synchronních žánrů typu chat, proto předpokládáme nižší frekvenci opakování a malé rozdíly mezi zkoumanými žánry.

## Metodologie

V následující analýze se zaměřujeme na opakování totožných slov bezprostředně za sebou v rámci věty či segmentu, tj. pomíjíme opakování víceslovných úseků a případy, ve kterých jsou mezi opakování téhož výrazu vložena jiná slova. Ve všech analyzovaných datech jsme zvláště vyhledali opakující se n-gramy obsahující 2–6 identických slov (myšleno slovních tvarů, nikoli lemmat), tj. dvojice (dále bigramy),<sup>139</sup> trojice (trigramy), čtveřice (tetragramy), pětice (pentagramy) a šestice (hexagramy). Při vyhledávání jsme tato slova nijak morfologicky nespecifikovali, např. omezením na stejný slovní druh. Pro detailnější analýzu lexikální bohatosti jsme využili především bigramy (tab. 7).

Opakování stejných slov potvrzuje první Zipfův zákon, který popisuje vztah mezi frekvencí např. slova a jeho pořadím (rankem) (Cvrček, 2017). V obou jazykových modech platí, že čím vyšší počet stejných slov za sebou, tím méně často se celé toto spojení vyskytuje (tab. 7). Častější výskyt n-gramů opakujících se slov v mluveném jazyce oproti psanému, byť v tomto případě redukovaném na elektronickou komunikaci, je podle našeho názoru dán větší růzností mimojazykových faktorů ovlivňujících mluvený projev.

---

<sup>139</sup> Dotaz pro psaný jazyk: 1:[] 2:[word="(?) [a-zšščřďtňůěýáíéó]+"] 3:[word="(?) [a-zšščřďtňůěýáíéó]+"] 4:[] & 1.word!= 2.word & 2.word = 3.word & 3.word!= 4.word within <s/>.

Dotaz pro mluvený jazyk: 1:[] 2:[word="(?) [a-zšščřďtňůěýáíéó]+"] 3:[word="(?) [a-zšščřďtňůěýáíéó]+"] 4:[] & 1.word!= 2.word & 2.word = 3.word & 3.word!= 4.word within <sp/>.

Příslušný počet n-gramů je v těchto dotazech označen na druhé a třetí pozici; pro vyhledání n-gramů je nutné přidat totožnou část dotazu vždy na další pozici a změnit číselné označení pozic.

	Koditex-web		Ortofon	
	lemmata	slovní tvary	lemmata	slovní tvary
bigramy	68	70	465	591
trigramy	12	12	91	104
tetragramy	1	1	25	27
pentagramy	0	0	11	13
hexagramy	0	0	1	6

Tab. 7. Počet n-gramů obsahující totožná slova

Opakování slov, mezi něž je v mluveném jazyce vložena pauza, se nevěnujeme ze dvou důvodů.<sup>140</sup> Prvním je snaha zachovat kontinuitu ve vztahu k formální podobě, tj. zápisu, mluveného a psaného jazyka. Jakákoli nefinální interpunkce vložená mezi tatáž slova ve větě či segmentu je brána jako přerušení řetězce. Pokud bychom do vyhledávání chtěli zahrnout i vložení pauzy, museli bychom vložit do dotazu pro psané elektronické texty nefinální interpunkční znaménka.<sup>141</sup> Tím bychom však opominuli případy označované jako parcelace textu, tj. rozdělení delší věty do několika kratších úseků oddělených finální interpunkcí. Druhý důvod je pragmatický: v obou jazykových modech se u bigramů snažíme určit důvod, proč k opakování slov došlo; počet těchto bigramů přitom nepřesahuje možnosti manuální analýzy.

Z výsledků byly ještě před analýzou manuálně odstraněny případy, ve kterých je opakování součástí názvu, např. filmu *Wild Wild West*, hudební skupiny *the Tap Tap* apod., nebo je dané chybou při zpracování dat do korpusu, např. @ *kat kat píše : Díky*. Vyřazeny byly také anonymizační značky používané v mluveném korpusu pro anonymizaci křestních jmen (značka NJ) a příjmení (značka NP), protože se pod každým výskytem může skrývat jiné jméno, např. segment *říkám zazvoň NP NP zazvonil*. Počet a různost opakujících se lexémů v bigramech ukazuje tabulka 8.

	lemmata		slovní tvary	
	ponecháno	vyřazeno	ponecháno	vyřazeno
Koditex-web	68	4	60	4
Ortofon	438	27	562	29

Tab. 8. Počty opakujících se bigramů

<sup>140</sup> Výskyt bigramů s vloženou pauzou je přibližně dvojnásobný. Kromě případů podobných př. 15 získáme výskyty, v nichž každé opakující se slovo má jinou funkci, např. v segmentu, reagujícím na otázku *myslíš Italy?*, *Italy no .. no ale právě že @ tam bude strašně moc lidí* je první *no* ukončovacím prvkem odpovědi, zatímco druhé *no* je startovacím výrazem nové informační části segmentu.

<sup>141</sup> Výskyt bigramů s nefinální interpunkcí vloženou do věty je přibližně dvojnásobný. Kromě žádoucích výskytů, např. *Fakta , fakta , fakta*, dostaneme větší množství případů dále označených jako gramatické opakování, např. *Pokud takový zákon bude , bude to jednoduché*.

## Motivace pro opakování v mluveném a psaném elektronickém jazyce

Odlišení záměrného a nezáměrného opakování včetně jejich podskupin (Karlík, 2017) je dáno tím, zda jsme schopni rekonstruovat z textu důvod, který motivoval mluvčího/pisatele k zopakování konkrétních slov. Kromě toho je možné vycházet z lexikálního významu, příp. významu morfologické kategorie: za vždy záměrné lze podle našeho názoru označit např. opakování slov, která jsou užívána k upoutání pozornosti adresáta, např. pomocí oslovení (vokativem i citoslovcem *hele*), imperativu (př. 15), sloves ve 2. osobě, či k signalizaci pozornosti věnované aktuálnímu mluvčímu, např. pomocí responzního slova (*backchannel*) *hmm*.

(15)

Iva\_804: jak může být ? [mluvíme o Karlově náměstí]

Bartoloměj\_805: [počkej . **počkej počkej**]

Vzhledem k tomu, že jsme při vyhledávání nijak nspecifikovali opakující se slova, dostaly se do výsledků i případy, v nichž se neopakují dvě slova na téže syntaktické pozici a v nichž je každé slovo jiným slovním druhem. Mimoto dotaz nspecifikuje ani případy dvouslovných citoslovcí skládající se z totožných slov, např. *řízy řízy*, příp. citoslovcí objevujících se nejčastěji v takovém bigramu, např. *haf haf*. Tyto případy pracovníě označujeme jako gramatické opakování, protože je součástí jazyka, tj. je lexikalizováno nebo dáno gramatickými pravidly. V protikladu ke gramatickému rozlišujeme opakování pragmatické, dále dělené na podskupiny záměrného a nezáměrného opakování (viz výše). Výskyt gramatického opakování jsme předpokládali především u bigramů, což bylo jedním z důvodů pro samostatné vyhledávání bigramů, trigramů atd.

Případy tzv. gramatického opakování se vyskytovaly ve všech analyzovaných žánrech, týkaly se však relativně malého počtu lexémů (viz dále). Mezi případy gramatického opakování řadíme frazémy (*mít něčeho*) *až až*, příp. *být* + adjektivum/pasivní participium + *až až*, (*být*) *fifty fifty*, (*jen*) *tak tak*,<sup>142</sup> a citoslovce. Z citoslovcí<sup>143</sup> sem patří onomatopoické výrazy napodobující zvířata (*haf*) a různé zvuky (*tuk, tů*), dále lexikalizované výrazy pro smích (*ha, hi*), řeč (*bla*), zpěv (*la, ta, tu, ty*)<sup>144</sup> či psychický stav (úleva – *uf*, starost – *běda, bože*). Další

---

<sup>142</sup> Za frazém považujeme pouze takové opakování stejných slov, u kterého je význam celku jiný než význam jeho částí, srov. rozdíl mezi *udělal to jen tak* (jednoduše, bez hlubšího důvodu) a *udělal to jen tak tak* (stěží).

<sup>143</sup> U opakování citoslovcí, zejm. onomatopoických, nejsme schopni určit, zda existuje sémantický rozdíl mezi jedním výskytem nebo bigramem. Pokud se jedná o více než dvojnásobné opakování, je podle našeho názoru motivováno snahou o vyjádření časového trvání zvuku nebo jeho intenzity.

<sup>144</sup> Zpěv se objevil pouze v mluveném jazyce, např. *jety jety ty ry ty ty .. tu tu dudu tutu tududu*.

skupinu tvoří případy souvýskytu dvou totožných slov v jiných funkcích, např. ve spojení *dohodnout se se zástupci* je každé z opakujících se slov jiným slovním druhem, ve větě *o to to bude horší* došlo k souvýskytu zájmena *to* v pozici předmětu a podmětu z rytmických důvodů. Slovo *to*, jehož frekvence výskytu je mezi bigramy v mluveném jazyce nejčastější a v psaném elektronickém jazyce druhá nejčastější, je v rámci všech opakujících se slov specifické, neboť jeho výskyty mohou být zařazeny mezi gramatické i pragmatické (viz dále).

U opakování označeného jako pragmatické uvažujeme o mimojazykových příčinách jeho výskytu. V psaných elektronických textech jsme podle výchozího předpokladu označili tento jev častěji za záměrný. Na rozdíl od mluvení jsou při psaní všechna produkovaná slova vidět, takže se pisatel může zastavit a přemýšlet. Záměrné opakování bylo v psaných textech použito k intenzifikaci (např. *tak moc moc držím palečky*) a ke zdůraznění (např. *Čili máte na výběr - buď houkačka u jednoho zaměstnavatele nebo nic , prostě nic nic nihil , velké mazlavé prd , soudruzi*).<sup>145</sup> K vyjádření vyšší míry intenzity byla použita adverbia *moc*, *mnohem* a adjektivum *nej* (ve spojení *vše nej nej*), která vyjadřují lexikálně totéž. Mimoto se v psaných datech objevilo opakování adverbia *pomalinku* pro naznačení menší intenzity. Jako intenzifikační jsme dále hodnotili opakování pozdravů *pa* a *pá* (v mluvených datech i *ahoj*, *čau*). Zdůrazňovací funkci jsme identifikovali u následujících výrazů: u responzního výrazu *jo*, sloves *nevím* a *vidíš*, adverbii *opravdu* a *enem* (ve větě *Toš to belo enem enem kuli tomu precedovi*), adjektiva *velký* a výrazů *hele* a *pane* sloužících k oslovení anebo žádosti o pozornost.

V případech, ve kterých jsme opakování nedokázali přiřadit k žádnému z výše jmenovaných důvodů, považujeme za příčinu výskytu tohoto jevu v psaných datech nedostatečnou (či zcela chybějící) editaci textu, a tyto příklady považujeme za nezáměrné opakování. Mezi nejčastějšími lemmaty se v tomto případě objevovaly jednopísmenná slova *a* (*Chtělo by to nějakou celospolečenskou diskusi a a možná by nebylo od věci*), *k*, *i*, *s*, *v* a jednoslabičné tvary sloves *být* a *mít*, např. *Pracovala jsem jsem 45 let , platila daně; Pan Jílek to měl měl u pana doktora taky jednu dobu těžký*. Jako nezáměrné byly označeny ještě dva případy z diskusních fór, v nichž se totožná slova objevují vedle sebe vinou chybějící interpunkce, resp. značek ohraničujících konce vět: *@ ladyjurasice přesně tak tak asi kvůli toho praštěného počasí nám stávkuje tělo; on má tak nějak vůbec pravdu pravdu téměř ve všem co říká*.

---

<sup>145</sup> Při určování rozdílu mezi intenzifikací a zdůrazněním v těchto příkladech jsme vycházeli z lexikálního významu opakujícího se slova. Pokud vyjadřovalo míru či velikost, kterou je možno stupňovat, označili jsme slovo a jeho opakování za intenzifikaci a vice versa.

Záměrné a nezáměrné opakování nelze v mluveném jazyce na rozdíl od psaného vždy jednoznačně odlišit, přestože si můžeme daná slova poslechnout. Příklad 16 ilustruje opakování výrazu *že*, jehož užití lze označit za záměrné i nezáměrné: mluvčí Kristina opakuje daný výraz, protože nemá předem připravenou formulaci, čímž si zároveň snaží udržet slovo a získává čas pro vyjádření myšlenky.

(16)

Kristina\_356: a Babiš říkal .. že chce do opozice .. [říkal . no a .. ale] **že že** [to] -

Alois\_360: [no do opozice ne . ale @ spíš] .. [spíš] to vypadá že .. socani budou mít vládu

Nezáměrné opakování, které je dáno především nepřipraveností projevu (v př. 17 a 18 je konkrétním důvodem neschopnost mluvčího vzpomenout si na konkrétní výraz, resp. navázat na předchozí kontext), se tedy z hlediska analýzy širšího interakčního kontextu může jevit i jako záměrné. Z toho důvodu se v mluveném jazyce dále nezabýváme rozlišováním záměrného a nezáměrného opakování. Toto dělení je kromě analýzy kontextu nutné usouvztažnit i s významem opakujících se lexémů.

(17)

dělá na té **na na** .. @ Táborské v Jechu

(18)

že si tam na mě udělají chvilku čas jako jest\* . **jak jak** a co a dál nebo jestli to mam nějak posunout

### **Srovnání lexikálního obsazení v mluveném a psaném jazyce**

Počet lexémů v bigramech je ovlivněn velikostí zkoumaných dat, tj. korpusů. Mluvený korpus Ortofon je přibližně čtyřikrát větší než psaný korpus Koditex-web (viz [3.1.1](#)). V bigramech z mluveného korpusu se vyskytuje 465 různých lemmat, zatímco v psaných elektronických textech pouze 68 (tab. 7). Dále se proto v této části zaměříme na srovnání 30 slovních tvarů, které se v bigramech objevovaly nejčastěji (tab. 9).

	Ortofon	Koditex-web		Ortofon	Koditex-web
1	<b>to</b>	<b>se</b>	16	<i>v</i>	<i>táta</i>
2	<i>a</i>	<b>to</b>	17	<i>ta</i>	<b>uf</b>
3	<i>že</i>	<i>a</i>	18	<i>je</i>	<i>spokojení</i>
4	<i>na</i>	<b>tak</b>	19	<i>ale</i>	<i>nakročeno</i>
5	<b>tak</b>	<i>moc</i>	20	<i>z</i>	<i>i</i>
6	<i>no</i>	<i>pa</i>	21	<i>jako</i>	<b>az</b>
7	<b>ty</b>	<b>šup</b>	22	<i>se</i>	<i>nasla</i>
8	<i>jo</i>	<b>až</b>	23	<i>už</i>	<b>fifty</b>
9	<i>jak</i>	<i>s</i>	24	<i>i</i>	<i>nine</i>
10	<i>ne</i>	<i>nevím</i>	25	<i>s</i>	<i>děláme</i>
11	<i>já</i>	<b>bla</b>	26	<i>hmm</i>	<i>je</i>
12	<i>co</i>	<i>atd</i>	27	<i>nebo</i>	<i>pane</i>
13	<i>ten</i>	<i>nic</i>	28	<i>ještě</i>	<i>romány</i>
14	<i>do</i>	<i>nula</i>	29	<i>jestli</i>	<i>hele</i>
15	<i>tam</i>	<i>jsem</i>	30	<i>za</i>	<i>mnoho</i>

Tab. 9. Třicet nejčastějších slovních tvarů objevujících se v bigramech s označením jejich příslušnosti do skupiny gramaticky (označeno tučně) a pragmaticky (označeno kurzivou) motivovaného užití.

Z tabulky 9 je patrný jednak rozdíl mezi mluveným a psaným modem v délce slov, jednak počet slov zařazený do skupiny gramaticky, nebo pragmaticky motivovaného užití. Zatímco v mluveném jazyce jsou mezi první dvacítkou maximálně třípísmenná slova, která jsou kromě spojky *ale* jednoslabičná, a teprve v následující desítce se objevují slova se čtyřmi a více písmeny, tedy dvouslabičná, v psaném jazyce je situace odlišná. Více než třípísmenná slova se vyskytují od desátého místa a v seznamu prvních dvaceti se vyskytují i dvě devítipísmenná, tedy čtyřslabičná slova. Častější užívání jednoslabičných slov v mluveném jazyce můžeme považovat za doklad principu jazykové ekonomie, který se v tomto modu projevuje mj. redukováním<sup>146</sup> víceslabičných slov. Výskyt delších slov v analyzovaných psaných textech není překvapivým zjištěním, jednak kvůli jejich asynchronní povaze, jednak kvůli možnému využití tzv. našeptávače při psaní na mobilu nebo tabletu.<sup>147</sup>

Na rozdílném počtu nejčastějších třiceti lexémů, jejichž opakování je motivováno gramaticky, se podle našeho názoru nejvíce projevuje právě velikost použitých korpusů. S větším objemem dat z psaného elektronického jazyka se dá očekávat vedle vyšší frekvence již nalezených lexémů i větší zastoupení pragmaticky motivovaného užití jiných opakujících se lexémů. Vyhledáme-li frazémy a citoslovce s dvěma totožnými lexémy z tabulky 9 ve stomilionovém psaném korpusu SYN2015, zjistíme, že jejich frekvence nepřekračuje stovky výskytů (konkrétně čtyř set výskytů). Jedná se sice o hrubé srovnání s odlišným typem psaných textů

<sup>146</sup> Redukování se může projevit např. vynecháním či spojením hlásek, ale i rychlejší výslovností dlouhého slova.

<sup>147</sup> Údaj o použitém elektronickém zařízení není v metadatech uveden.

(SYN2015 obsahuje publikované fikční, odborné a publicistické texty), přesto je podle našeho názoru i takováto ukázka dostatečná.

Ve srovnání nejčastějších třiceti slov v bigramech se ukazuje rovněž rozdíl mezi slovními druhy v mluvených a psaných elektronických textech, byť je příslušnost ke slovním druhům relativizována funkcemi, které bigram v textu plní (viz rozdělení na gramaticky a pragmaticky motivované užití). Při srovnání všech bigramů se ukazuje, že častěji se v obou komunikačních modech objevují bigramy složené z funkčních slov (zájmen, předložek, spojek, částic a citoslovcí). Výrazný rozdíl mezi oběma mody je v častějším zastoupení substantiv v psaných textech.

Rozdíl mezi žánry psaných elektronických textů nebyl ani v lexikálním, ani slovnědruhovém obsazení bigramů téměř žádný. V textech z diskusí a Facebooku se mezi prvními třemi opakovanými slovy objevovaly *to, se, a*, ve fórech se do nejčastější trojice dostalo slovo *moc* místo *a*, které je v pořadí čtvrté. Slova *to, se, a* představují jediné tři lexémy ve všech žánrech. Jedinou výraznější spojnicí mezi seznamem slov v bigramech tvoří v textech z Facebooku výrazy, která jsou odrazem interakce – responzní výraz *jo* a kontaktní výrazy *hele, pa, pá, pane*. Tím se facebookové texty podobají mluveným textům, v nichž je nápadně časté opakování responzních<sup>148</sup> slov (*jo, ne, no, hmm*),<sup>149</sup> vyjadřujících jak explicitní (ne)souhlas, tak pozornost věnovanou řeči jiného mluvčího (*backchannel*).

Ve stejnoslovných tri- až hexagramech jsme jak u mluveného, tak u psaného jazyka nenašli žádný případ, který bychom označili za motivovaný gramaticky. V seznamech opakujících se tří, čtyř, pěti i šesti stejných slov v mluveném jazyce se shodně vyskytovaly výrazy *jo, ne, no*. Z dalších jednoslabičných slov, která se objevovala i v seznamu nejčastějších bigramů, lze najít např. *a, co, hmm, jak, je, na, ta, to, tak, ty, že*. Z víceslabičných slov byl nejčastěji opakován rozkazovací tvar *počkej* a responzní výraz *jasně* (k funkcím oběma výrazů viz Čermáková et al., 2019: 270–274, 280–283).

## Shrnutí

Opakování stejných slov bezprostředně za sebou je např. Müllerovou (2011) označováno za jednu z charakteristik spontánního, nepřipraveného mluveného jazyka, což potvrdila i MDA

---

<sup>148</sup> Označení přejímáme podle Čermákové et al. (2019), které analyzovaly funkce daných výrazů v mluveném jazyce.

<sup>149</sup> Slova *jo, ne, no*, resp. jejich čtyřnásobné opakování, se objevují v seznamu nejčastějších čtyřslovných kombinací (bez vložené pauzy) z korpusu ORAL, který je pětinašobně větší než korpus Ortofon (Čermáková & Koprivová, 2018).

(Cvrček et al., 2018b). Opakování slov může být kromě nepřipravenosti způsobeno i jinými důvody, např. Karlík (2017) rozlišuje dvě skupiny podle toho, zda je opakováním užíváno záměrně k realizaci určité funkce. Jak jsme v této kapitole ukázali, takto pojímanou záměrnost je možné aplikovat na psané texty, ale u mluvených textů je situace komplikovaná vlivem opakování na časový průběh interakce. Ukázali jsme, že v mluveném jazyce je možné i tzv. nezáměrnému opakování přisoudit záměr, resp. mluvčí dokáže i nezáměrné opakování využít ve svůj prospěch – udrží si slovo, získá delší čas pro formulaci myšlenky nebo pro nalezení vhodného výrazu.

Hledáním motivace pro opakování slov v psaných elektronických textech a analýzou lexikálního obsazení opakujících se n-gramů jsme se snažili najít podobnosti s nepřipraveným mluveným jazykem a také vzájemné rozdíly. V psaném jazyce jsme vzhledem k vizuálnímu kanálu očekávali pouze opakování, kterému lze na základě kontextu přisoudit záměr, opakování lexikalizované (frazémy a dvouslovná citoslovce) a souvškyt dvou totožných slov v různých syntaktických funkcích. V případech záměrného opakování jsme dále rozlišili dvě hlavní funkce: zesilovací a zdůrazňovací. Opakování, jemuž jsme v elektronických textech nedokázali přisoudit důvod, podle našeho názoru poukazuje na skutečnost, že pisatel již vyprodukovaný text nečetl. Tyto případy, označené v naší analýze za nezáměrné opakování, byly však méně časté (zahrnovaly 7 různých slovních tvarů z celkově nalezených 70) než případy záměrného opakování. Můžeme tedy konstatovat, že analyzované psané texty vykazují z hlediska opakování slov určitou míru spontánnosti, ale vzhledem k nízké frekvenci nemůžeme tento jev označit za tak výrazný indikátor spontánnosti jako u mluveného jazyka.

### **3.4.2 Fragmenty slov, překlepy a přeřeknutí**

Součástí produkce mluveného i psaného textu jsou vždy úpravy, jejichž možnosti se výrazně liší především mezi jazykovými mody. U psaného textu vytvářeného na počítači nebo jiném elektronickém zařízení je možné stávající text přepsat, doplnit, posunout apod., zatímco u textu psaného rukou jsou možnosti úprav závislé na použitých psacích potřebách (jak na látkách, kterými se píše, např. inkoustové pero, obyčejná tužka, propisovací tužka, tak na látkách, na které se píše, např. čtvrtka, kancelářský papír, tkanina). Při produkci mluveného textu není možné již vzít zpět to, co bylo vyřčeno, takže jakékoli úpravy, opravy, doplnění následují po již vyřčeném a je v zájmu mluvčího, aby takové úpravy signalizoval (např. lexikálně, intonačně). V této studii se nezaměřujeme na úpravy, opravy či doplnění, ale na jejich pravý opak, tedy na to, co upraveno nebylo, ačkoli bychom to očekávali. Pohybujeme se pouze na úrovni slov, tj. shluku hlásek ohraničeného z obou stran mezerou, a zaměřujeme se pouze na



jejich formální stránku, tzn. neposuzujeme jejich spojování do vyšších celků ani adekvátnost užití v daném kontextu.

Budeme se zabývat výskyty slov, které se dvojnásobným způsobem odchylně od své standardní podoby: 1) ve slově došlo k přesmyku hlásek (pro psaný jazyk dále označováno jako překlepy, pro mluvený jazyk jako přerěknutí, např. *zkušensot*), 2) v textu se vyskytuje nedokončené slovo (dále tyto případy nazýváme fragmenty slov, např. *hemoglobin rozváš\* roznáší kyslík po těle*). V psaném jazyce řadíme k překlépům i případy, ve kterých došlo k chybnému umístění mezery, a tedy k rozpojení slova na fragmenty (např. *nech ápu*), nebo ke spojení dvou a více slov do jednoho celku (např. *žeby*). Ve všech případech předpokládáme, že ke vzniku překlépů/přerěknutí i fragmentů slov dochází neúmyslně a jejich ponechání v psaném textu interpretujeme ve shodě s Čmejrkovou (1997) jako projev lineárního přístupu k textu: pisatel již nečte (a tedy ani neupravuje) již napsané. Vznik těchto jevů v mluveném jazyce může být motivován různě: produkčními obtížemi zapříčiněnými fyziologicky (např. koktáním), psychicky (např. nervozitou), situačně (např. vyrušením jiným mluvčím); při jejich posuzování však musíme mít na paměti, že mluvčí neměl možnost „vyjmout“ tyto jevy z již vyprodukovaného textu. Záměrem této studie není odhalit faktory, které se na vzniku jevů podílejí, ale jevy popsat a zkusit nalézt podobnosti a odlišnosti mezi mluveným a psaným jazykem.

### **Metodologie**

Vyhledání obou jevů probíhalo v každém komunikačním modu různě. V psaných elektronických textech jsme využili morfologické značkování. Zajímala nás neoznačovaná<sup>150</sup> slova, která se v korpusové lingvistice většinou označují jako nerozpoznaná (v tagu je uvedeno X a lemma je totožné s tvarem slova). Seznam nerozpoznáných slov byl poté manuálně roztríděn do skupin podle pravděpodobného důvodu jejich nerozpoznání.

Při třídění jsme nepracovali s celým seznamem, ale pouze s jeho částí, protože jsme slova třídili do skupin ručně. Nejprve jsme analyzovali lemmata s minimální frekvencí 2 v daném žánru (v tab. 10 řádek nazvaný mezisoučet). Vzhledem k nízkému zastoupení žádoucí skupiny, tj. překlépů a fragmentů slov (tab. 11), jsme poté analyzovali ještě 5 % náhodně vybraných nerozpoznáných lemmat s frekvencí 1.

---

<sup>150</sup> U obou korpusů jsme použili značkování provedené systémem MorphoDita (Straková et al., 2014), tj. dotaz vypadal takto: [mtag="X.\*"]. Morfologické značkování jakožto automatická procedura není nikdy stoprocentní. Za problematické lze označit především vlastní jména, cizí slova a novotvary, které jsou příznačné pro idiolekt autora či mluvčího. K popisu procesu lemmatizace a značkování viz Petkevič (2017).

		diskuse	fóra	Facebook
počet nerozpoznaných lemmat s frekvencí <sup>151</sup>	≥ 10	9	64	16
	< 10 & ≥ 5	49	154	46
	< 5 & ≥ 3	145	285	102
	= 2	289	543	239
	mezisoučet	492	1 046	403
	= 1	3 665	5 511	3 054
celkem lemmat		4 157	6 557	3 457
<b>celkem tříděno lemmat</b>		<b>697</b>	<b>1342</b>	<b>577</b>

Tab. 10. Počet různých nerozpoznaných tvarů v analyzovaných psaných datech

Při třídění bylo kromě překlepů a fragmentů slov identifikováno patnáct dalších skupin (tab. 11). Hledané skupiny byly ve všech žánrech elektronické komunikace zastoupeny méně než v 10 %, což znamená, že se řádově jednalo o jednotky či desítky výskytů. Nejvíce byla v seznamu nerozpoznaných slov zastoupena slova zapsaná bez diakritiky (např. *zrejme*) a propria (např. *Kkklarka*, *Tsipras*). Pro představu uvádíme příklady slov v ostatních skupinách: citoslovce (např. *mno*, *muhehe*), nářeční výrazy (např. *ešče*, *chodija*), zdomácnělá slova cizího původu (např. *dislajkiju*, *nickem*), slova v cizím jazyce (např. *review*, *Waffen*), slova patřící do idiolektu (např. *hodinkovnice*, *zlistinil*), slova s pravopisnou chybou (např. *taki*, *zapoměla*), slova s fonetizovaným pravopisem, tj. zapsaná podle výslovnosti (např. *mooc*, *přeceda*, *uš*), symboly (např. °, ¨) univerbizáty<sup>152</sup> (např. *hysterák*, *živoiko*), vlastní jména s chybným malým počátečním písmenem (např. *evropu*), zkratky a zkratková slova (např. *cz*, *mmchdm*).

	diskuse	fóra	Facebook
citoslovce	1 %	0,5 %	0,9 %
cizí jazyk	13,2 %	4,8 %	7,8 %
diakritika	30 %	49 %	23 %
fonetizace	2,3 %	2,8 %	3 %
<b>chybná mezera</b>	<b>4,7 %</b>	<b>5,6 %</b>	<b>3,1 %</b>
chyba	1,6 %	0,6 %	4,5 %
idiolekt	7,7 %	4 %	6,8 %
nářečí	1,3 %	1 %	8,3 %
nevím	0,1 %	0 %	0 %
odkaz	1,1 %	2,5 %	1,2 %
propria	16,9 %	19,2 %	15,3 %
<b>překlep</b>	<b>3,9 %</b>	<b>2,4 %</b>	<b>7,3 %</b>
symbol	0,6 %	0,2 %	1,6 %
univerbizát	0,1 %	0,5 %	0,2 %

<sup>151</sup> Matematický zápis vytvořen podle <https://www.karlin.mff.cuni.cz/~portal/logika/?page=konjunkce>.

<sup>152</sup> Pojem převzat od Kolářové a Štíchy (2019).

velké písmeno u proprií	3,7 %	0,5 %	4,7 %
zdomácnělé cizí slovo	2,4 %	1,7 %	3,8 %
zkratka	9,3 %	4,8 %	8,7 %

Tab. 11. Procentuální zastoupení skupin v seznamu nerozpoznaných slov v psaném jazyce

V mluveném jazyce jsou oba jevy značeny již při transkripci.<sup>153</sup> V korpusu Ortofon je pak toto značení zveřejněno následovně: u přerěknutí pomocí strukturní značky *para type*,<sup>154</sup> u fragmentů jako tag F (v transkriptu jsou všechny fragmenty označeny hvězdičkou, která je přidána k vysloveným hláskám, např. *to sis mohl v\* do\* vyloženě vybrat*). Na rozdíl od fragmentů může značka pro přerěknutí zahrnovat i více než jedno slovo, protože v některých případech je těžké rozhodnout, ve kterém slově přesně k přerěknutí došlo, a tvůrci korpusu chtěli zachovat lokaci tohoto jevu i za cenu nepřesnosti a vágnosti značení. Součástí přerěknutí tak mohou být i fragmenty, např. *tak jak by se líbilo cizincům když nerozuměj <PR t\* tem t\* tem> .. těm textům*. Frekvenci obou jevů ukazuje tabulka 12, zatímco tabulka 13 uvádí počet slov včetně jejich slovnědruhového tagu v rámci přerěknutí. V následující analýze se u přerěknutí budeme zabývat pouze nerozpoznanými slovy.

		frekvence
fragменты		13 708 (11 086,06 ipm)
přerěknutí	značka PR	556 (449,65 ipm)
	počet slov v rámci značky	809 (654,26 ipm)

Tab. 12. Frekvence obou jevů v mluveném korpusu Ortofon

	slovní druh	absolutní frekvence
1	nerozpoznáno	268
2	fragment	164
3	předložka	72
4	substantivum	71
5	zájmeno	68
6	sloveso	59
7	spojka	29
8	příslowce	27
9	adjektivum	17
10	interpunkce	15
11	částice	12

<sup>153</sup> Cílem této studie není ukázat, nakolik je značení těchto jevů správně.

<sup>154</sup> Dotaz tedy vypadá následovně: `<para type="přerěknutí"/>`. V příkladech z mluveného korpusu Ortofon zahrnujeme přerěknutí do špičatých závorek se značkou PR, tj. `<PR takto>`.

12	číslovka	5
13	hezitace	1
14	citoslovce	1

Tab. 13. Frekvence slov podle slovnědruhového značení v rámci přeřeknutí

## Analýza

### Překlepy a přeřeknutí

Při porovnání jednotlivých skupin vyčleněných v rámci nerozpoznaných slov se ukazuje, že překlepy tvoří ve všech třech žánrech elektronické komunikace málo frekventovanou skupinu (tab. 11). Toto srovnání nemůžeme v mluveném jazyce udělat, protože přeřeknutí byla hledána jiným způsobem (viz s. 81). Překlepy a přeřeknutí můžeme vzájemně porovnat podle toho, v čem došlo k odchylce od standardní podoby slova. Zatímco u překlepů byly tyto odchylky rozděleny do čtyř skupin, u přeřeknutí do třinácti včetně skupiny sdružující taková přeřeknutí, u kterých jsme nedokázali říci, čím se odchylují od standardní podoby, tj. jak slovo standardně vypadá.

Odchylky identifikované u překlepů byly zároveň nalezeny i u přeřeknutí. Jedná se o 1) chybějící písmeno/hlásku, 2) záměnu písmene/hlásky, 3) nadbytečné písmeno/nadbytečnou hlásku, 4) změnu v pořadí písmen/hlásek. Záměna písmene/hlásky byla nejčastější skupinou všude kromě diskusních fór, ve kterých byly stejným počtem zastoupeny případy chybějícího písmene a změny pořadí písmen ve slově (tab. 14). V mluveném jazyce tvořila záměna hlásek téměř polovinu všech přeřeknutí. Záměny hlásek jsme už dál netřídili podle konkrétních hlásek, uvádíme zde pouze příklady z mluveného jazyka: záměna krátkého a dlouhého vokálu, např. *vázáný* (*vázany*), záměna vokálů stejné kvantity, např. *všichny* (*všechny*), *bedou* (*budou*), záměna hlásek lišících se měkkostí např. *úplňá* (*úplná*), *zjistili* (*zjistili*), *vysle* (*vyšle*), záměna alveolár *r* a *l*, např. *akolát* (*akorát*), *faláři* (*faráři*), *Berlinare* (*Berlinale*), záměna sykavek a polosykavek, např. *netučím* (*netuším*). V psaném jazyce se záměna týká především těch písmen, která se na klávesnici nacházejí vedle sebe, např. *nuset* (*muset*), *ptoblémy* (*problémy*), *šetla* (*četla*), *tesy* (*tedy*). Slovo *Ruzyňské* ukazuje, že k záměně může dojít i kvůli složitějšímu psaní na klávesnici (pro ň je nutno stisknout kombinaci kláves).

popis odchylky	diskuse	fóra	Facebook	Ortofon
chybějící písmeno/hláška	28 % (7)	31 % (10)	33 % (14)	5 % (11)
chybějící slabika	-	-	-	1 % (2)
nadbytečná slabika	-	-	-	9 % (19)
nadbytečné/á písmeno/hláška	24 % (6)	13 % (4)	12 % (5)	18 % (39)
nahrazení jiným slovem	-	-	-	0,5 % (1)
neurčeno	-	-	-	8 % (17)
spojení několika slov	-	-	-	0,5 % (1)
zakotání	-	-	-	1 % (3)
záměna písmene/hlásky	44 % (11)	25 % (8)	38 % (16)	48 % (102)
záměna slabiky	-	-	-	1 % (3)
zkomolení slova	-	-	-	4,5 % (9)
změna v pořadí písmen/hlásek	4 % (1)	31 % (10)	17 % (7)	2,5 % (5)

Tab. 14. Rozdělení překlepů/přeřeknutí podle typu odchylky od standardní podoby slova včetně jejich frekvence v datech (čísla uvedená v závorce ukazují absolutní frekvenci)

U dvou skupin společných pro oba komunikační mody – slov s chybějícím písmenem/hláskou a se změnou pořadí písmen/hlásek ve slově – byl zjištěn vyšší podíl v psaném jazyce oproti mluvenému. Zatímco u psaného jazyka můžeme obě odchylky přiřknout rychlosti psaní, u mluveného jazyka se tento důvod jeví jako pravděpodobnější spíše u první skupiny (ačkoli vynechání hlásky se týká i slova *trece*, tj. *trenclé*, které bylo vysloveno zřetelně, neredukovaně a relativně pomalu). Ve skupině zde zmiňované jako druhé se objevují tři původem cizí slova – *iniciativy* (*iniciativy*), *prementku* (*permanentku*), *recivisty* (*recivisty*), u nichž nelze vyloučit, že k záměně hlásek došlo vinou neznalosti, a dvě česká slova – *aokrát* (*akorát*), *sleským* (*selským*).

O poslední skupině společné pro oba komunikační mody (nadbytečné písmeno ve slově) se nedá říci, pro který modus je typičtější z hlediska frekvenčního podílu mezi nerozpoznanými slovy. V psaném jazyce je nadbytečným písmenem často to, které se nachází na klávesnici vedle písmena objevujícího se ve slově náležitě, např. *váchova* (*výchova*), nebo je totožné písmeno použito dvakrát, např. *intelligence*, *tto*. U slov s nadbytečnou hláskou, která byla nalezena v mluveném jazyce, byla při poslechu často patrná změna tempa a zaváhání. Co se týče

konkrétních hlásek, vokál *u* byl několikrát vysloven jako dvojhláska *ou*, např. *koupuju* (*kupuju*) a sykavka *s* byla protažena do *š*, např. *společensšší* (společenšší). V několika dalších případech jsme si povšimli tendence zopakovat hlásku, která se ve slově vyskytuje náležitě na jiném místě, např. *plodlouhlý* (*podlouhlý*), *regdiguju* (*rediguju*), *ztpátky* (*zpátky*).

Ze skupin identifikovaných pouze v mluveném jazyce bylo nejčastější přidání slabiky. Jednalo se jak o zdvojení téže slabiky, např. *vzdálenonost* (*vzdálenost*), tak o přidání hláskově zcela jiné slabiky, např. *nedochápu* (*nechápu*). Přidání hláskově odlišné slabiky je často dáno opravou předchozí slabiky a přepis se v těchto případech může různit, srov. *nevyvěděla* (*nevěděla*), *Božena ně* (*Boženě*), *leháka ka . tka* (*lehátka*). Nekonzistence zápisu a značení přereknutí (a slovního fragmentu) je patrná v následujícím příkladu, ve kterém se mluvčí přerekl na začátku první slabiky nového slova a hned se opravil: *pod kolenem dole .. že mě něco <PR chle> lechtá*.

Jak bylo řečeno výše, k přereknutí i překlepům dochází u původem cizích slov. Mimoto se týká i proprií. V mluveném jazyce jsme si povšimli, že mluvčí reflektuje svou nejistotu ohledně správného tvaru slova (př. 19), příp. je opraven druhým mluvčím (př. 20), zatímco v psaných textech jsme obdobné případy nenašli.

(19)

Naděžda\_325: takže .. ať to řeknu dobře . kvard\* kva\* .. to ty ostrovy <PR Kvardepské> ? .. ne jak se jaké jsou ka\* ? +

Soňa\_148: + kap\* . Kapverdské

Naděžda\_325: Kapverdské já to vždycky to .. no takže . @ Natálka měla .. touhu tam jet na dovolenou

(20)

Augustin\_683: on to byl nějakej . gauner svině

Petra\_682: a to říkali včera že korve\* . n\* ko\* jak se to říká ?

Augustin\_683: <PR konvektoval>

Petra\_682: konvertoval k islámu ?

Mezi překlapy tvoří zvláštní skupinu chybně umístěná mezera,<sup>155</sup> která zapříčiňuje buď rozdělení slova, např. *je ště*, nebo spojení dvou slov dohromady, např. *docelaceličký*, příp. dvou slov a interpunkčním znaménkem, *pracuje.To*. Ve všech žánrech elektronické komunikace

---

<sup>155</sup> Vzhledem k tomu, že se jedná o jev potenciálně se vyskytující pouze v psaném jazyce, protože transkripty mluveného jazyka jsou několikrát zkontrolované, uvádíme ho v tabulce 11 samostatně.

převažovalo spojení slov (diskuse: 22×, fóra: 60×, Facebook: 12×) nad rozdělením (diskuse: 10×, fóra: 14×, Facebook: 6×).

### Fragmenty slov

Mezi překlepy i přeroknutími jsme vyčlenili skupinu slov, ve kterých chybí písmeno/hlásky. Ačkoli je odlišnost této skupiny od fragmentů daná především způsobem zápisu v mluveném korpusu Ortofon, je mezi nimi rozdíl i v části slova, ve kterém hláska chybí. U překlepů/přeroknutí se většinou jedná o prostředek slova, méně často o jeho konec, nikdy ne o začátek. Naopak fragmenty jsou nejčastěji pouze začátky slov (soudě podle umístění symbolu \* za vyslovené hlásky). Fragmenty, které nejsou začátkem slova, tj. mají hvězdičku před vyslovenými hláskami, tvoří pouze 2 % případů a většinou se jedná o slova rozdělená v řeči na dvě nebo více částí včetně vložené pauzy nebo výplňkového výrazu, např. *nas\* na\* . \*sekat jim ořechy, no autosalonu .. no .. z auto\* prostě . \*prodejny no.*

Podle frekvenčního seznamu dominuje mezi fragmenty vyslovení jedné hlásky (třemi nejčastějšími fragmenty jsou *s\**, *t\**, *z\**). Ačkoli se první slabika (*na\**) vyskytuje v tomto seznamu na čtvrtém místě, mezi deseti nejčastějšími fragmenty je jediná, další se objevují až od 12. místa. První vyslovený<sup>156</sup> dvojslabičný fragment (*tako\**) se objevuje na 118. místě. Při analýze jsme se nezabývali tím, zda se po fragmentu vyskytuje v segmentu ono slovo, ke kterému by fragment patřil, ani ověřováním spolehlivosti přepisu.

Kontextová analýza ukázala, že ve 21 % dojde po vyslovení fragmentu k pauze v řeči. K určení důvodů pro vznik fragmentu však kontextová analýza příliš nepomáhá, fragment může vzniknout v zásadě kdekoli a na jakémkoli slově. Samozřejmě může být důvodem vyrušení jiným mluvčím (v př. 21 souhlasné přitakávání, v př. 22 skočení do řeči), překryv (pouze 21 % fragmentů bylo vysloveno v překryvu), (sebe)oprava (např. *to je takovej ten . fial\* f\* taková ta fialov\* modrofialová mastička*), rychlá řeč, ve které se může mluvčí několikrát „zaseknout“ (např. *nebo t\* snad se to t\* jako ty traktor\* traktoras\* traktoristi*), cizí nebo obtížně vyslovitelné slovo (např. *eur\* euroformuláře*), řečová porucha apod.

(21)

Emil\_1048: přitom by asi měly bejt nepromokavý ne když si to [na\* . \*impregnuješ] a je to pravá kůže  
Vítězslav\_1047: [hmm]

---

<sup>156</sup> Toto adjektivum přidáváme kvůli fragmentům *něja\**, *nějak\**, které jsou sice ve frekvenčním seznamu výše (54. a 110. místo), ale vysloveny jsou jednoslabičně [ňá], [ňák].

(22)

Václav\_89: potom na svačinu mám [rýžov\*] ..

Michal\_88 : [je to .. je to dobré] ?

Václav\_89: ne je to hnusné navaluje se mně z toho ..

Fragmenty se jeví jako součást idiolektu. V mluveném korpusu Ortofon se vyskytují nejčastěji u mluvčího Břetislav\_1045, který je hlavním mluvčím (tj. jeho řeč tvoří většinu, zde 85 % hovoru) také v nahrávce s nejvyšším počtem fragmentů. V korpusu je obsaženo pět nahrávek s tímto mluvčím, přičemž ve čtyřech z nich je Břetislav hlavním mluvčím (v těchto nahrávkách celkem vystupují tři mluvčí, kromě Břetislava ženy).

### **Shrnutí**

V této studii jsme se zaměřili na výskyt slov, která se odchylojí od své standardní podoby jak svou pozměněnou hláskovou podobou (ty označujeme za překlepy a přeřeknutí), tak svou nekompletností (ty označujeme za fragmenty slov). Jedná se o jevy, na jejichž výskytu se podílí množství faktorů, které nejsme s to v úplnosti podchytit, přesto jsme se pokusili popsat oba jevy v mluvených i v psaných elektronických textech. Následující závěry jsou spíše než exaktně dosaženými výsledky pozorováními, k nimž přispěla kvantitativní analýza.

Mezi překlepy a přeřeknutí jsme identifikovali čtyři styčné skupiny rozdělené podle typu odchylky od standardní podoby slova. V psaném i mluveném jazyce se nejčastěji jednalo o záměnu hlásky/písmene za jiné, které ze slova činí neexistující nebo zcela jiný lexém, dále o chybějící nebo nadbytečné/nadbytečnou písmeno/hlásku a změnu pořadí písmen/hlásek. Výskyt všech identifikovaných typů překlepů lze odůvodnit rychlostí psaní a absencí následného přečtení, nemůžeme však vyloučit ani jiné důvody (např. nejistotu či nevědomost u psaní cizích, dlouhých apod. slov, vliv cizího jazyka nebo nářečí).

Výskyt obou jevů v mluveném jazyce je těžké posoudit, protože jsme řeč neanalyzovali fonetickými metodami (zaměřenými např. na proměny tempa řeči). U přeřeknutí jsme si všímali především záměn mezi hláskami s určitou vlastností (např. délka vokálu, místo artikulace a měkkost konsonantů). Výskyt fragmentů jsme usouvztažnili s výskytem pauzy jako dalšího znamení formulačních obtíží. Při analýze fragmentů jsme dále odhalili mluvčího, pro kterého je nedořečení slov možno vnímat jako výrazný rys idiolektu.

### **3.4.3 Shrnutí kapitoly**

Jako jazykové projevy spontánnosti jsme vybrali opakování slov a výskyt překlepů, přeřeknutí a fragmentů slov. Na rozdíl od druhého rysu lze motivaci pro opakování stejných slov za sebou



postihnout jednoznačně: v mluvené komunikaci se jedná o spolupůsobení časového tlaku se snahou vyjádřit se nebo zachovat si slovo, v psané elektronické komunikaci o nepozornost pisatele ve smyslu absentujícího nebo nedostatečně pozorného přečtení svých vlastních slov. Tyto důvody jsou platné i pro druhý rys, avšak zejména v mluveném jazyce je výskyt přeráznutí a slovních fragmentů ovlivněn dalšími faktory, z nichž jsme v analýze poukázali např. na hláskové složení slova, individuální řečové chování a průběh komunikace.

V obou komunikačních modech jsme našli podobné charakteristiky pro oba rysy, přestože frekvenčně se liší (převažuje vždy výskyt v mluveném jazyce). Ačkoli je četnost rysu často prvním a zpravidla nejzásadnějším ukazatelem odlišností mezi komunikačními módy, při analýze rysů spjatých se spontánností bylo nutné usouvztažnit ji s možnostmi daného komunikačního modu a potenciálními důvody pro výskyt obou jevů. Z toho hlediska se oba vybrané rysy jeví jako vhodně vybrané, neboť jsou skutečně projevem spontánnosti.

### **3.5 Rysy související s naplněním pragmatických potřeb**

V oblasti běžného dorozumívání je komunikační záměr každého mluvčího/pisatele různý, odvislý od komunikační situace, tématu apod. Během jakékoli komunikace může u mluvčího/pisatele vyvstat aktuální potřeba nesouvisející přímo s obsahem textu, ale s jeho recepcí u adresáta, příp. s recepcí samotnou. Jedná se např. o vyzvání adresáta k reakci, nasměrování jeho pozornosti k určitému jazykovému prvku. Takové jazykové chování označujeme za pragmatické.

V této části se budeme zabývat třemi jazykovými rysy, které jsou mluvčími/pisateli užívány k naplnění pragmatických potřeb, a to diskurzními markery, iterací grafémů a jevem označeným jako pronoun non-dropping. Ani jeden rys není využíván k naplnění pragmatických potřeb určitého typu, ale jedná se vždy o širší spektrum. Cílem analýzy tedy bude nejenom tyto potřeby identifikovat, ale také porovnat v mluvené a psané elektronické komunikaci.

#### **3.5.1 Diskurzní markery**

Označení a charakteristika těchto výrazů se v literatuře různí<sup>157</sup> (např. Fraser, 2006), budeme proto vycházet z nejnovějšího výzkumu provedeného na češtině (Čermáková et al., 2019). Diskurzní markery (dále DM) můžeme nalézt v každém jazyce, vzhledem k jejich různorodé funkci však nejsou snadno přeložitelné (nemají jeden přímý ekvivalent, zpravidla se překládají právě podle funkce). Nemají lexikální význam, nebo je jejich původní význam silně oslaben,

---

<sup>157</sup> Tuto problematiku shrnuje Čermáková et al. (2019: 246n.).

stojí mimo tradiční syntaktické struktury, proto je nelze přiřadit k tradičním slovním druhům (pomineme-li možnost zařadit je mezi částice). Za jejich doménu bývá považován mluvený jazyk interakčního charakteru; v něm je vodítkem k jejich identifikaci prozodie (Aijmer, 2013).

DM plní množství různých funkcí, podle nichž bývají kategorizovány (např. Brinton, 1996). S funkcí souvisí i různorodost jejich pojmenování, např. Müllerová (1979, 2011) používá pojem členicí signály a dále je dělí na konstrukční, kontaktní a komentářové. Naopak Čermáková et al. (2019) rozlišují tři skupiny DM podle toho, k jakému prvku komunikačního procesu se vztahují: k adresátovi – interakční DM, ke sdělení – modální DM, k organizaci projevu – konektivní DM. V první a poslední makroskupině je použito ještě poziční kritérium (iniciální a terminální DM). U interakčních DM jsou navíc rozlišeny dvě skupiny podle převažující funkce slova (kontaktní a responzní DM). V MDA jsou DM přítomny v několika jazykových rysech, z nichž jsou pro tuto kapitolu relevantní následující dva: kontaktní slova a výplňkové výrazy. Funkcí prvně zmíněných by mělo být navázání a udržení kontaktu s adresátem, u druhé skupiny slovní vyplnění projevu ve chvíli, kdy mluvčí tápe a hledá správné slovo, kdy se v komunikaci snaží zabrat větší prostor apod. DM jsou charakteristické také svou polyfunkčností (Čermáková et al., 2019), rozdělení do skupin podle převažující funkce bez uvedení rozhodovacích kritérií proto může být zavádějící. Z těchto důvodů jsme zařadili oba jazykové rysy do kapitoly o DM. Při analýze bývají DM zpravidla popisovány po jednotlivých lexikálních jednotkách, viz např. studie Schiffrinové (1987) či Aijmerové (2013). Tímto přístupem se inspirovala také nejnovější studie o DM v mluvené češtině (Čermáková et al., 2019),<sup>158</sup> na kterou v této kapitole navazujeme. Pro analýzu si vybíráme tři výrazy, které jsou obsaženy v seznamu potenciálních DM u Čermákové et al. (2019: 257) a v seznamu kontaktních výrazů nebo výplňkových slov v MDA, a jednu kolokaci, která je také v seznamech pro MDA. Analyzované výrazy také patří mezi tisíc nejfrekventovanějších slov v mluveném i psaném jazyce obecně.<sup>159</sup> Ze slovnědruhového hlediska představují DM otevřenou skupinu, proto jsme pro případové studie vybrali jednoslovné výrazy z různých slovních druhů: substantivum/částici *fakt*, částici *jo*, adverbium *pak*, sloveso *viš*. Na rozdíl od Čermákové et al. (2019) jsme stejným postupem analyzovali ještě kolokaci *že jo*.

Následující případové studie mají jednotnou strukturu. Nejprve jsou shrnuty informace o analyzovaném slově z odborné literatury a dvou výkladových slovníků (*Slovník spisovného*

---

<sup>158</sup> Převzata byly pouze analýza, na jejímž ručním zpracování se autorka této práce přímo podílela. Viz 3.5.1.1.

<sup>159</sup> Měřeno na tisíce nejfrekventovanějších tvarech z výrazně větších korpusů: psaného korpusu SYN2015, a sloučení mluvených korpusů ORAL a Ortofon (Škrabal & Komrsková, 2019).

*jazyka českého*, Havránek et al., 1960–71; *Slovník spisovné češtiny*, Filipec et al., 1994), poté je uvedena frekvence slova v korpusech Ortofon a Koditex-web. K analýze byl ve většině případů použit náhodný vzorek o velikosti 100 výskytů z každého typu zkoumaných dat, tj. 4 vzorky. Pokud se slovo vyskytovalo méně než stokrát, byly použity všechny výskyty. Vlastní analýza je zaměřena na popis funkcí daného slova; kritériem pro rozlišení funkcí byl okolní kontext, pozice slova ve větě/segmentu, prozódie.

Cíle této kapitoly se dají shrnout do následujících bodů:

- ověřit, zda lze o DM uvažovat i v psané elektronické komunikaci,
- doplnit pro češtinu seznam analyzovaných DM,
- aplikovat metodologii Čermákové et al. (2019) na psaných datech.

### 3.5.1.1 Jo

Výraz *jo* je ve studii Čermákové et al. (2019) označen za interakční responzní DM.<sup>160</sup> V MDA je zařazen do seznamu kontaktních výrazů.

Ve výkladových slovnících SSJČ a SSČ je výraz *jo* rozdělen do dvou hesel. Rozdíl mezi slovníky je patrný u významu souhlasu a přisvědčení, který SSJČ označuje za adverbialní a SSČ za částicový (společně s významem zdůraznění). Další významy, např. vyjadřování odporu, povzdechu, překvapení, expresivity, kategorizují oba slovníky jako citoslovce. SSČ u citoslovce *jo* ještě uvádí, že bývá zpravidla reduplikované či emfaticky prodloužené. Slovníky v exemplifikaci uvádějí také kolokace *že jo* (SSJČ) a *ach jo* (SSČ). Kolokaci *že jo* je věnována samostatná případová studie (viz dále).

V této studii se věnujeme pouze samostatně stojícímu výrazu *jo*.<sup>161</sup> Z vyhledávání byly vyňaty kolokace *ach jo*, *no jo*, *ty jo*, *že jo*. Ponechány však byly reduplikované výskyty *jo jo*. Samostatně byly analyzovány výskyty reduplikovaného *jo* zkráceného pouze na konsonanty: *jj*.<sup>162</sup> Tento jev je popisován v kontextu elektronické komunikace, v níž hraje roli rychlost odeslání příspěvku, např. na chatu.

---

<sup>160</sup> Souhrnná míra výskytu u výrazu *jo* je 0,39. Podobné hodnoty dosahují další interakční responzní výrazy *ok* (0,41), *ano* (0,36).

<sup>161</sup> Dotaz použitý v psaném korpusu: [lemma!="a?ch|no|ty|že"] [lemma="jo"] within <s/>.

Dotaz použitý v mluveném korpusu: [lemma!="ach|no|ty|že"] [lemma="jo"] within <sp/>.

<sup>162</sup> Vyhledány pomocí dotazu: [word="(?)jj+"].

	frekvence
Ortofon	10 471 (8 468,2 ipm)
diskuse	41 (170,23 ipm)
fóra	140 (577,26 ipm)
Facebook	37 (152,4 ipm)

Tab. 15. Přehled frekvence výrazu *jo* v použitých datech

## Analýza

### 1. Funkce výrazu *jo*

Primárním významem výrazu *jo* je souhlas, který považujeme za specifický typ responzního chování. Vzhledem k povaze analyzovaných dat nemůžeme vždy rozlišovat mezi silným a slabým souhlasem (přítakáváním) jako Čermáková et al. (2019: 285n.). Všechny výskyty v souhlasné funkci (1.1) chápeme jako DM, mimoto rozlišujeme další tři funkce DM. O zdůrazňovací funkci (1.2) se zmiňuje SSČ, ale ve studii od Čermákové et al. (2019) se na rozdíl od zbylých dvou skupin – výzva k reakci (1.3), signalizace změny tématu (1.4) – neobjevuje.

#### 1.1 Souhlas

Vyjádření souhlasu implikuje přítomnost otázky či jiné výzvy v předchozím kontextu. Vzhledem k podobě korpusu Koditex jsme to nemohli ověřit ve všech výskytech, v nichž jsme o dané funkci uvažovali, řídili jsme se proto kontextem výrazu. V případech, v nichž byl výraz *jo* modifikován zájmenem *to* nebo vytýkací částicí, např. *asi, možná, určitě*, jsme označili jeho funkci za souhlasnou (př. 23). Řadili jsme sem i výskyty, ve kterých se objevila kolokace *jako jo* ve smyslu váhavého souhlasu.

(23)

- **jo jo jo jo . jo to jo** jasný .. **jo jo jo** (Ortofon)
- Jestli to tak opravdu je , nemůžu posoudit , v jednotlivých případech možná **jo** , všeobecně pochybuji , ale nevím . (diskuse)
- @ belladonna píše : **jo** , nemá , to máš pravdu . (fóra)
- no , to je pokrok , to **jo** teda (Facebook)

Do této skupině byly dále zahrnuty výskyty, ve kterých *jo* zastupovalo kladnou variantu věty nebo kladnou možnost rozhodnutí (př. 24). Celkově se jednalo o nejpočetnější skupiny v rámci všech žánrů.

(24)

- a to já už nedělám , teda dnes **jo** , ale včera a předtím ne (fóra)

## 1.2 Zdůraznění

Tuto funkci jsme přisoudili pouze jednomu výskytu (př. 25). Výraz *jo* slouží ke zdůraznění či k vyjádření vyšší míry vlastnosti vyjadřované následujícím výrazem.

(23)

- když chcete být **jo** sprostý (diskuse)

## 1.3 Výzva k reakci

Výraz *jo* ve funkci výzvy směřuje od pisatele/mluvčího k adresátovi, často je také doplněn otazníkem nebo v mluveném jazyce i stoupavou intonací<sup>163</sup> či protažením vokálu. V mluveném jazyce není vždy jasné, zda se mluvčí obrací k adresátovi s cílem získat od něj reakci, nebo zda jde pouze o návyk bez očekávání reakce. Mluvčí tímto výrazem zároveň člení segment. Pro tuto skupinu je charakteristická pozice na konci věty/segmentu, příp. užití v samostatné replice (př. 26).

(26)

- tam je to je prostě všechno vydlážděný a to a teče . teče to dolů **jo** .. (Ortofon)

- Já za tebe musím platit daně , **jo** ? (diskuse)

- Buď od té dobroty a nesud' mě podle sebe , **jo** ? (fóra)

## 1.4 Signalizace změny tématu

Do této skupiny jsme zařadili případy, v nichž se výraz *jo* objevoval nejčastěji na začátku věty/segmentu a uvozoval nové téma či další informaci, kterou pisatel/mluvčí zapomněl a nyní ji chce dodat (př. 27). Souhlasná funkce je stále přítomna, výraz *jo* slouží kromě zavedení nového tématu také k navázání a propojení vět/segmentů.

(27)

- vono to bylo uděláno režijně taky zajímavě že to bylo vlastně zpětně . @ **jo** ta ta předehra měla krásnej doprovod (Ortofon)

- Nebo nám spíš celou dobu přehřátí sluníčkáři v čele s Merkel kecají , jak si myslí drtivá část rozumné populace , **jo** a Hnízdo , kolik už máš u sebe ubytovaných těch chudáčků ? (diskuse)

---

<sup>163</sup> Stoupavá intonace vede k užití otazníku.

- Přítel jí to zatajil a řekl jí to , až vypila dva panáky , nee ? **Jo** a k dr . ji odvezl v komatu  
(fóra)

- Jinak u mne celkem OK , **jo** ještě srovnat ty kšandy , kšandy my baby nosíme víc na bok  
(Facebook)

### 1.5 Nejasné případy

Žádnou funkci jsme nepřisoudili těm výskytům, u nichž jsme ani z kontextu nemohli jednoznačně rozhodnout (př. 28). Tato situace se týká pouze elektronické komunikace a souvisí s tvorbou korpusu Koditex-web. Omezení, která vyplývají ze sdružování příspěvků pouze od jednoho autora, byla vzata v úvahu již před analýzou výrazu *jo*.

(28)

Evelyn , **jo** , tak zhubnout 30 kg by se mi hodilo  
(fóra)

### 1.6. Reduplikace výrazu *jo*

V analyzovaných datech se objevila reduplikace pouze ve variantě *jj*, ačkoli je možné užít neomezený počet grafému *j*. Reduplikace se v elektronické komunikaci vyskytovala spíše marginálně: ve fórech 28×, v diskusích 9×, na Facebooku 7×. Výskyty byly rozloženy mezi větší počet pisatelů: 20 ve fórech, 5 v diskusích, 5 na Facebooku.

Všechny výskyty považujeme za vyjádření souhlasu, a to i v případech, v nichž kvůli chybějící reakci komunikačního partnera nevíme, k čemu se *jj* vztahuje (př. 29). Vycházíme ze situace v mluveném jazyce, v němž se reduplikace *jo jo* objevila nejčastěji při vyjádření souhlasu a přitakání, které můžeme chápat jako bezmyšlenkovité vyjádření zpětné vazby signalizující pozornost.<sup>164</sup> Ve většině případů bylo rozhodnutí o souhlasné funkci podpořeno kontextem, v němž byl souhlas opakován jinými slovy (př. 30).

(29)

- nejlépe o Sýrii , o tu nyní jde , nejvíce , jak se měli , jaké měli práva a zákony a jak tam lidé opravdu žili </s><s> :-> </s> <s> **jj** můžu Vám říci , žena měla prim , ale to nechcete vědět jinak byste se vzdělal  
(Facebook)

(30)

- **Jj** , přesně tak .  
(fóra)

---

<sup>164</sup> V literatuře označované také jako backchannelling (např. Peters & Wong, 2015).

*Jj* se ve většině případů vyskytovalo na začátku věty (pravděpodobně celého příspěvku). Poukazuje to na jeho využití v interakci jako prostředku, jímž je vyjádřen souhlas s předchozím příspěvkem a jímž je zároveň umožněno plynule navázat vlastním textem.

### Shrnutí analýzy

Všechny výše popsané funkce výrazu *jo* chápeme jako funkce DM (srov. Čermáková et al., 2019). Při analýze se jako nejčastější ukázala ve všech žánrech funkce souhlasu (tab. 16). Tato funkce byla přiznána i samostatně analyzovaným výskytům reduplikovaného *jo jo* v psané elektronické komunikaci. Další výrazná diskurzněmarkerová funkce – výzva adresáta k reakci – se objevila nejčastěji v mluveném jazyce. Naopak v elektronické komunikaci, zejm. v příspěvcích z diskusí a fór, byl výraz *jo* použit jako signál změny tématu. Každá z těchto dvou funkcí je spojena s jinou pozicí výrazu: výzva s finální pozicí, změna tématu s iniciální pozicí.

		Ortofon	Koditex-web		
			diskuse	fóra	Facebook
diskurzni marker	souhlas	50 %	54 %	64 %	89 %
	zdůraznění	0 %	2 %	0 %	0 %
	výzva k reakci	38 %	6 %	11 %	0 %
	signalizace změny tématu	12 %	17 %	20 %	8 %
nejasné		0 %	21 %	5 %	3 %

Tab. 16. Souhrn nalezených funkcí u výrazu *jo*

Analýza výrazu *jo* byla v elektronické komunikaci limitována absencí celého diskusního vlákna, tj. příspěvků od všech pisatelů. Výhoda tohoto přístupu – důraz na idiolekt – byla využita při analýze *jj*. Tato reduplikovaná varianta je ve zkoumaných datech rovnoměrně rozložena mezi různé pisatele, závěry týkající se jejího užívání jsou proto obecně platné.

#### 3.5.1.2 Víš

Tento výraz je v Čermákové et al. (2019: 257) označen za potenciální DM, ale samostatně zde analyzován nebyl. V souladu s metodou<sup>165</sup> použitou v Čermákové et al. (2019) bychom ho zařadili do skupiny interakčních kontaktních terminálních DM vedle výrazů *chápeš*, *rozumíš*, *vid'*. V MDA byl tento výraz součástí seznamu kontaktních výrazů jako samostatný tvar a zároveň v kolokaci *to víš*. V naší analýze jsme pracovali pouze se tvarem *víš*.<sup>166</sup>

<sup>165</sup> Souhrnná míra výskytu, na jejímž základě autorky určovaly, zda výraz tenduje k výskytu na začátku, nebo na konci segmentu, je u výrazu *víš* 0,51. Uvedené interakční kontaktní terminální výrazy mají souhrnnou míru výskytu výrazně odlišnou: *rozumíš* (0,06), *vid'* (-1,95), *chápeš* (nelze vypočítat, v korpusu Ortofon se vyskytuje pouze na začátku segmentů).

<sup>166</sup> Vyhledaným dotazem [word="( ?i) víš" ].

Výraz se neobjevuje jako samostatné heslo ve výkladových slovnících, je součástí hesla *vědět*. V SSSJČ je ke tvarům *viš*, *víte* uvedena poznámka o jejich expresivním využití ke zdůraznění a vyjádření naléhavosti toho, co je vyslovováno.

*Viš* je druhým nejčastějším tvarem (po tvaru *nevím*) lemmatu *vědět* v mluveném jazyce, tvoří 32 % všech jeho výskytů (*nevím* zahrnuje 45 %). V elektronické komunikaci jsou častější tvary *nevím*, *vím*, *vi*, *víte*, *nevi*. Nižší frekvence tvarů 2. osoby je zapříčiněna častějším uváděním vlastního názoru pisatele nebo jiné osoby oproti explicitnímu obracení se na komunikačního partnera, resp. komunikační partnery.

	frekvence
Ortofon	2 717 (2 197,32 ipm)
Diskuse	21 (87,19 ipm)
Fóra	47 (193,79 ipm)
Facebook	13 (53,54 ipm)

Tab. 17. Frekvence výrazu *viš* v použitých datech

## Analýza

### 1. Původní slovesný význam

Do této skupiny byly řazeny výskyty, v nichž se výraz *viš* objevil s valenčním doplněním v akuzativu (příp. s jeho nahrazením vedlejší větou) nebo ve spojení s předložkou *o* + LOK (př. 31). Tato skupina byla nejčastější ve všech třech žánrech elektronické komunikace (tab. 17).

(31)

- jestlipak **viš** u koho ten Honza spal jo ? (Ortofon)
- drtivou většinu největších swináren - a to doufám taky dobře **viš** - si dělají bílí vohnouti navzájem mezi sebou (diskuse)
- Njn , kdyby jsi přinesla prachy , alkohol nebo cigarety , pak by jsi více potěšilaale to mě můžou **viš** , co ... nas.at ! (fóra)
- **Viš** o tom , že totoje auto nasmradí za minutu asi tolik co pětset kuřáků (Facebook)

### 2. Oslabení slovesného významu

O oslabení původního slovesného významu uvažujeme pouze v případech kolokací s výrazy *co* (*viš co*), *jak* (*viš jak*) a *to* (*to viš*), pokud není z kontextu možné nahradit tyto výrazy konkrétním vyjádřením. K pochopení je často nutná znalost delšího předchozího kontextu nebo



mimotextových skutečností. Kolokace také mohou referovat k obecně sdíleným názorům, které není třeba rozvádět. Z těchto důvodů neřadíme tyto případy k DM.

(32)

- ale tak to ne to já se projdu ale .. **víš** jak – (Ortofon)
- A to jak ve vyznání , tak ve způsobu a podstatě To **víš** že křesťané válčí a dobývají , to **víš** , že ano . (diskuse)
- na Fb , se ptá tady a zase nic nevím **víš** co , to dělával chlap (fóra)
- protože sem se živil rukama , neměl sem hrůzu z lopaty a od té se nevyhazovalo - nebylo kam - To **víš** , socialismus osumdesátých let (Facebook)

### 3. Diskurzni marker

Základním kritériem pro zařazení do této skupiny byla absence valenčního doplnění. Dalším kritériem byl test, zda lze výraz vynechat, aniž by se změnila informační struktura textu. Výraz *víš* klasifikovaný jako DM slouží především k navození či zesílení kontaktu s adresátem (i ve spojení s oslovením, viz diskuse v př. 33) nebo jeho pozornosti (lze dokonce říci, že mluvčí se dovolává adresátovy znalosti). Kontaktní funkce se v některých případech z mluveného jazyka kombinuje s členicí funkcí; výraz *víš* stojí zpravidla na konci věty (př. 34) nebo před výrazem, který chce mluvčí zdůraznit (př. 35).

(33)

- **Víš** Polický , já přebírám pouze data . (diskuse)
- možná první půlrok byl lepší . . **víš** nedovedu si představit upřímně druhý rozvod . . (fóra)
- Jak se před dětmi chováme my zde v Česku a jiných státech , to v Norsku vůbec nepřipadá v úvahu . Tak že mi říkal . **Víš** . Kdo ví , jak to tam bylo . (Facebook)

(34)

prošel jsem si vodopády **víš** a @ pak jsem šel nahoru do Rešova (Ortofon)

(35)

já myslel že je to nějaká . kostka a že tam budou **víš** vokoło čtyři vobchody (Ortofon)

### Shrnutí analýzy

Analýza ukázala na posun v užívání tvaru *víš* od původního slovesa k DM. Kritériem pro odlišení byla přítomnost, či ztráta valenčních doplnění. Samostatně byly vyčleněny kolokace

tvaru *víš* s valenčním doplněním odkazujícím neurčitě ke skutečnostem v širším kontextu. Právě toto užívání ilustruje podle našeho názoru proces postupného přechodu k DM.

*Víš* jako DM bylo nejčastěji zastoupeno v mluveném jazyce (39 %). V elektronické komunikaci jako celku se jednalo o druhou nejčastější skupinu (33 %), v jednotlivých žánrech se však podíly lišily (tab. 18). Rozdíl mezi mluveným a psaným jazykem je patrný při procentním srovnání všech tří skupin a pozici výrazu ve větě. V mluveném jazyce je zastoupení skupin vyrovnanější, zatímco v psaném výrazně převažuje původní slovesný význam. V mluveném jazyce se výraz častěji objevoval na konci vět, v elektronické komunikaci se vyskytoval nejčastěji na začátku věty.

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
původní slovesný význam	32 %	52 %	66 %	54 %
oslabení původního významu	29 %	9 %	15 %	31 %
diskurzivní marker	39 %	29 %	9 %	15 %

Tab. 18. Srovnání analýzy u výrazu *víš*

### 3.5.1.3 Fakt

Tento výraz je u Čermákové et al. (2019: 257) označen také za potenciální DM.<sup>167</sup> V MDA je do seznamu kontaktovalých výrazů zařazen výraz s následujícím otazníkem, tj. *fakt?* Pro naši analýzu bereme v úvahu všechny výskyty tvaru<sup>168</sup> *fakt*, protože frekvence *fakt?* je v datech nízká.

Při vyhledávání tedy nerozlišujeme mezi významy uvedenými v SSJČ a SSČ, v nichž je jako první význam uvedeno substantivum označující skutečnost. Další význam, který nese příznak hovorový, je v SSJČ považován za příslovce, v SSČ za částici modální s odstínem dotvrzování a přesvědčení.

Tabulka 19 ukazuje, že výraz *fakt* je nejčastěji užíván v mluveném jazyce a v diskusních fórech. Z korpusu Ortofon se dovídáme, že *fakt* užívají častěji ženy (56 %) oproti mužům, lidé s vysokoškolským vzděláním (66 %) oproti lidem s nižším vzděláním, lidé mladší 35 let (54 %) oproti starším 35 let.

<sup>167</sup> Souhrnná míra výskytu u výrazu *fakt* je 1,51.

<sup>168</sup> Vyhledáno pomocí dotazu [word="(?) fakt"].

	frekvence
Ortofon	1 908 (1 543,06 ipm)
Diskuse	85 (352,91 ipm)
Fóra	306 (1 261,72 ipm)
Facebook	118 (486,02 ipm)

Tab. 19. Přehled frekvence výrazu *fakt* v použitých datech

## Analýza

### 1. Substantivum

Při vyhledávání jsme nepoužívali morfologické značkování kvůli chybovosti.<sup>169</sup> Zařazeny sem byly případy, ve kterých se výraz *fakt* dal nahradit synonymem skutečnost, pravda (či pravdivá skutečnost). Substantivum může být rozvito přívlastkem, např. adjektivy *nezpochybnitelný*, *pozoruhodný*, *zřejmý* nebo zájmenem *tento*. Často kolokuje se slovesy vyjadřujícími myšlenkový proces, např. *brát jako*, *opomíjet*, *svědčit o*, *zpochybnit*. Výraz se také objevuje v kolokaci (*ale*) *je fakt*, *že* používané pro uvození jiného názoru.

### 2. Diskurzni marker

U výskytů považovaných za DM jsme se pokusili o detailnější klasifikaci jeho funkcí. Pokud bylo z kontextu patrné, že se platnost výrazu vztahuje pouze k jinému výrazu v rámci segmentu/věty, označili jsme funkci výrazu za dotvrzovací (2.1). V případě, že se za výrazem vyskytoval otazník a tato otázka směřovala k zjištění nějaké skutečnosti, přisoudili jsme těmto výskytům funkci kontaktnou (2.2). Další skupinu tvoří případy, v nichž výraz *fakt* vyjadřuje souhlas (2.3). Poslední skupina, označená jako výplněk (2.4), je tvořena pouze výskytů z mluveného jazyka, protože vznikla na základě poslechu. Výraz *fakt* je použit k získání času pro produkci.

#### 2.1 Dotvrzování

Výraz *fakt* ve funkci dotvrzování či přesvědčování o platnosti může být nahrazen synonymy *opravdu*, *skutečně*, *vážně*. Dotvrzování zároveň může vyjadřovat vyšší míru intenzity.

<sup>169</sup> Nejvíce chyb bylo přítomno v datech z internetových diskusí. Z celkových 85 výskytů jsme manuálně označili 24 jako substantivum, přičemž z tohoto počtu bylo pouze 12 určeno automaticky správně.

(36)

- ty byli jako **fakt** opilý a jako oplzlý (Ortofon)
- To jste mne potěšil ( ale **fakt** - nedělám si legraci (diskuse)
- Jako **fakt** to bylo pěkný , já ty lístky dostala ke 40 inám . (fóra)
- vy **fakt** nekoukáte na noviny ? (Facebook)

*Fakt* v této funkci nemá stabilní pozici v segmentu/větě. Může stát před výrazem, k němuž se vztahuje bezprostředně (př. z Facebooku a Ortofonu) nebo v delším kontextu (př. z fór). Může také následovat za výrazem, k němuž se vztahuje, a tím ho zdůraznit (př. z diskusí).

## 2.2 Kontaktovost

Zařazení většiny výskytů výrazu *fakt* s následujícím otazníkem do této skupiny nebylo provedeno pouze na základě tohoto souvýskytu, ale vyplynulo z otázkové platnosti daného výrazu. Jím se mluvčí/pisatel obrací k adresátovi s žádostí, aby potvrdil, či vyvrátil platnost skutečnosti zpravidla zmíněné v předchozím kontextu. Otázka *fakt?* vyjadřuje zároveň překvapení mluvčího/pisatele.

(37)

- s halapartnou nebo tam stál ? . **fakt** ? (Ortofon)
- pokud člověk řídí auto jen on sám , tak el . sedadlo je mu úplně k ničemu . To jako **fakt** ? (fóra)
- my sdílíme nějaké extra lepší hodnoty než oni ? **Fakt** ? Wow ! (Facebook)

Výraz v této funkci tvořil ve většině nalezených dokladů samostatnou otázku (př. z Facebooku a Ortofonu). Pokud se otázka skládala z více výrazů, stál výraz *fakt* jako poslední (př. z fór).

## 2.3 Souhlas

Vyjádření souhlasu prostřednictvím výrazu *fakt* jsme identifikovali na základě předchozího kontextu. V př. 38 se jednalo o kontext předchozí věty v rámci příspěvku jednoho pisatele. V př. 39 byl výraz použit po otázce vyjadřující překvapení (srov. předchozí skupinu).

(38)

- NORMÁLNÍ člověk ví , že chce - li někdo kuřecí , musí se ZABÍT ( ó díjo :-) kuře . **Fakt** , šunka zatím neroste na stromech (diskuse)

(39)

- Rastislav\_715: jako přímo originál baterky .. ty tam byly čtyři roky
- David\_903: jo ?

Rastislav\_715: **fakt** .. ty už další žádný baterky co jsme koupili už tam tak dlouho nevydržely  
(Ortofon)

Výskyty vyjadřující souhlas se v analyzovaných vzorcích nacházely vždy na začátku segmentu/věty.

## 2.4 Výplněk

Tato funkce byla identifikována v případech, ve kterých nebylo výrazu *fakt* možné přisoudit žádnou z předchozích funkcí. Výraz je použit k získání času pro (správnou) formulaci sdělení, v jeho bezprostředním okolí se mohou vyskytovat jiné výrazy v platnosti DM, pauzy, opakování týchž výrazů (př. 40). Tuto skupinu jsme vyčlenili pouze v dokladech z mluveného jazyka, což neznamená, že by se nemohla objevit i v psaném jazyce. Její nepřítomnost v analyzovaných datech může být ovlivněno velikostí vzorku.

(40)

oni prostě . loktem do toho .. do okýnka .. pak tam něco poštelovali a odjeli si .. ale . to to byli **fakt** .. to byli Romové bylo jim asi fakt jako . čtrnáct let jako ..  
(Ortofon)

## Shrnutí analýzy

Analýza ukázala, že výraz *fakt* je v mluveném i v psaném elektronickém jazyce výrazně častěji užíván jako DM s různými funkcemi než jako substantivum (tab. 20). Mezi funkcemi DM dominuje dotvrzování, o němž se zmiňují i výkladové slovníky a které zároveň zvyšuje míru či intenzitu vyjádřenou zpravidla následujícím výrazem. Další tři funkce DM – kontakto-  
vost, souhlasná a výplňková – byly ve vzorcích málo frekventované, nejvíce výskytů bylo nalezeno v mluveném jazyce. Pozičně se funkce DM liší: dotvrzovací *fakt* bylo nejčastější uvnitř věty před výrazem, k němuž se vztahovalo, kontakto-  
vost se vyskytovalo na konci otázky nebo tvoří samostatnou otázku, souhlasné bylo na začátku segmentu/věty. Výplňkové *fakt* se nacházelo uvnitř větné struktury v místě, kde došlo k přerušení, po němž následovalo zopakování části věty nebo reformulace.

		Ortofon	Koditex-web		
			diskuse	fóra	Facebook
substantivum		2 %	28 %	13 %	24 %
diskurzí- marker	dotvrzování	86 %	71 %	85 %	75 %
	kontakto- vost	4 %	0 %	2 %	1 %
	souhlas	5 %	1 %	0 %	0 %
	výplněk	3 %	0 %	0 %	0 %

Tab. 20. Souhrn nalezených funkcí výrazu *fakt*

Pokud bychom chtěli výraz zařadit do jedné ze skupin DM navrhovaných Čermákovou et al. (2019), volili bychom skupinu konektivních DM, v níž se vyskytují výrazy *ale, jako, prostě, protože, teď, vyloženě*. DM v této skupině se vztahují k organizaci, strukturaci projevu a napomáhají formální textové soudržnosti. Výraz *fakt* ve své nejčastější funkci – dotvrzování – je zároveň strukturačním prostředkem, zejména v postavení před výrazem nebo částí věty, která přináší nové informace.<sup>170</sup>

### 3.5.1.4 Třeba

Výraz *třeba* je stejně jako předchozí *fakt* označen v Čermákové et al. (2019) za potenciální DM<sup>171</sup> a je zahrnut do seznamu výplňkových slov v MDA. Ve výkladových slovnících SSJČ a SSČ je rozlišováno příslovce ve významu nutnosti, nezbytnosti, potřeby a částice s významem nerozhodnosti, neurčitosti a libovolnosti.

Výraz je rovnoměrně zastoupen ve všech sociolingvistických charakteristikách, které byly kritériem pro sestavení korpusu Ortofon, tj. v pohlaví, věku, nejvyšším dosaženém vzdělání a nářeční oblasti pobytu mluvčího do 15 let.

	frekvence
Ortofon	3 180 (2 571,76 ipm)
diskuse	335 (1 390,88 ipm)
fóra	420 (1 731,77 ipm)
Facebook	252 (1 037,93 ipm)

Tab. 21. Přehled frekvence výrazu *třeba* v použitých datech

## Analýza

### 1. Příslovce

Jako příslovce byly na základě slovníkových údajů označeny výskyty, ve kterých je výraz *třeba* součástí verbonominálního predikátu *je třeba*. Ve všech analyzovaných vzorcích se jednalo o nefrekventovanou skupinu.

### 2. Diskurzivní marker

V rámci této skupiny jsme se snažili o detailnější kategorizaci na základě kontextu. *Třeba* se objevuje jako uvozovací výraz při uvádění konkretizaci sdělení pomocí příkladu (2.1), dále

<sup>170</sup> Vzhledem k absenci širšího kontextu u vícesměrné elektronické komunikace zde záměrně neuvádíme termín réma.

<sup>171</sup> Souhrnná míra výskytu u výrazu *třeba* je 1,10.

slouží k vyjádření nejistoty (2.2) a ke zvýšení koherence řeči v případě, že je ohrožena plynulost projevu (2.3).

## 2.1 Uvozování

Tato skupina se ve všech analyzovaných žánrech ukázala jako nejfrekventovanější. Výrazem *třeba* je uvozena skutečnost, která konkretizuje nebo specifikuje předchozí kontext. Výraz s touto funkcí nemá pevnou pozici, může se vyskytovat jak před příkladem, tak po něm (př. 41). Pokud mluvčí/pisatel zdůrazňuje zájmemem svou osobu, nestojí výraz *třeba* před zájmemem *já*, ale po něm.

(41)

- to je tak s\* . dělaj si to tak já nevím **třeba** sto padesát metrů i s garáží **třeba** ..

(Ortofon)

- Já **třeba** rozumím tomu

(diskuse)

- Když dostane dárky tři ( plus další tři od babiček **třeba** ) , tak už to snad na chlubení stačí , ne ?

(fóra)

- Já jsem **třeba** liberál , takže můj pohled je jasný .

(Facebook)

## 2.2 Nejistota

Výraz *třeba* je dále používán k vyjádření nejistoty. Kritériem pro odlišení od předchozí skupiny je náhrada synonymem *možná*. Tento význam uvádí i SSJČ jako jeden z podtypů částice *třeba*. Výraz s touto funkcí je umístován před část sdělení, k níž se vztahuje – zpravidla na začátek věty.

(42)

- najde se tam prostě sto devadesát jako . **třeba** jo .. hloubka ale jako nic

(Ortofon)

- Můžete to navrhnout , **třeba** to projde .

(diskuse)

- Fuj , jestli je pravda , že ji admini takhle zamázli ... fakt svoboda projevuAle **třeba** se to vše vysvětlí a znovu začne fungovat ... bez omezování

(fóra)

- Tak vzbudte Havla , **třeba** nám poradí , když nám LIDUMILSKY zrušil VOJNU

(Facebook)

## 2.3. Výplněk

Tato skupina byla identifikována pouze v mluveném jazyce. Zařazeny sem byly výskyty, které se objevily v místě, kdy došlo k porušení plynulé řeči např. z důvodu reformulace nebo zastavení (př. 43). Dále sem byly přidány výskyty, ve kterých je sice funkce výrazu odhadnutelná z následujícího delšího kontextu (první výskyt výrazu v př. 44), ale další signály

řečové dysfluence (pauzy, hezitace, jiná slova s výplňkovou funkcí) ukazují, že je výraz *třeba* použit k získání času pro formulaci myšlenky.

(43)

to zase třeba oblečení to **třeba** .. neříkám to ne ale jako .. v těch potravinách voni se dost snaží

(Ortofon)

(44)

věci co si myslíš že jsou zakořeněný .. jako naše jako **třeba** že .. @@ .. že . **třeba** příklad že .. @@ člověk je je . dělanej z dvou půlek že jo ..

(Ortofon)

### Shrnutí analýzy

Analýza ukázala rozdíl mezi mluveným a psaným komunikačním modem spočívající v častějším užívání výrazu *třeba* v původním příslovečném významu v psaném jazyce (tab. 22). Ve všech analyzovaných žánrech se výraz nejčastěji objevoval jako uvození následujícího příkladu. S ohledem na to bychom ho mohli zařadit do skupiny konektivních DM definovaných Čermákovou et al. (2019: 260) jako výrazy, které se podílejí na organizaci projevu. Výraz *třeba* může ještě sloužit k vyjádření nejistoty (nejčastěji se tak dělo v příspěvcích z diskusních fór) a v mluveném jazyce jako výplňkový výraz.

		Ortofon	Koditex-web		
			diskuse	fóra	Facebook
příslovce		5 %	21 %	11 %	22 %
diskurzní marker	uvození	89 %	73 %	75 %	66 %
	nejistota	3 %	6 %	14 %	8 %
	výplněk	3 %	0 %	0 %	0 %

Tab. 22. Souhrn funkcí nalezených u výrazu *třeba*

#### 3.5.1.5 Že jo

Zařazení kolokace *že jo*<sup>172</sup> mezi výrazy analyzované v této kapitole bylo motivováno její přítomností v seznamu kontaktních výrazů v MDA. Kolokace by se na základě svých funkcí, popsaných na materiálu mluveného jazyka (Müllerová, 2007; Komrsková, 2017), dala považovat za DM. Ve výkladových slovnících SSJČ a SSČ se objevuje v exemplifikaci hesla *jo* u významu souhlasu, přitakání.

<sup>172</sup> Jsme si vědomi skutečnosti, že označení kolokace může být v tomto případě sporné, neboť někteří uživatelé píšou *že jo* jako jedno slovo (v současném největším psaném korpusu SYN v7 se objevuje 260×, tj. 0,05 ipm). V této studii vycházíme ze zápisu v mluveném korpusu Ortofon.



Kolokace *že jo* je někdy psána jako jedno slovo, proto jsme do vyhledávání zahrnuli i tuto variantu.<sup>173</sup> Ačkoli zde píšeme o kolokaci, při vyhledávání samém ji nemůžeme odlišit od pouhého souvýskytu<sup>174</sup> spojky a částice. Tabulka 23 ukazuje jak celkovou frekvenci *že jo*, tak frekvenci kolokace po odstranění souvýskytů. V mluveném jazyce nebyly kvůli vysoké frekvenci posuzovány všechny výskyty *že jo*. Na základě výsledků anotace náhodného vzorku o 100 konkordančních řádků<sup>175</sup> a poznatků z Komrskové (2017) jsme vyloučili ty výskyty, u nichž se v bezprostředním levém kontextu vyskytovala lemmata *doufat, myslet, obávat (se), představit (si), tvrdit, vědět, věřit, zdát* a lemma *možná*. Nejčastěji se jednalo o spojení *myslím že jo* a *to víš že jo*.

	všechny výskyty	kolokace
Ortofon	2 607 (2 108,36 ipm)	2 407 (1 946,61 ipm)
diskuse	12 (49,82 ipm)	8 (33,22 ipm)
fóra	20 (82,47 ipm)	16 (82,47 ipm)
Facebook	6 (24,71 ipm)	4 (16,48 ipm)

Tab. 23. Frekvence kolokace *že jo* v použitých datech

## Analýza

### 1. Výzva k reakci

Tato skupina byla identifikována v podobných případech jako u výrazu *jo* (s. 93). Výzvodový charakter je zejména v psaném jazyce posílen připojením otazníku či doplňující otázkou (př. 45).

(45)

- Je to pro tebe hodně blbý , že k tématu nic nemáš , **že jo** ? Sžirá tě to ? (diskuse)
- Si děláš srandu , **že jo** ? (fóra)
- co by kohokoliv zajímalo nějaké de iure natož de facto když se prostě dá fantazírovat celý den a noc , **že jo** ? (Facebook)
- von .. a ten je v Plzni ne Botanicus **že jo** že je v Plzni . tam v tej jednej ulici (Ortofon)

K odlišení této a následující skupiny napomohla v mluveném jazyce intonace (převažuje stoupavá) a případná reakce komunikačního partnera. V př. 46 není na první pohled možno

<sup>173</sup> Dotaz použitý v psaném korpusu: ([lemma="že"] [lemma="jo"]) | [lemma="žejo"] within <s/>. Dotaz použitý v mluveném korpusu: [lemma="že"] [lemma="jo"] within <sp/>.

<sup>174</sup> Za souvýskyt považujeme případy, v nichž vedle sebe stojící výrazy nefungují jako lexikální jednotka, ale každý zachovává vlastní význam.

<sup>175</sup> Souvýskyt se objevil 15×.

určit funkci kolokace, nicméně po poslechnutí je patrný jednoznačný výzvolový charakter: kolokace *že jo* je vyřčena se stoupavou intonací a těsně po ní je slyšet pousmání komunikačního partnera, kterým reaguje na uvedený segment.

(46)

ale já úplně čumím jak voni jsou vnímavý ty děti **že jo** .. to . jí něco řekneš a ona okamžitě ví . přesně  
(Ortofon)

## 2. Sumarizace

Ve druhé skupině se vyskytují případy, ve kterých kolokace následuje po výčtu faktů či informací, které mohou být důležité pro pochopení vyprávěného, nebo po shrnutí, příp. generalizaci (př. 47).

(47)

- udělal cestu spodek z toho betonu víš a to .. to byly stotunový náklady .. **že jo** .. no .. a tohle tam vozil .. **že jo** .. a .. prostě .. ta cesta .. byla pro koně ta byla úzká  
(Ortofon)

- jedno to ovoce jsem včera darovala tetě a mám ještě skoro celé , snad to nevyhodím , **že jo**  
(fóra)

- To je úplně stejně blbý argument jakože ve velké británii kradou a flákaj se češi . A my víme , že to jsou ano sic češi , ale naši milí nepřizpůsobiví , **že jo** .  
(Facebook)

Ve většině výskytů se kolokace objevuje na konci věty či klauze, což souvisí s označením funkce – sumarizace. Stejně jako v předchozí analýze mluveného jazyka (Komrsková, 2017) se v rámci této skupiny vyskytují případy, ve kterých je kolokaci možné považovat za uvozovací prvek při opakování týchž slov (př. 48) nebo pro propojení<sup>176</sup> částí segmentu (př. 49). Vložním kolokace *že jo* před klíčové slovo v př. 49 zároveň mluvčí docílil zdůraznění tohoto slova.

(48)

- jenže víš co tam jde vo to že musí mít nějaký zase **že jo** musí mít .  
(Ortofon)

(49)

- jak se říká **že jo** panoramata jo takhle hlavně ty panoramata jsem chtěl vidět ..  
(Ortofon)

---

<sup>176</sup> Příklad je vyřčen bez zaváhání a jiných zvukových signálů typických pro situace, ve kterých si mluvčí není schopen vybavit vhodný výraz.

## Shrnutí analýzy

V analýze jsme se zaměřili na kolokaci *že jo*, která v mluveném i psaném elektronickém jazyce převažovala nad souvýskytem *že jo*, tj. nad případy, ve kterých si každé slovo zachovává vlastní význam. U kolokace jsme na základě odborné literatury a vlastních předchozích analýz předpokládaly pouze funkce DM (tab. 24). Kolokace se objevuje většinou na konci věty, klauze či segmentu, při odlišení funkcí jsme vycházeli z předchozího kontextu a v mluveném jazyce i z intonace. Funkce označená jako výzva k reakci je spojena se stoupavou intonací, která je v psaném modu značena otazníkem. V případě druhé funkce – sumarizace – následuje kolokace *že jo* po uvedení výčtu informací, jejich shrnutí nebo po uvedení obecně známých či tradovaných faktů. Sumarizační funkce byla v analyzovaných datech s výjimkou internetových diskusí častější. Vzhledem k nízké frekvenci kolokace ve všech třech žánrech elektronické komunikace rozdíl mezi funkcemi a komunikačními mody dále neporovnáváme.

		Ortofon	Koditex-web		
			diskuse	fóra	Facebook
diskurzivní marker	výzva k reakci	18 %	100 %	31 %	25 %
	sumarizace	82 %	0 %	69 %	75 %

Tab. 24. Souhrn funkcí nalezených u kolokace *že jo*

## Shrnutí studie

Předmětem této studie byla funkční analýza výrazů označovaných jako DM. V návaznosti na nejnovější výzkum DM v češtině (Čermáková et al., 2019) a zastoupení této skupiny mezi jazykovými rysy vyděluje spontánní texty v MDA (Cvrček et al., 2018b) jsme analyzovali čtyři výrazy a jednu kolokaci ve vzorcích z mluveného a psaného elektronického jazyka. Cílem kapitoly bylo především srovnání funkcí v daných žánrech, resp. komunikačních modech, neboť právě na základě funkcí výrazu v textu je skupina DM definována. Zároveň jsme se pokusili ověřit metodologii použitou v Čermákové et al. (2019) pro psaná data a rozšířit seznam detailně analyzovaných výrazů v češtině.

Ve všech případových studiích se ukázalo, že DM jsou obecně platnou kategorií, tj. nejsou vlastní pouze mluvenému jazyku. Rozdíl mezi mluveným a psaným elektronickým jazykem se projevil zejména v podílu určité diskurzivněmarkerové funkce (např. u *že jo*), nikdy se však neukázalo, že by se daný výraz choval v jednom žánru jako DM a v jiném ne. Rozdílné podíly mohou být zapříčiněny zvolenou metodologií, tj. velikostí náhodných vzorků.

Celkem bylo detekováno jedenáct diskurzivněmarkerových funkcí, které se z většiny shodují s funkcemi popsány v Čermákové et al. (2019). Na základě převažující funkce jsme daný

výraz zařadili do tří skupin DM postulovaných tamtéž: *jo*, *viš* a *že jo* mezi interakční markery, *fakt* a *třeba* mezi konektivní. U všech analyzovaných výrazů s výjimkou *jo* a kolokace *že jo* jsou funkce DM dokladem posunu v užívání výrazu od funkcí vlastních jeho lexikálnímu významu. Tento proces je u každého výrazu v jiném stadiu: u výrazů *fakt* a *třeba* gramatikalizace rozdílů mezi částicí (tj. DM) a substantivem, resp. příslovcem, proběhla, kdežto u výrazu *viš* teprve probíhá. Výraz *jo* se od ostatních výrazů liší v tom, že za DM považujeme všechny jeho funkce včetně vyjádření souhlasu, což je podle výkladových slovníků jeho primární funkcí. Na rozdíl od Čermákové et al. (2019) chápeme souhlas jako specifickou skupinu verbální reakce (responze) adresáta. Pro analýzu *že jo* jsme zužitkovali výsledky předchozích studií (Müllerová, 2007; Komrsková, 2017). Při vyhledávání jsme odfiltrovali tzv. souvšytky od kolokací, takže jsme dále analyzovali pouze diskurzivněmarkerové funkce vlastních kolokací.

Z výsledků všech analýz vyplývá, že analyzované výrazy a kolokace jsou oprávněně zastoupeny v seznamech kontaktních slov a výplňkových výrazů použitých při MDA. Cvrček et al. (2018b, v tisku) uvádějí, že tyto seznamy byly pořízeny excerpcí odborné literatury a mluvníc, v řadě publikací se však vyskytuje pouze neúplný seznam nejčastěji uváděných slov, proto se domníváme, že naše snaha o rozšíření detailně analyzovaných potenciálních DM je smysluplná.

### 3.5.2 Pronoun non-dropping

Tento jazykový rys použil jako první N. Chomsky v knize z r. 1981 *Lectures on Government and Binding* k rozdělení jazyků na ty, ve kterých může být subjekt finitních klauzí fonologicky prázdný, tedy tzv. *pro-drop* jazyky, a ostatní, v nichž subjekt prázdný být nemůže. Další výzkumy jevu označovaného v teorii řízení a vázání jako *pro-drop* parametr ukázaly, že se v různých jazycích, v nichž se tento parametr objevuje, neprojeví stejné jevy na něj navázané (Emonds et al., 2017). Některé jazyky navíc připouštějí vynechání (*dropping*) subjektu pouze za určitých podmínek.

V této práci přebíráme pojem *pronoun non-dropping*<sup>177</sup> a jeho vymezení z MDA provedené na češtině (Cvrček et al., 2018b). Pojmem rozumíme užívání zájmena ve funkci podmětu společně s příslušným slovesným tvarem v přísudku. V češtině nese finitní sloveso (dále jen V<sub>fin</sub>) morfologické kategorie osoby a čísla, pro porozumění toho, která osoba je aktérem děje, tedy

---

<sup>177</sup> Pojmenování rysu ponecháváme v angličtině a v textu užíváme jeho konvencionalizovanou podobu *pro-drop* nebo *pro-non-drop*, a to zejména kvůli explicitnosti při porovnání vynechaných, tj. *drop*, a nevynechaných, tj. *non-drop*, výskytů zájmena.

plně postačuje slovesný tvar, zatímco např. v angličtině není z většiny slovesných tvarů tato informace vyvoditelná. Nevyjádřený podmět je v češtině podle Grepla a Karlíka (1998: 226) pokládán za základní, bezpříznakový. Jak ukázali např. Hirschová (1984) či Nekula (2003b), je třeba odlišit vyjadřování podmětu zájmeny 1. a 2. os. od pronominalizace podmětu zájmeny 3. os. Rozdíl mezi nimi tkví podle Zimové (1988) v tom, že zájmena 1. a 2. os. nezastupují plnovýznamová pojmenování ani neplní funkci textově odkazovací, ale rozlišují role komunikantů (tj. mluvčího a adresáta). Faktory ovlivňující vyjádření 1. a 2. os. zájmenem rozdělují Grepl – Karlík (1998: 226n.) na komunikační, konstrukční a věcné. Za nejdůležitější je např. u Zimové (1988) a Štíchy (2010) označeno aktuální členění, které řadí Grepl a Karlík (1998) mezi faktory komunikační. Podle Zimové (1988) je stejně významným faktorem gramatická výstavba výpovědi, jež PMČ zařazuje do skupiny konstrukčních faktorů. Nakonec za věcné faktory považuje PMČ užívání zájmena v případech, ve kterých není osoba dostatečně určena koncovkou V<sub>fin</sub>, např. ve větě *Já to věděla.*, nebo u kontextové elipsy přísudku, např. ve větě *Ty si to možná myslíš, já ne.* Z dalších faktorů uváděných ve zde citované literatuře jmenujme zřetel k modální výstavbě výpovědi, emocionální aktualizaci, rytmus<sup>178</sup> a vliv funkční stylové oblasti. K distribuci tohoto jevu v mluveném jazyce Grepl a Karlík (1998: 227) poznamenávají, že zájmena *já, ty, my, vy* se vyskytují častěji než v psaných textech a slouží k dosažení důvěrného tónu nebo se vyskytují ve vzrušené řeči.

V MDA se tento jazykový rys projevil mezi charakteristikami spontánních nepřipravených textů, frekvenci jeho výskytů tedy můžeme považovat za signál poukazující na míru spontánnosti. Kromě srovnání mluveného jazyka a vybraných žánrů elektronické komunikace podle míry spontaneity se v této studii zaměříme na to, jakým způsobem je spontánnost prostřednictvím tohoto rysu konstruována. Pokusíme se na základě kontextu v širším smyslu, tj. nejen verbálního, ale i na základě znalostí o průběhu rozhovoru, tématu textu apod., popsat důvody, proč pisatel/mluvčí zájmeno nevynechal.

### **Metodologie**

U tohoto jevu jsme se nejprve zaměřili na kvantitativní rozdíl mezi případy, ve kterých a) před V<sub>fin</sub> nepředchází osobní zájmeno pro příslušnou slovesnou osobu (*drop*) a b) V<sub>fin</sub> předchází příslušné osobní zájmeno (*non-drop*). V MDA je tento jev operacionalizován pro případ b) ve

---

<sup>178</sup> Roli rytmu při užívání pronominálního podmětu se věnoval i Mathesius (1947). Pronominální podmět, kterým věta začíná, podle jeho názoru odpovídá tendenci češtiny začínat klesavým rytmem, od něhož se přechází k důraznějším částem věty, uvádějící nové informace (réma). Tato tendence se projevuje i v zástupnosti pronominálního podmětu za jiné stejně dlouhé slovo, nejčastěji *to* nebo jednoslabičné příslovce. Mathesius dále upozornil na častost jevu v hovorové, tj. mluvené češtině.

všech slovesných osobách, záběr této studie jsme však zúžili ve dvou směrech, a to pouze na 1. a 2. slovesnou osobu jednotného čísla ve všech slovesných způsobech kromě imperativu. Mimoto přejímáme z MDA zaměření pouze na anteponované, nikoli postponované zájmeno.<sup>179</sup> Posledně jmenované omezení vychází z předpokladu o menší příznakovosti pronominalizovaného subjektu v antepozici přísudku v oznamovacích větách (Nekula, 2003a), a tudíž i jeho vyšší frekvenci.<sup>180</sup> Přestože se zaměřujeme pouze na dvě slovesné osoby v singuláru, přikládáme tabulku 25 pro představu o frekvenci slovesných osob v rámci analyzovaných dat.<sup>181</sup>

	Ortofon	Koditex-web		
		diskuse	fóra	Facebook
1. os. sg	36 724 (29 699 ipm)	3 778 (15 685 ipm)	10 141 (41 813 ipm)	5 889 (24 255 ipm)
2. os. sg	18 256 (14 764 ipm)	993 (4 122 ipm)	2 434 (10 036 ipm)	598 (2 463 ipm)
3. os. sg	763 498 (70 170 ipm)	17 042 (70 756 ipm)	17 308 (71 365 ipm)	16 242 (66 897 ipm)
1. os. pl	73 125 (6 720 ipm)	1 057 (4 388 ipm)	1 572 (6 481 ipm)	1 876 (7 726 ipm)
2. os. pl	55 200 (5 073 ipm)	1 731 (7 186 ipm)	659 (2 717 ipm)	2 112 (8 698 ipm)
3. os. pl	214 176 (19 684 ipm)	5 786 (24 022 ipm)	3 307 (13 635 ipm)	5 383 (22 171 ipm)

Tab. 25. Frekvence slovesných osob v analyzovaných datech

V analýze jsme se soustředili na rozdíly mezi mluveným a psaným jazykem včetně jeho žánrů ve frekvenci případů a) a b) a v analýze bezprostředního kontextu. Zajímalo nás, jaké lexémy se objevují na pozici V<sub>fin</sub>, zda se objevují častěji s non-drop, nebo drop zájmenem, na kolikáté pozici v segmentu/větě se V<sub>fin</sub> nachází a zda jsme schopni z kontextu vyčíst důvody pro použití non-drop zájmena.

K naplnění posledně jmenovaného cíle této studie byla využita kvalitativní analýza případů a) a b) vyskytujících se v textech od jednoho autora, resp. mluvčího. Tuto mikroanalýzu jsme

<sup>179</sup> Zúžením zkoumaného jevu se snažíme jednak dosáhnout co nejlepších výsledků (zejména v mluveném jazyce není vždy jasné, zda se zájmeno vztahuje k V<sub>fin</sub>, nebo je vyřčeno z jiných důvodů), jednak zúžit množství faktorů, které mohou ovlivňovat pořadí subjektu a predikátu a užívání pronominálního subjektu.

<sup>180</sup> Postup pro vyhledání jevu byl pro mluvený a psaný jazyk stejný. Nejprve jsme vyhledali všechna V<sub>fin</sub> v 1. a 2. os. sg. pomocí dotazu [tag="V..S...[12].\*"]. Pro získání případů a) jsme použili negativní filtr v podobě ([lc="já" within [lc="já" [pos!="[ZJV]" ]\* [tag="V..S...1.\*" within <s/>] | [lc="ty" within [lc="ty" [pos!="[ZJV]" ]\* [tag="V..S...2.\*" within <s/>] | [lc="tys" within [lc="tys" [pos!="[NAPCJZXV]" ]\* [tag="V..S...2.\*" within <s/>] pro psaný jazyk (pro mluvený jazyk byla namísto struktury <s> struktura <sp>). Pro získání případů b) jsme stejný dotaz použili jako pozitivní filtr.

<sup>181</sup> Čísla v tabulce vycházejí z aktuálního morfologického značkování.

provedli pouze u mluveného jazyka, kde jsme mohli díky většímu objemu dat oproti psané elektronické komunikaci vybrat texty, které bez jakékoli pochybnosti pocházejí od jednoho mluvčího. Mikroanalýza byla provedena na třech nahrávkách,<sup>182</sup> v nichž vystupuje mluvčí Věra\_19. Zvolena byla ze dvou důvodů: v jejích segmentech patřila frekvence případů a) i b) mezi tři nejfrekventovanější, vedle toho byly k dispozici nahrávky, ve kterých hovoří Věra\_19 pouze s jedním dalším mluvčím. Rozhovor probíhající pouze mezi dvěma mluvčími je pro analýzu, která má k dispozici pouze zvukovou nahrávku, jednodušší, protože je vždy zřejmé, komu je určen právě pronášený segment a který z mluvčích právě mluví. V takových nahrávkách je také menší počet nesrozumitelných úseků, tj. úseků, v nichž nejsme schopni rekonstruovat vyřčená slova. Ve dvou nahrávkách je s Věrou\_19 zachycen její manžel Jan\_20, přičemž v každé nahrávce mluví více jiný mluvčí, ve třetí nahrávce je zachycen Věřin rozhovor s kamarádkou Šárkou\_93 (tab. 27).

Při mikroanalýze jsme případy V<sub>fin</sub> s drop i non-drop zájmenem anotovali ručně v celých nahrávkách. Na rozdíl od automatického vyhledávání, použitého pro frekvenční srovnání lexikálního obsazení a pozice jevu v segmentu/větě, jsme v mikroanalýze vzali v úvahu anteponované i postponované zájmeno.

### **Analýza**

Nejprve jsme se zaměřili na kvantitativní porovnání V<sub>fin</sub> s drop a non-drop zájmenem. Z tabulky 26 je patrné, že obecně častější je vynechání pronominálního zájmena, tj. případ V<sub>fin</sub> s drop zájmenem. Rozdíl mezi frekvencí je výraznější v psaném jazyce. Frekvenční srovnání žánrových rozdílů v rámci drop a non-drop skupin ukazuje, že non-drop výskyty pronominálního zájmena jsou typičtější pro mluvený jazyk, zatímco v psaném jazyce se jejich výskyt jeví jako nevýrazný.

---

<sup>182</sup> Využity byly celé nahrávky dostupné z repozitáře Lindat/Clarín: <https://lindat.mff.cuni.cz/repository/xmlui/handle/11234/1-2580>.

		Ortofon	Koditex-web		
			diskuse	fóra	Facebook
Vfin1	non-drop	33,9 % 12 441 (10 061,4 ipm)	9,3 % 350 (1 453,16 ipm)	10,6 % 1 070 (833,69 ipm)	7,6 % 446 (1 836,98 ipm)
	drop	66,1 % 24 283 (19 638,37 ipm)	90,7 % 3 428 (14 232,69 ipm)	89,4 % 9 071 (40 980,21 ipm)	92,4 % 5 443 (22 418,55 ipm)
Vfin2	non-drop	9,2 % 1 681 (1 359,47 ipm)	7 % 70 (290,63 ipm)	5,8 % 141 (581,38 ipm)	2,8 % 17 (70,02 ipm)
	drop	90,8 % 16 575 (13 404,68 ipm)	93 % 924 (3 836,35 ipm)	94,2 % 2 293 (9 454,62 ipm)	97,3 % 581 (2 393,01 ipm)

Tab. 26. Frekvence jevu pro 1. a 2. osobu v analyzovaných datech (v prvním řádku je uveden podíl daného případu, ve druhém absolutní frekvence, ve třetím ipm)

Frekvenční rozdíl je patrný nejen mezi žánry, ale i mezi 1. a 2. os. sg. (tab. 26). v mluveném i psaném jazyce je častější 1. os. sg. oproti 2. os. sg. Nižší frekvenci 2. os. sg. v psaném jazyce oproti mluvenému můžeme interpretovat tak, že ve vybraných psaných elektronických textech se jedná především o sebevyjádření autora, ve kterém není důležité obracet se na recipienta pomocí 2. os. sg., zároveň se může jednat o zkreslení způsobené zaměřením pouze na jeden jazykový prostředek, kterým se může autor obracet k recipientovi.

Při analýze lexikálního obsazení Vfin se ukázalo, že v psaném i mluveném jazyce se mezi deseti nejčastějšími Vfin nezávisle na drop a non-drop výskytu zájmena objevují tvary slovesa *být* pro přezens, futurum a kondicionál, tj. *jsem, budu, bych, jsi, budeš, bys*. Tvary *jsem* a *jsi* jsou zároveň ve všech případech s výjimkou Vfin2 v mluveném jazyce na prvním místě, což je dáno zejména tím, že se jedná o výskyty plnovýznamového *být* v přítomnosti i pomocného slovesa (zejména pro minulý čas). Jako nejčastější tvar se v mluveném jazyce pro Vfin2 objevuje tvar slovesa *být* v podobě příklonného -s (v transkriptech zapisovaný jako #s). Kromě něj se mezi prvními deseti tvary v mluveném jazyce objevil nestandardní tvar 2. os. sg. od slovesa *být seš*, který se v použitých psaných datech vyskytoval minimálně (nejvíce 11× v diskusních fórech v podobách *seš* i *jseš*). Mezi deseti nejčastějšími tvary pro 1. a 2. os. sg. byl v mluveném i psaném jazyce dále afirmativní i negativní tvary od sloves *mít* a *vědět* (tvar *nevím* je frekventovanější než *vím*). U Vfin1 se ještě objevuje sloveso *myslet* (v afirmativu), jehož tvary ve Vfin2 se pohybují mezi nejčastějšími dvaceti tvary.



Lexikální obsazení V<sub>fin1</sub> a V<sub>fin2</sub> se odlišuje vyšší mírou shody mezi prvními deseti lexémy pro V<sub>fin1</sub>, což souvisí s vyšší frekvencí této slovesné osoby (tab. 25). Největší odlišností mezi mluveným a psaným jazykem u V<sub>fin1</sub> je přítomnost tvaru *řikám* mezi deseti nejčastějšími V<sub>fin</sub> v mluveném jazyce, zatímco v psaném jazyce je tento tvar, pokud se objevuje, mezi třiceti nejčastějšími. Tvar *prosím* se v první desítce kromě mluveného jazyka vyskytuje ještě v příspěvcích z Facebooku, v dalších psaných žánrech patří k nejčastějším dvaceti V<sub>fin</sub>. Přítomnost tvaru *řikám* je ovlivněna jeho použitím ve funkci uvozovacího prvku reprodukované řeči (např. Hoffmannová et al., 2017), zatímco *prosím* je využito přímo v dané komunikační situaci pro zdvořilou žádost. Největší odlišností u V<sub>fin2</sub> byla přítomnost imperativu *počkej* v mluveném jazyce a *viz* v textech z internetových diskusí a Facebooku. Mezi dvaceti nejčastějšími lexémy se v mluveném i psaném jazyce vyskytovaly také tvary od sloves *moci* (v afirmativu i negativu), *muset* (v afirmativu) a *chtít* (v afirmativu i negativu).

Výše zmíněné V<sub>fin</sub> se objevovaly u případu drop i non-drop zájmena. Rozdíly v lexikálním zastoupení V<sub>fin</sub> se u příslušné osoby projevovaly v závislosti na celkové frekvenci jevu (tab. 26). U V<sub>fin1</sub> byly největší rozdíly mezi mluvenými a psaným modem; žánry psané elektronické komunikace se vzájemně odlišovaly méně výrazně. Zaměříme-li se zvláště na V<sub>fin</sub> vyskytující se s non-drop zájmenem a zvláště s drop zájmenem, můžeme potvrdit větší lexikální různorodost mezi oběma případy u V<sub>fin2</sub> oproti V<sub>fin1</sub> a také větší rozdílnost v rámci psaných žánrů.

Rozdílnost se začíná ukazovat v pořadí (ranku) i frekvenci V<sub>fin</sub> následujících po první desítce (viz výše). Například v mluveném jazyce u V<sub>fin1</sub> se slovesa pohybu (*jdu, jedu, pojedu, půjdu*), slovesa vyjadřující vnímání (*chápu, pamatuju, podívám, vidím, znám*) a slovesa *dám, dělám, potřebuju* vyskytují v podobném pořadí, tj. s maximálním rozdílem 10 pozic, s non-drop i drop zájmenem. Naopak slovesa *doufám, těším se, kecám* jsou výrazně častější (tj. rozdíl v pořadí je více než 10 pozic) s drop zájmenem, zatímco V<sub>fin</sub> *su, napíšu, povídám* s non-drop zájmenem. Spojnicí mezi psanými žánry a mluveným jazykem je V<sub>fin</sub> *prosím*, který se vyskytuje výrazně častěji s drop zájmenem než s non-drop. Z dalších V<sub>fin1</sub> výrazných u psaných žánrů jmenujme následující: na diskusních fórech je častější *věřím* s non-drop zájmenem a *chci* s drop zájmenem, v internetových diskusích se V<sub>fin</sub> *souhlasím* vyskytuje podobně v drop i non-drop případech, ale *nesouhlasím* jen s drop zájmenem, zatímco V<sub>fin</sub> *doufám* s non-drop zájmenem, ve facebookových příspěvcích je častější *dávám* a *netvrdím* s non-drop zájmenem a V<sub>fin</sub> *děkuji, musím, řikám* s drop zájmenem. Slovesa mluvení *řikat* a *řict* se v tvarech přítomného či

minulého času sice vyskytují nejčastěji v mluveném jazyce, avšak objevují se i v psaných žánrech, všude častěji s drop než non-drop zájmenem.

Jak už bylo řečeno výše, u V<sub>fin2</sub> je lexikální různost mezi všemi analyzovanými žánry větší, a to především v souvislosti s nižší frekvencí. Projevuje se to zejména u seznamu V<sub>fin2</sub> s non-drop zájmenem, u něhož se už mezi deseti nejfrekventovanějšími tvary objevují frekvence v jednotkách. Lexikální srovnání je v takovém případě téměř zbytečné, proto se omezíme pouze na dva postřehy. Prvním z nich je přítomnost sloves vztahujících se k argumentaci jiných pisatelů na diskusních fórech a v internetových diskusích (např. *doložit, pochopit, reagovat, srovnávat, tvrdit, vyblít*), zatímco v mluveném jazyce se objevují lexémy vztahující se k tématu hovoru, nikoli způsobu komunikace. Druhým je vyšší frekvence imperativních tvarů v mluveném jazyce (v úvahu bereme pouze drop zájmeno) a rozdílná paleta těchto tvarů: psaných žánrech se objevily např. *ber, (ne)dej, měj, piš, věř, zkus*, zatímco v mluveném jazyce např. *běž, dej, jdi, neboj, nech, neříkej, nekecej, počkej/počki, podívej, promiň, řekni, sedni*. Lexikální srovnání V<sub>fin2</sub> s non-drop a drop výskyty neukazuje ani v mluveném jazyce výrazné rozdíly; za zmínku stojí pouze častější použití V<sub>fin</sub> *koukáš* s non-drop zájmenem.

Položili jsme si také otázku, na jaké pozici v segmentu/větě se zkoumaný jev vyskytuje. Zaměřili jsme se na první dvě pozice na začátku a na konci segmentu/věty, resp. na vzdálenost jednotlivých výrazů od začátku a konce segmentu. V mluveném jazyce se na první pozici<sup>183</sup> v segmentu téměř stejně často objevily výskyty V<sub>fin1</sub> s non-drop a drop zájmenem (1922 vs. 2093), což je překvapivé, protože spojení non-drop zájmena s V<sub>fin</sub> je vždy pouze malým výsekem ze všech slov, která mohou V<sub>fin</sub> předcházet. Ve všech ostatních případech – tj. u V<sub>fin2</sub> v mluveném jazyce, u V<sub>fin1</sub> a V<sub>fin2</sub> v psaném jazyce – byla situace stejná: na první pozici výrazně převažovaly výskyty s drop zájmenem nad výskyty s non-drop zájmenem (v mluveném jazyce se vyskytovalo drop zájmeno u V<sub>fin2</sub> v 11 %, v psaném jazyce u V<sub>fin1</sub> v 14,7 %, u V<sub>fin2</sub> v 5 %).

Pokud byl zkoumaný jev na druhé pozici od začátku segmentu/věty, ve všech případech převažovaly případy s non-drop zájmenem. Na první pozici se pak shodně v mluveném i psaném jazyce nejčastěji objevovala jedno- a dvouslabičná funkční slova, která slouží k provázání textu: *a, ale, no, tak, to*.

---

<sup>183</sup> Pro zjednodušení považujeme spojení non-drop zájmena s V<sub>fin</sub> za jednu pozici. Při analýze pozic jsme brali v úvahu pouze výskyty těsného spojení, tj. mezi zájmenem a V<sub>fin</sub> nebylo žádné jiné slovo ani interpunkce.

Výskyt jevu na konci segmentu/věty vykazoval jednotnou tendenci: V<sub>fin</sub> s drop zájmenem byl v mluveném i psaném jazyce častější oproti případům s non-drop zájmenem jak na poslední pozici v segmentu/větě, tak na předposlední. Pokud byl analyzovaný jev na předposlední pozici, na té poslední se v mluveném jazyce nejčastěji objevovala jednoslabičná slova *co*, *no*, *že*. Naproti tomu v psaném jazyce nemá smysl jakákoli slova uvádět, protože společný průnik stejných slov pro (non)-drop V<sub>fin1</sub> a V<sub>fin2</sub> byl malý a řádově v jednotkách výskytů.

Posledním bodem analýzy bylo porovnání výskytů drop a non-drop zájmena s V<sub>fin</sub> v textech od jednoho mluvčího. Zajímalo nás, zda je možné rekonstruovat u případů V<sub>fin</sub> s non-drop zájmenem za pomoci analýzy širšího kontextu, zvuku, znalostí o nahrávané situaci a jejím průběhu důvody, proč se v nich objevuje zájmeno. Tabulka 27 ukazuje kromě informací o nahrávkách frekvenci hledaného jevu; do frekvence jsou započítány výskyty non-drop zájmena v prepozici i postpozici. Frekvence slovesných osob je v souladu s automatickým vyhledáním na větším objemu dat.

označení nahrávky	označení mluvčího	kolik % slov zaujímá mluvčí	frekvence drop V <sub>fin1</sub>	frekvence non-drop V <sub>fin1</sub>	frekvence drop V <sub>fin2</sub>	frekvence non-drop V <sub>fin2</sub>
12A004N	Věra_19	59 %	20	22	29	0
	Jan_20	41 %	23	13	4	2
13A011N	Věra_19	54 %	71	83	19	3
	Šárka_93	46 %	79	46	50	1
13A015N	Věra_19	45 %	58	36	34	4
	Jan_20	55 %	66	34	18	1

Tab. 27. Souhrnné informace o nahrávkách v subkorpusu Věra

Zájmeno *já* a *ty* v postpozici se v celém subkorpusu vyskytovalo 7×, což tvoří 1 % všech výskytů non-drop zájmen. Non-drop zájmeno v postpozici bylo použito ke zdůraznění subjektu mluvčího, např. *na s\* na schodech jsem stál já když jsem přišel . takhle, kde ji máš ty?* Dalším důvodem pro ponechání (místo vynechání) zájmena bylo odlišení jednotlivých mluvčích a rozdělení jejich činností (př. 50). S tím je často spojeno zdůraznění subjektu jednoho mluvčího v kontrastu ke druhému (př. 51, 52). V př. 52 je tento důraz posílen opakováním různých tvarů zájmena *já* u každého V<sub>fin</sub>, které se vztahuje k dané mluvčí.

(50)

Jan\_20: a hlavně se těším na pak na věcu .. pojedu . tam **pojedu já** .. [ty pojeděš zpátky]

Věra\_19:

[hmm .. no] to seš velmi vtipnej

teda takže **já budu .. já budu abstinovat** zase

(51)

Věra\_19: dobrý kafe .. ten velvet je dobrej [prostě]

Jan\_20: [já mám] taky dobrý no to nový .. nový nový z Lidlu

(52)

Věra\_19: hmm a je tohle to je dobrý ? vo čem to vůbec je ? .. **já jsem** to **neviděla** nikdy

Jan\_20: no jak @ kdysi dávali ten seriál v televizi

Věra\_19: no jo tak kdysi dávali seriál **já jsem se** na to **nekoukala** .. mě to nezajímalo takovýdle věci .. a nezajímá mě to ani teď

Hlasové zdůraznění zájmena se v subkorpusu vyskytovalo spíše okrajově, častější byly případy, kdy se hlasové zdůraznění týkalo spojení zájmena se slovesem. V př. 53 dochází na spojení *já nevím* k rychlému stoupenutí a klesnutí hlasem. V př. 54 už je hlasově zdůrazněno pouze *Vfin nevím*, na kterém dochází jednak ke zpomalení tempa a jednak ke stoupenutí a následnému klesnutí hlasem.

(53)

Věra\_19: tak vona třeba půjde pěšky že jo ? to **já nevím**

(54)

Věra\_19: hele notebook supr jako .. úplně mi to vyhovuje .. protože **já jsem se** tamhle .. u toho počítače **hrbila** . **já nevím** jak tam můžeš u toho sedět [takovou dobu]

Poměrně často se non-drop zájmeno objevovalo při reprodukci rozhovoru (př. 55). V takovém případě slouží jednak k odlišení konkrétních osob, jednak jako uvozovací prvek předcházející vlastní reprodukci. Zároveň může být přímo součástí řeči, která je reprodukována (př. 56).

(55)

Věra\_19: a on říká .. ty jo vy máte snad od Dánikena úplně všechno ne ? .. [a **já jsem říkala** zdaleka] @ **já jsem říká\*** .. no **já nevím** kde na to přišel .. protože **jsem říkala** zdaleka ne **já jsem říkala** . nesplet sis to . se Sitchinem ?

(56)

Jan\_20: ale říkala tak @ .. ale @ . a **ty už máš** třetí dítě ne ? ne říkám druhý a @ . a . poslední ..

V další části mikroanalýzy jsme se zabývali porovnáním výskytů nejfrekventovanějšího plnovýznamového *Vfin* s drop a non-drop zájmenem. Vzhledem k nízké četnosti non-drop zájmena *ty* jsme se zaměřili pouze na zájmeno *já*. Nejfrekventovanějším slovesným tvarem u všech mluvčích v subkorpusu bylo *nevím*. Počet výskytů s drop a non-drop zájmenem byl

podobný: 23× drop, 21× non-drop. Pro užívání delšího spojení *já nevím* platí výše zmiňované důvody, kromě nich jsme identifikovali další typizovaný způsob užívání, za jehož motivaci považujeme získávání času pro další produkci (př. 57). Týká se jak umístění non-drop zájmena s V<sub>fin</sub> do relativně plynulého proudu řeči, tak na začátek (př. 58) nebo konec úseku, který je zvukově (pauzou, rozdílnou intonací) vydělen od následujícího nebo předchozího (př. 59).

(57)

Jan\_20: ne tam jsou ty Mu\* to **já nevím** jak .. tam .. mužský tam je nějaký to .. @ ty skály nějaký tam jsou

(58)

Věra\_19: von teda vstává fakt vo půl sedmý tak nějak .. on se posadí na posteli a prostě .. **já nevím** já když se ráno probudím tak jsem nejdřív taková trošku jako malátná a von hnedka běží někam a hned chce jíst

(59)

Věra\_19: něco takovýho nebylo to za tebe ještě ? .. nebo **já nevím**

Poslední částí mikroanalýzy bylo porovnání rozdílu v idiolektu jednotlivých mluvčích vzhledem ke zkoumanému jevu. Ve všech nahrávkách se oba mluvčí podílejí na komunikaci proporcčně podobně, největší rozdíl je v nahrávce s Věrou\_19 a Janem\_20, ve které její řeč zabírá 59 % řeči v celé nahrávce, a jeho tedy 41 % (tab. 27). Vliv podílu řeči jednoho mluvčího na zkoumaný jev jsme však nepozorovali. Proměny idiolektu můžeme díky složení subkorpusu sledovat nejlépe u Věry\_19, zatímco u Šárky\_93 a Jana\_20 můžeme konstruovat pouze jeho podobu, ale ne proměny související s komunikací v jiných situacích a s jinými mluvčími.

První charakteristika zkoumaného jevu ve Věřině idiolektu je patrná ze spojení V<sub>fin1</sub> s drop zájmenem, která se u Věry\_19 pohybuje vždy v podobné frekvenci jako u jejího komunikačního partnera (tab. 27). Předpokládáme, že je to ovlivněno situačně a tematicky. Ženy se setkaly na pravidelné návštěvě, při které porovnávaly prostřednictvím vyprávění v 1. os. sg. svoje zážitky a zkušenosti; frekvence zkoumaného jevu je u obou žen vysoká. Naopak v manželském rozhovoru, který se sice v první nahrávce týkal každodenní rutiny a organizačních záležitostí, ale ve druhé se zaměřil také na vyprávění o procházce s dítětem a jednání na úradě, chyběla komparace názorů a zkušeností jako u žen a při části vyprávění bylo použito plurálu. V nahrávkách manželů je Věra\_19 tím, kdo se aktivněji obrací na druhého prostřednictvím 2. os. sg. Dalo by se říci, že se jedná pouze o stereotypní snahu manželky zatáhnout manžela do rozhovoru, avšak Věřina snaha je zároveň ovlivněna skutečností, že rozhovory nahrávala,

takže bylo v jejím zájmu udržovat komunikaci. Tato skutečnost se naopak marginalizuje v rozhovoru s kamarádkou, kterou nemusela vůbec pobízet; frekvence 2. os. sg. je u Věry\_19 výrazně nižší než u Šárky\_93.

Kromě výrazného využití 2. os. sg. je v Šárčině idiolektu patrná tendence užívat finitní sloveso bez explicitně zmíněného zájmena. U Jana\_20 se díky dvěma různým nahrávkám tato tendence nejeví tak silně jako u Šárky\_93, u níž jsme analyzovaly jednu nahrávku.

Při porovnání drop a non-drop výskytů zájmen u slovesa *nevím* v idiolektech všech mluvčích se ukazuje, že drop výskyty jsou častěji vysloveny redukovane. Podle našeho názoru to souvisí s výše zmiňovanými důvody, kdy je použito *já nevím* a kdy pouze *nevím*. Ve většině případů se dal výskyt non-drop zájmena v porovnání s drop výskytem odůvodnit, pro kompletnost však uvádíme i př. 60 a 61, u kterých neumíme ani za pomoci zvuku nebo kontextu určit, proč tatáž mluvčí v tomtéž vyprávění o lékařské kontrole jednou použije *já nevím* a podruhé *nevím*.

(60)

Šárka\_93: a tak třeba zkouší to oddělení nebo **já nevím** jak to maj na nějakou praxi jestli .. taky obcházej nějaký kolečka nebo kd\* jak to maj

(61)

Šárka\_93: já si myslím že mě píchali do prstu a pak .. ty jo **nevím** jestli m\* asi mi taky brali krev normálně

U Věry\_19 jsme na rozdíl od ostatních mluvčích konstatovali tendenci užívat *nevím* s drop zájmenem tam, kde si není jista odpovědí. V tomto ohledu je zajímavý př. 62, ve kterém vyjadřuje Šárka\_93 svou neznalost zopakováním a modifikací fráze *nevím vůbec – vůbec nevím*, kterou následně postaví do kontrastu s informací, kterou ví, prostřednictvím zájmena *já* u slovesa *vím*.

(62)

Věra\_19: a to je v porodnici nebo @ v nemocnici tam někde [neviš]

Šárka\_93: **[nevím]** vůbec .. vůbec **nevím** .. **já vím** že někdo chodil na rehabilitace tam jak je to Papírový náměstí

## Shrnutí

V této studii jsme se zabývali jevem nazvaným pronoun non-dropping u sloves v 1. a 2. os. sg. Porovnávali jsme frekvenční a lexikální rozdíly mezi oběma osobami, mezi mluveným a psaným jazykem (ve třech žánrech elektronické komunikace) a mezi finitními slovesnými tvary, u kterých se v prepozici vyskytuje (non-drop), či nevyskytuje (drop) pronominální

subjekt. Výsledky kvantitativní analýzy, které vycházejí z automatického vyhledávání jevu, se týkají pouze anteponovaného zájmena. Při kvalitativní ruční anotaci jevu v mluveném subkorpusu Věra byla postpozice zájmena identifikována v 1 % všech výskytů non-drop zájmena u V<sub>fin</sub>. Tento údaj je platný pouze pro daný subkorpus, ve větším objemu dat se frekvence může radikálně změnit. Při analýze postpozice by však bylo třeba uvažovat i aktuální větné členění pro různé typy vět, např. otázky, vedlejší věty, což by v analyzovaném typu mluveného jazyka vyžadovalo jednotnou koncepci pro rozhodnutí hranic vnitrosegmentálních (či vnitroreplikových) struktur. Kvůli snaze o jednotnost v přístupu k datům jsme proto v analýze vycházející z automatického vyhledání jevu pracovali pouze se zájmenou antepozicí.

Frekvenčně vyšší zastoupení 1. os. sg. oproti 2. os. sg. ve všech analyzovaných textech ovlivnilo analýzu rozdílů v lexikálním zastoupení V<sub>fin2</sub> mezi případy s non-drop a drop zájmenem. Větší lexikální různost byla u jevu s celkovou nižší frekvencí v daném žánru, proto nebyla tato část analýzy nijak obsáhlá. Konstatovat můžeme vysokou míru shody nejčastějších 10 V<sub>fin</sub> ve všech žánrech nezávisle na výskytu příslušného anteponovaného zájmena. Rozdíl mezi psaným a mluveným jazykem se v tomto případě týkaly častějšího zastoupení v mluveném jazyce oproti psanému: V<sub>fin2</sub> slovesa *být* v podobě příklonného *-s* nebo nestandardního tvaru *seš*, V<sub>fin1</sub> *prosím*, *říkám*, *počkej*. Při komparaci V<sub>fin</sub> ve spojení s drop a non-drop zájmenem se jako shodné pro všechna data ukázalo častější používání slova *prosím* bez anteponovaného zájmena. Při určování dílčích lexikálních rozdílů mezi žánry navzájem i mezi případy V<sub>fin</sub> s drop a non-drop zájmenem jsme kromě frekvence příslušných V<sub>fin</sub> brali v úvahu i jejich pořadí v rámci seznamu V<sub>fin</sub> pro každý žánr. Zajímavá jsou především V<sub>fin</sub> vyskytující se s non-drop zájmenem. V mluveném jazyce se jednalo o *koukáš*, *napišu*, *povídám* (naopak V<sub>fin</sub> od sloves *říkat*, *říct* je ve všech žánrech používáno častěji s drop než non-drop zájmenem), *su* a v psaném jazyce o *dávám*, *doufám*, *netvrdím*, *věřím*. U psané elektronické komunikace se zejména ve 2. os. sg. ukázalo významné zastoupení sloves vztahujících se ke způsobu komunikace, resp. argumentace pisatelů, naopak v mluveném jazyce se slovesa týkala spíše tématu hovoru a zachycené komunikační situace.

Rozdíl v použití non-drop a drop zájmena jsme se dále pokusili vztáhnout k pozici jevu na hranicích segmentu/věty, tj. na dvou prvních a posledních pozic od hranic. Frekvenční analýza ukázala, že na první pozici se s výjimkou V<sub>fin1</sub> v mluveném jazyce vyskytuje nejčastěji případ s drop zájmenem. V případě poslední pozice byla tendence jednotná: případy s drop zájmenem jsou zde častější než případy s non-drop zájmenem. Při porovnání *já nevím* vs. *nevím*

v mluveném subkorpusu Věra se ukázalo, že případ s drop zájmenem je častěji vyslovován redukovaně. Tato zjištění nás vedou k úvaze o tom, že mluvčí přisuzuje začátku svého segmentu větší důležitost než jeho konci; nechce-li, aby byla informace nesená  $V_{fin1}$ , která se vztahuje k jemu samému, upozaděna, zdůrazní ji přidáním zájmena *já*.

$V_{fin}$  *nevím* také sloužilo jako příklad pro porovnání stejných lexémů lišících se pouze (ne)přítomností zájmena. Užívání tohoto  $V_{fin}$  společně s non-drop zájmenem odráží motivace pro užívání non-drop zájmena s jakýmkoli  $V_{fin}$ : zdůraznění subjektu mluvčího implicitně i explicitně v kontrastu ke druhému mluvčímu, např. kvůli jejich odlišení nebo rozdělení úloh (a to i v rámci reprodukované řeči), a použití delšího jazykového prostředku z důvodu získání delšího času pro produkci mluvené řeči. K těmto dvěma základním motivacím jsme dospěli analýzou idiolektu mluvčí Věry\_19 se dvěma rozdílnými mluvčími.

Zatímco první z nich není třeba v souvislosti se psaným jazykem rozporovat, u druhé to neplatí. Produkce psaných elektronických textů analyzovaných v této práci nejspíš nebyla ovlivněna časovým tlakem, protože dané elektronické žánry jsou asynchronní povahy. Na druhou stranu je tato komunikace dialogická, většina textů je reakcí na jiný, dříve zveřejněný text, a také needitovaná jiným pisatelem před svým zveřejněním, což jsou dvě charakteristiky, ve kterých se tyto psané texty podobají analyzovanému typu mluveného jazyka. Výskyt non-drop zájmena s  $V_{fin}$  v psaném jazyce, kterému neumíme přisoudit žádný ze způsobů zdůraznění, pak může být jednoduše ovlivněn častějším užíváním tohoto spojení v mluveném jazyce, odkud je přenesen do psaného textu.

### 3.5.3 Iterace grafémů

Zmnožení čili iterace (multiplikace) grafémů je v literatuře uváděno mezi typickými rysy elektronické komunikace (např. Jílková, 2017; Herring, 2012). Tento jev je vykládán v souvislosti s expresí emocí ve vizuálním kanálu: iterací grafémů se pisatel snaží napodobit mluvení prodchnuté emocemi tak, že způsobí prodloužení hlásek (např. Bieswanger, 2013). Při mluvení však není vyjádření emocí jediným důvodem, proč k prodloužení hlásky dochází, může se jednat o způsob, jakým získat čas při přemýšlení nad dalšími slovy, jak naznačit formulační obtíže či absenci informace, jak oddálit výměnu mluvčích apod. Prodloužení tedy může plnit podobné funkce jako některé diskurzivní markery (srov. 3.5.1).

Iteraci lze také použít k ozvláštňení psaného textu. Ačkoli může pisatel iterovat jakákoli písmena, v reálných datech se tak většinou neděje: Jílková (2017) uvádí, že iteraci podléhají vokály a souhlásky trvací, tedy hlásky, které lze prodloužit i ve výslovnosti. Tím je



odůvodňována motivace pro iteraci. Předmětem této studie bude ukázat, v jakých případech k iteraci v elektronické komunikaci dochází, jaké hlásky jsou iterovány a v jakých slovech. Výsledky budou následně porovnány s mluvenými daty.

### Metodologie

Slova, ve kterých se objevuje alespoň jeden zmnožený grafém byla vyhledána v korpusu Koditex-web pomocí série dotazů cílících na iterovaný grafém, který může být na jakékoli pozici ve slově. V úvahu byly brány všechny grafémy používané v českém pravopisu, tj. včetně y, ý, ů.<sup>184</sup>

Z nalezených výsledků byly ručně odstraněny případy, ve kterých je iterace grafémů náležitá, konstituje dané slovo a nemůže se jednat ani o překlep, ani o iteraci za účelem signalizace emocí. Vyloučena byla tedy cizí slova (bez rozlišení, zda se jedná o slovo začleněné do české slovní zásoby, nebo vyskytující se v cizojazyčné větě), např. *jazz*,<sup>185</sup> *komm*, *off*, vlastní jména včetně přezdivek, zkratky, např. *jj* pro *jo jo*, *mm* pro milimetr, označení politické strany *ČSSD*, a slova, ve kterých se zdvojený grafém vyskytuje na morfematickém švu, např. *nejjasnější*, *povinnost*. Ve výsledcích jsme ponechali pouze případy, v nichž se multiplikovaný grafém objevuje navíc nad počet náležitých grafémů ve slově, tj. např. *bduu* nebo *prootže* jsme odstranili, byť obsahují iteraci, ale *buduu* nebo *prootože* jsme ponechali. Taktéž jsme odstranili slovo *jeiich*, protože se v něm nevyskytují všechna písmena z náležitého tvaru, kterým je *jejich*. V případě slov, kde se iterace objevila důsledkem chybějící mezery mezi slovy, jsme postupovali tak, jako by byla slova rozdělena správně.

V mluveném korpusu Ortofon byla iterace grafémů vyhledána stejným způsobem jako v psané elektronické komunikaci, ovšem hledání probíhalo ve fonetické transkripci, která zachycuje reálnou výslovnost. Zaměřili jsme se přitom pouze na konsonanty, protože změny vokalické kvantity jsou možné a známé u všech vokálů a jsou mj. ovlivněny nářečím.

### Analýza

Ve zkoumaných žánrech elektronické komunikace byla zjištěna iterace všech krátkých i dlouhých vokálů včetně *ů* a 9 konsonantů (tab. 28, 29). Častější iterace u vokálů podporuje pozorování Jílkové (2017), tedy fakt, že k zmnožení dochází u písmen/hlásek, které lze v řeči protáhnout. Pro ověření tohoto tvrzení jsme ještě využili fonetickou transkripci korpusu Ortofon. Mezi multiplikovanými konsonanty se sice objevily i okluzivy orální, tj. *p b t t' d d' k*,

---

<sup>184</sup> Dotaz vypadal následovně: [word=".+aa+.\*"]. Místo grafému *a* byla postupně dosazována další písmena.

<sup>185</sup> V řadě případů pomohl k identifikaci základového slova kontext, např. u slova *najazzlý* se jeho význam vyjeví až ve větě *u toho klasického najazzlého aranžmá s trubkou*.

avšak při poslechu se ukázalo, že se jedná o případy slov, v nichž je zdvojení hlásky ve výslovnosti vedeno zdvojením v pravopisu, např. *oddělení, jakkoliv, předtím*, nebo v případě *b p* o krátké protažení první fáze artikulace dané hlásky či celkově nezřetelnou výslovnost. U zbylých 16 konsonantů je jejich prodloužení v řeči reálně možné.<sup>186</sup> Tabulka 29 ukazuje, že v psané elektronické komunikaci byla nalezena pouze polovina konsonantů, které se v neformální mluvě prodlužují.

iterovaný vokál	frekvence iterovaného vokálu	počet lexémů s iterací
a	15	12
á	28	17
e	17	7
é	28	8
i	2	2
í	8	8
o	48	21
ó	4	3
u	12	7
ú	2	2
ů	6	4
y	3	3
ý	4	3
celkem	177	97

Tab. 28. Přehled vokálů obsahujících iteraci včetně počtu lexémů a jejich frekvence

iterovaný konsonant	frekvence iterovaného konsonantu	počet lexémů s iterací
f	4	2
h	10	4
l	7	6
m	23	4
r	9	4
s	6	6
š	3	2
ž	1	1
celkem	63	29

Tab. 29. Přehled konsonantů obsahujících iteraci včetně počtu lexémů a jejich frekvence

<sup>186</sup> Z hlediska způsobu tvoření se jedná o okluzivy nazální *m n ň*, semiokluzivy *c č*, konstriktivy *f v s z š ž h r ř l* a aproximantu *j* (terminologie podle Krčmové, 2017). Hlásky, u kterých bylo na fonetické rovině zaznamenána multiplikace ve smyslu prodloužení, jsme dále neanalyzovali, a to především z pocitu nekompetence pro takovou analýzu (domníváme se, že by bylo vhodné použít instrumentální analýzu řeči, doplněnou o percepční test provedený aspoň dvěma různými osobami). Problematika „nadstandardního“ prodloužení hlásky se také týká idiolektu a regionálních specifik.

Tabulka 28 dále ukazuje, že se iterace týká stejně často krátkých i dlouhých samohlásek. Iterace dlouhých vokálů se na první pohled může jevit paradoxem, protože znásobení vokálu samo o sobě značí jeho prodloužení. Nabízí se otázka, zda prodloužení již dlouhého vokálu by tedy mělo indikovat ještě větší prodloužení, nebo zda se jedná pouze o grafickou signalizaci délky, resp. trvání, bez rozdílu v tom, zda je iterován krátký, nebo dlouhý vokál. Z pohledu na data soudíme, že pravděpodobnější je druhá možnost. Příkladem mohou být lexémy *ale, hodně, jo, moc, ne, tak*, které se v datech vyskytují s multiplikovaným (různým počtem) krátkým i dlouhým vokálem (př. 63).

(63)

- Jo **taaak** , tak to že nejsi do něj zamilovaná , jsem asi přehlídla . (fóra)
- Mně se do toho **taaak** nechce , ale M . se zatím po jídle vůbec nevrhá . (fóra)
- Jenže bych pak byla **táák** uvolněná a bez stresu , že by moje pracovní morálka šla k šípku (diskuse)
- Jo , ty tábory by nebyly špatným řešením , však už se to také navrhovalo , ale všechno je **tááááááááááák** pomalé . . . (diskuse)

U dvojhlásek může být zmnožen jakýkoli z daných grafémů, příp. oba, např. *dlooooouuuhý, dloooooouho, dlooooouuuhý*. Mezi výskyty krátkých vokálů jsou zařazeny i případy psaní bez diakritiky, tento jev se však týká hlavně písmene i, jehož náležité zmnožení bylo nalezeno pouze jednou ve slově *kosiiiiik* (viz dále). Při iteraci také hrají roli snadnost zápisu grafému na klávesnici, na což poukazuje nízký výskyt grafému *ó* oproti ostatním dlouhým vokálům. V př. 64 jsou označena všechna slova s iterovaným grafémem. Absence grafémů *d'* (ve slově *ted'*) a *ň* (ve slově *mňá(á)m*) podporuje tuto domněnku.

(64)

Já taky miluju knedlíky zítra budu **míít** . . jinak s tím zelím kvůli mně se k plněným musí dělat velké hrnce , protože já ho taky miluju **mnáám** už se těším na zítra . . jinak jsem ted měla **pomeloo** mnám dneska mám vitamínků akorát si myslím (fóra)

Při posuzování funkce tohoto jevu se nelze vyhnout pojmu překlep. Přestože tabulky 28 a 29 ukazují počty lexémů, v nichž se vyskytuje alespoň o jeden grafém více, než je náležitý počet v daném slově, i zde jsou zahrnuty potenciální překlapy. Tak označujeme všechny případy, ve kterých se objevuje iterace právě dvou grafémů a které se vyskytují v kontextu s dalším překlepem (např. *Tohle bylo myšleno čistě na psrtnerý , jinal mám kplem sebe saamý úžasný lidi*). K hodnocení iterace jako překlepu se kloníme také v případech, v nichž nedokážeme iteraci z kontextu přisoudit význam, ačkoli nemůžeme vyloučit ani neznalost pisatele o správné

podobě slova (př. 65). Potenciální překlepy se týkají 53 % lexémů obsahujících iteraci, přičemž větší podíl zaujímá iterace konsonantů (srov. procentuální podíl překlepů ve slovech s iterovaným vokálem vs. konsonantem v tab. 30).

(65)

- Já v **těchhle** lamentech vždycky cítím touhu toho dárce setřít (fóra)
- Holky . . přeji klidný večer . . za chvíli přijede **manžžel** z rachoty . . (fóra)
- v usa či kanadě mají plyn - třeba ze **zássobníků** , **ellektrínu** mají všude kde je více obydlí (diskuse)

- když už jsme u toho Babiše tak jsem zjistila , že 80 tis . **sii** vyinkasuje ochranka za jednu směnu (diskuse)

	iterované vokály		iterované konsonanty	
	počet lexémů	frekvence výskytů	počet lexémů	frekvence výskytů
citoslovce	30,6 %	50,8 %	51,7 %	68,3 %
překlep	15,3 %	8,5 %	37,9 %	19 %
ostatní	54,1 %	40,7 %	10,4 %	12,7 %

Tab. 30. Podíl citoslovcí a překlepů na iteraci grafémů v korpusu Koditex-web (v procentech)

Lexémy neoznačené jako potenciální překlepy patří z 82 % k citoslovcím (tab. 30), např. *auuuu*, *bleee*, *booóže*, *hahaha*, *chjoóoo*, *jaaaj*, *pšššt*, *uff*, *vžummm*. K citoslovcím řadíme i responzní výrazy *ano* (v 1 variantě), *jo* (v 5 variantách), *ne* (ve 14 variantách), *no* (ve 3 variantách), které jsou příčinou většího podílu vokalické iterace na celkové frekvenci tohoto jevu.<sup>187</sup> U citoslovcí je funkce iterace jasná: prezentace emocí nebo simulace zvuku, včetně jeho intenzity či délky trvání. Totéž se dá říci i o ostatních slovech s multiplikovaným grafémem, která nebyla označena za překlepy. Interpretace iterace pak odráží také slovní druh či funkci, v jaké se daný lexém používá, např. zvolání *mamíííí*, *jedéém*, emotivní hodnocení *náádhera*, *špatnýý*, *úúúúzasně*. Příklad 66 výběrově ukazuje další funkce.

(66)

Nápodoba délky objektu:

- Jako že když budu mít kuchyni do L , do dvou rovnoběžek , do jedné **dlooooouhé** linky . . . tak se přenášení vyhnu ? (diskuse)

<sup>187</sup> Lexém *hm* hodnotíme ve všech variantách jako citoslovce vyjadřující váhání (na rozdíl od korpusu Ortofon, v němž *hmm* představuje kód pro souhlasnou responzi).

Nápodoba vzdálenosti:

- nekolikrát jsem zazila situaci , kdy jsem psy nechala u obchodu , privazane na kratko - a stojim u pokladny a vidim , jak se k malemu hrne nejaky prcek - klasicky 2 - 3 roky a rodic nekde **vzaaada** ...  
(fóra)

Nápodoba váhání:

- . . . . . a navíc je ten židama zaplaceney trouba tak hloupej , že to klidně pustí ven . . . . **ledažeeee** . . . .  
. . . . . á , plány v plánech a spiknutí ?  
(diskuse)

Zdůraznění:

- Nikdo není takový psychopat , aby si vyrobil falešné doklady a pokusil se převést tak obrovské sumy peněz - - - - - Tak to by jste se u Ransdorfa **móóóc** spletli .  
(diskuse)

- igelin: Přeju brzké uzdravení . Kojená miminka **netloustnouuuu** .  
(fóra)

- Proto já studovala a makala , abych ji a její dítě zajistila ? ? ? **Tooo** snad ne ku \* va .  
(Facebook)

## Shrnutí

V této studii jsme analyzovali jev označený jako iterace grafémů, který patří k tradovaným charakteristikám elektronické komunikace. Frekvenčně se v našich datech jedná o marginální jev (týkal se pouze 126 lexémů s frekvencí 240 výskytů, tj. 0,06 % všech lexémů/lemmat v korpusu Koditex), přesto si zaslouží pozornost, neboť se jedná o způsob, jakým je do vizuálního kanálu včleněna auditivní složka. Empiricky jsme zde ověřili tvrzení Jílkové (2017), že se iterace týká hlásek, jež lze v řeči prodloužit. V mluveném korpusu jsme našli doklady prodloužení 16 konsonantů, ale v elektronické komunikaci byla iterace zastoupena pouze u 8 z nich. Domníváme se, že je to zapříčiněno velikostí zkoumaných dat spíše než vlastnostmi zkoumaných žánrů.

Při analýze jsme se pokusili oddělit záměrnou a nezáměrnou iteraci, tj. označit překlepy. Přestože jsme o nich uvažovali pouze u duplikovaného grafému, za překlep jsme označili téměř polovinu lexémů s iterovaným konsonantem a 15 % lexémů s iterovaným vokálem. Výraznou skupinou slov obsahujících iteraci obou typů hlásek byla citoslovce, která tvořila 82 % všech lexémů. Z funkčního hlediska se iterace nejen u citoslovcí týkala nápodoby zvuku a zvukové realizace asociované s emocemi či typickým chováním, např. váháním. Prodloužení hlásky přitom může především u vokálů představovat i zvýšení hlasitosti, intenzity a metaforické zobrazení délky objektu či vzdálenosti od něj.

### 3.5.4 Shrnutí kapitoly

V této části jsme analyzovali tři jazykové rysy (diskurzí markery – DM, pronoun non-dropping, iteraci grafémů), jejichž výskyt je ovlivněn vnětetovými – pragmatickými – důvody. Ačkoli každý rys náleží jiné rovině (DM lexikální, pronoun non-dropping syntaktické, iterace grafémů grafické), důvody ovlivňující jejich výskyt jsou podobné, byť v případě iterovaných grafémů se často jedná o grafickou nápodobu či simulaci zvuku. Důvody můžeme rozdělit do tří skupin podle toho, k čemu se bezprostředně vztahují: 1) vzájemný kontakt mluvčího/pisatele s adresátem (např. souhlas, výzva adresáta k reakci), 2) modifikace komunikačního procesu (např. prodlužování segmentu za účelem získat čas), 3) modifikace textu (např. zdůraznění části textu, nejistota ohledně pravdivosti sdělení). Jedná se o vzájemně provázané skupiny (např. při modifikaci textu může zároveň dojít k modifikaci komunikačního procesu), které se uplatňují v různé míře v mluveném a psaném elektronickém jazyce.

Rozdíl mezi jazykovými mody se výrazně projevil u rysu pronoun non-dropping v jeho vyšší frekvenci v mluveném jazyce. U rysu DM se rozdíly mezi mody a žánry liší v závislosti na konkrétním DM. Pouze u slov *fakt* a *třeba* můžeme říci, že existuje výrazný rozdíl v zastoupení jedné a téže funkce těchto výrazů v mluveném jazyce oproti psanému elektronickému jazyku. Naopak u slova *jo* byly v obou komunikačních modech nalezeny stejné funkce, ale v mluveném jazyce bylo jejich zastoupení vyrovnanější než v psaných žánrech, ve kterých převažovala jedna funkce. U dalších dvou DM se vždy výrazněji odlišoval jeden žánr: u slova *víš* diskusní fóra, u kolokace *že jo* internetové diskuse. U posledního jevu – iterace grafémů – jsme mluvená data použili pro ověření, zda má tento jev oporu v mluveném jazyce.

### 3.6 Neverbální komunikace na příkladu smíchu

Za neverbální komunikaci (či neverbální složku komunikace) lze považovat v širokém slova smyslu vše, co může být chápáno jako významonosné, tj. čemu lze přisoudit význam, v mluveném dialogu tedy např. postoj komunikantů, jejich gesta a mimika, intonace, výška hlasu, v psaném textu např. velikost, styl a barva písma, členění do odstavců. Pro účely této práce jsme vybrali jeden neverbální prostředek, který se objevuje v analyzovaných datech mluveného i psaného modu, byť v různé míře. Jedná se o smích.

Smích představuje specifický druh mimického chování, který podle Glenna a Holtové (2016: 6) postrádá sémantiku a lingvistický obsah, avšak jehož různé formy (tj. způsoby, jak se člověk směje) mají komunikační efekt. Problematika toho, co lze za smích považovat, přesahuje rámec

této práce (více viz např. Potter & Hepburn, 2010; Glenn & Holt, 2016: 5n.), proto se omezíme na výběrové uvedení informací o tomto jevu.

Distinktivním rysem smíchu je v porovnání s řečí jeho sdílení mezi účastníky komunikace, byť ne vždy je smích souběžný (Jeffersonová et al., 1977). S vymezením smíchu souvisí i popis toho, co je úsměv. V naší práci postačí laické odlišení pomocí percepce: úsměv<sup>188</sup> je vnímán vizuálně, zatímco smích také akusticky.

Smích může plnit různé funkce, z nichž se nejčastěji uvádí signalizace humoru. Jeffersonová (1984) upozornila na další funkci, která se objevuje zejména v delikátních situacích, např. při konfliktu, řešení problému či v trapných chvílích: smích se v nich uplatňuje jako prostředek pro zdoání obtíží. Ačkoli se v principu lze smát čemukoli a kdykoli, hledání příčiny smíchu začíná vždy analýzou daného projevu – zda a kde se v něm vyskytuje něco, čemu se lze smát. Typicky se jedná o souběžné nebo bezprostředně následující umístění příčiny a smíchu samého (např. Glenn, 2003).

Smích lze lexikálně vyjádřit dvěma způsoby: pomocí lexémů, které přímo pojmenovávají toto chování (např. sloveso *smát se*), nebo pomocí lexikalizovaných citoslovcí obsahujících slabiky *ha/cha, he/che, hi/chi, ho/cho*. V psané elektronické komunikaci se však uplatňuje další způsoby záznamu – např. emotikony a akronymy. Jejich využití se však neomezuje pouze na smích. Smích je většinou zaznamenáván i při přepisu mluveného projevu, kromě dat určených pro jazykovou analýzu můžeme tyto záznamy najít např. také v novinách a časopisech v přepisech rozhovorů. V nich je smích a jeho typy zpravidla zaznamenán jako komentář v kulatých závorkách (srov. Komrsková & Škrabal, 2018), byť v některých periodikách je tento způsob v současné době nahrazen emotikonom :-).

V následujících dvou studiích se budeme zabývat záznamem smíchu. První studie se týká způsobu, jakým je smích zaznamenán v mluveném korpusu Ortofon. Ničím nezprostředkovaná mluvená komunikace je primárním prostředím smíchu, zatímco do psaného textu se smích dostává až sekundárně prostřednictvím různých konvencionalizovaných způsobů a jeho výskyt je ovlivněn vědomým rozhodnutím pisatele, které zahrnuje nejen způsob kódování (záznamu) smíchu, ale i jeho pozici v textu. Porovnání smíchu v mluvené a psané komunikaci by proto nebylo možné. Předmětem první studie je z tohoto důvodu snaha zjistit, jakou funkci má smích pro průběh komunikace a čím je jeho výskyt ovlivněn. Poznatky z pozorování smíchu

---

<sup>188</sup> Také u úsměvu je možno rozlišovat různé typy, např. Duchenneův úsměv představuje souběžný stah hlavního lícního svalu a kruhového očního svalu (více viz Duchenne, 1990).

v mluveném jazyce se pokusíme přenést do studie realizované na psaných datech. V ní se budeme zabývat záznamem smíchu v psaných elektronických textech pomocí emotikonů. Smích vyjádřený akronymy v této kapitole analyzovat nebudeme kvůli nedostatku dat.<sup>189</sup>

### 3.6.1 Smích v korpusu Ortofon

Korpus Ortofon je složen z nahrávek a jejich transkriptů. Přepisy zachycují kromě všech vyslovených slov také nejazykové projevy mluvčích a zvuky z okolí, které jsou jednak v nahrávce slyšet a jednak mohou ovlivňovat či modifikovat řeč. Jedná se například o nádechy, výdechy, zakašlání, zívání a hlavně smích. Tyto projevy jsou značeny na dvou samostatných rovinách podle toho, zda jsou produkovány společně s konkrétními slovy – pak jsou zachyceny do ortografického přepisu,<sup>190</sup> nebo ne – pak jsou pouze v tzv. metajazykové rovině<sup>191</sup> (obr. 5). Tyto dva způsoby značení se navzájem nevyklučují: modifikace hlasu smíchem je často patrná i po skončení samotného zasmání. Kromě toho může být hlas modifikován pouhým úsměvem, tj. protažením rtů bez zapojení hlasivek jako při smíchu (více viz Kohler, 2006). V korpusu Ortofon se takové detailní rozlišení nebere v úvahu, modifikace hlasu, které je percepčně přisouzen (spíše) úsměv či smích, jsou kódovány stejně.

Následující obrázek ukazuje situaci, ve které se oba mluvčí při rozhovoru smějí. Ačkoli je způsob zobrazení ve zveřejněném korpusu Ortofon jiný (viz s. 48), je obr. 5 lepší ukázkou pro to, čemu se budeme v této studii věnovat. V ortografických rovinách (na obrázku označených jako 0 ort, 1 ort) je modifikace hlasu smíchem či úsměvem označena kódem SM a všechna slova, u kterých je tato modifikace patrná, jsou daným kódem označena. V metajazykových rovinách (0 meta, 1 meta) jsou naopak vyznačeny úseky, ve kterých je slyšet pouze smích bez jakékoli slovní produkce. Smích může vzniknout i v pauze mezi slovy (viz stopu 1 ort a 1 meta v čase 4:31).

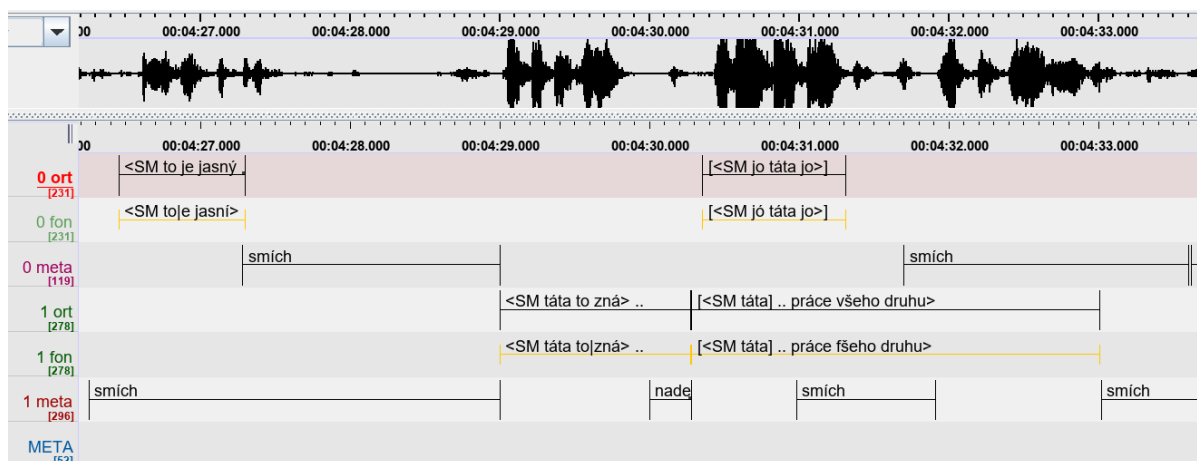
---

<sup>189</sup> K akronymům s významem smíchu se řadí např. *lol* (*laugh out loud/lot of laughing*), *lmao* (*laugh my ass off*), *rofl* (*rolling on floor laughing*). Ačkoli je *lol* uváděno jako jeden z nejčastějších akronymů v elektronické komunikaci (např. Tagliamonte & Denis, 2008), v korpusu Koditex-web se vyskytoval pouze dvakrát. Zbylé dva akronymy se v něm neobjevily vůbec.

<sup>190</sup> V anglickojazyčné literatuře se tento smích objevuje pod pojmy *smiled speech*, *smile voice*.

<sup>191</sup> V přepisech dostupných v Elanu jsou rozlišeny dvě metajazykové roviny: pro nejazykové projevy mluvčích a pro zvuky z okolí. Ve druhé skupině se objevuje značka smích více mluvčích najednou, kterou v této studii neanalyzujeme. V seznamu všech metajazykových projevů se objevuje na dvacátém místě z celkového počtu 75 různých projevů či jejich kombinací.





Obr. 5 Ukázka přepisu, ve kterém je značen smích (sonda 15C006N z korpusu Ortofon)  
 Na ortografické rovině patří označení „se smíchem“ za nejčastější modifikaci hlasu při řeči  
 v celém korpusu (tab. 31).<sup>192</sup> Za ním následuje imitace a přeroknutí. Značky se mohou vzájemně  
 kombinovat, avšak mezi deseti nejčastějšími se žádná kombinace nevyskytuje.

	značka v ortografické rovině	frekvence
1	se smíchem	4 325 (3 497,75 ipm)
2	imitace	603 (487,66 ipm)
3	přeroknutí	556 (449,65 ipm)
4	potichu	470 (380,1 ipm)
5	šeptem	463 (374,44 ipm)
6	s plnou pusou	357 (288,72 ipm)
7	cizojazyčný projev	262 (211,89 ipm)
8	čtený text	248 (200,56 ipm)
9	nahlas	198 (160,13 ipm)
10	s povzdechem	194 (156,89 ipm)

Tab. 31. Nejčastějších deset značek v korpusu Ortofon pro modifikaci hlasu při řeči

V metajazykové rovině se při přepisu rozlišuje smích, nebo pousmání podle toho, o jak výrazný a jak dlouhý projev se jedná. Rozlišování obou dvou typů není založeno na empiricky zjistitelných vlastnostech daného projevu, ale na individuálním posouzení přepisovatelů a

<sup>192</sup> Pro hledání byl použit následující dotaz: <para type=="."/ >.

kontrolorů korpusu Ortofon. Mezi metajazykovými projevy se smích vyskytuje jako třetí nejčastější (tab. 32), pousmání jako páté.<sup>193</sup>

	komentář v metajazykové rovině	frekvence
1	(nadechnutí)	10 517 (8 505,4 ipm)
2	(rušivý zvuk)	7 543 (6 100,24 ipm)
3	(smích)	5 371 (4 343,68 ipm)
4	(hluk v pozadí)	3 217 (2 601,68 ipm)
5	(pousmání)	2 913 (2 355,83 ipm)
6	(mlasknutí)	2 168 (1 753,32 ipm)
7	(cinkání nádobí)	1 878 (1 518,79 ipm)
8	(odkašláni)	1 026 (829,76 ipm)
9	(vydechnutí)	989 (799,83 ipm)
10	(citoslovce)	710 (574,2 ipm)

Tab. 32. Nejčastějších deset komentářů v metajazykové rovině korpusu Ortofon

V obou tabulkách je patrné, že smích (a pousmání) je v mluveném projevu výrazně frekventovaný. Z toho se dá usuzovat na jeho důležitost v neformální spontánní komunikaci. Frekvence záznamu smíchu z tabulek 31 a 32 dále ukazuje, že častěji je v datech zaznamenán smích na metajazykové rovině. Neznamená to však, že by se nahraní mluvčí smáli více bez mluvení než s mluvením; obr. 5 ukazuje, že před analýzou je nutné seznámit se se způsobem anotace takového jevu. Vlastní segment je na metajazykovou rovinu přidán vždy, když se mluvčí při smíchu odmlčí, ačkoli se dá říci, že se směje pořád, pouze u toho občas něco řekne. Frekvenční údaje je proto nutno brát s rezervou a při jakékoli analýze pracovat se značením smíchu na obou anotačních vrstvách.

Předmětem této studie není komplexní výzkum smíchu založený na datech korpusu Ortofon, ale ověření, zda se pouze ze zvukového záznamu nahrané situace dá odvodit příčina smíchu v daný moment a vliv smíchu na rozvíjení rozhovoru.

<sup>193</sup> K vyhledání byl použit následující dotaz: [pos="M"].

## Metodologie

Tato studie je založena na kvalitativním přístupu, avšak konkrétní nahrávka je pro následnou analýzu vybrána podle kvantitativního hlediska.

V korpusu Ortofon jsme vyhledali všechny záznamy smíchu na ortografické rovině<sup>194</sup>. Ze seznamu nahrávek s těmito záznamy jsme poté vybrali tu, v níž byla jejich frekvence nejvyšší a v níž zároveň vystupovali pouze dva mluvčí. Omezením počtu mluvčích v nahrávce jsme se snažili eliminovat nejasnosti, které by mohli nastat při zjišťování, kdo na koho reaguje ve větším počtu účastníků. V takto vybrané sondě (12A009N) jsme poté ručně anotovali všechny úseky, ve kterých se vyskytuje smích, a snažili se pojmenovat jeho příčinu.

Sonda 12A009N trvá 10 minut 57 vteřin a vystupuje v ní partnerská dvojice muže (Pavel\_35) a ženy (Alena\_36). Hovor se odehrává doma. Větší proporce řeči zaujímá v rozhovoru žena (55 %), ale modifikace řeči smíchem se objevuje častěji u muže (pokrývá 69 % všech těchto případů). Smích anotovaný na metajazykové rovině se také vyskytuje častěji u muže (zaujímá 84 % všech těchto případů).

## Analýza

V sondě 12A009N jsme označili 27 úseků, ve kterých se vyskytuje smích. Jedná se o minisituace ohraničené začátkem smíchu/slova vyřčeného se smíchem po doznění smíchu/slova vyřčeného se smíchem. Souhrnné údaje o každé minisituaci ukazuje následující tabulka.

	téma smíchu	kdo se směje	iniciátor smíchu	délka úseku se smíchem (v sekundách)
1	absurdní jídlo	oba	žena	9,06
2	kdo bude vařit	oba	muž	3,70
3	slovní obrana: zapomenutý počet a jména spoluhráčů	muž	muž	4,41
4	počet aut k odvozu z fotbalu	oba	žena	4,41
5	počet gólů	oba	muž	1,95
6	vzpomínka na fotbalový zápas	muž	muž	2,45
7	vzpomínka na spoluhráče	muž	muž	4,35
8	vzpomínka na cestu z fotbalu	muž	muž	1,74
9	výčitka: muž nedovezl dárek	oba	žena	4,00
10	nakupování, rozměňování hotovosti	oba	muž	12,13
11	vyprodaný obchod	oba	muž	5,42
12	telefon od kamaráda	muž	muž	3,38
13	kolikrát musí jet muž na dřevo	muž	žena	1,89

<sup>194</sup> Dotazem <para type="se smíchem"/>.

14	termín dokončení zásob dřeva na zimu	oba	muž	1,58
15	termín dokončení zásob dřeva na zimu	oba	žena	3,30
16	mužova přítomnost při řezání dříví	oba	muž	3,92
17	komu lze svěřit tajemství	oba	žena	4,63
18	svatba jako tzv. akce století	oba	žena	17,02
19	prosazování nového tématu – svatby	oba	žena	4,22
20	slovní obrana: termín rozlučky se svobodou	muž	muž	1,86
21	budoucí ženich	oba	žena	6,06
22	zaplacení svatby	oba	žena	6,60
23	jídlo u známých	muž	žena	2,88
24	dojem, který budí tento rozhovor	muž	žena	2,52
25	práce	oba	muž	11,54
26	představa bruslení ve větru	oba	žena	5,68
27	kdo je nefachčenko	oba	žena	5,53

Tab. 33. Souhrnné údaje o úsecích se smíchem v sondě 12A009N

Témata, k nimž se smích vztahuje, jsou v tabulce 33 popsána tak, aby si čtenář vytvořil představu o situaci. V zásadě se však jedná o několik tematických okruhů: jídlo (úseky 1, 2, 23, 24), fotbal (úseky 3–8), nakupování (úseky 9–11), dříví (úseky 12–16), svatba (úseky 17–22), bruslení (26), práce (úseky 25, 27).

Vznik smíchu v rozhovoru a jeho průběh ukazuje př. 67 (anotovaný úsek je v příkladu označen tučně). Smích je zapříčiněn návrhem Aleny\_36 dát si vajíčko. Pavel\_35 to zřejmě posuzuje jako vtip, protože se rozesměje a slovně na to reaguje (např. *samotné vajíčko, co su?*). Smích záhy odezní, Alena\_36 podává jiný, serióznější návrh na jídlo. Úsek, ve kterém zněl smích, trval cca 9 vteřin (zatímco předchozí část od otázky *Co budeme dneska vařit?* trvala 12 vteřin).

(67)

Pavel\_35: Ady co budem dneska vařit ?

Alena\_36: já nevím co bych si dala .. dal by sis maso ?

Pavel\_35: ne

Alena\_36: tak omáčku ?

Pavel\_35: emm

Alena\_36: vajíčko ?

Pavel\_35: <SM vajíčko ? .. samotné vajíčko budu jest nebo co ?>

Alena\_36: (smích)

Pavel\_35: (smích)

Pavel\_35: <SM co jako ? . co su ?>

Pavel\_35: (smích)

Alena\_36: tak třeba guláš bramborový

V př. 67 není příčina smíchu zahrnuta do úseku, ve kterém zaznívá smích. Většina anotovaných úseků byla právě taková, což souvisí s tím, že mluvčí se většinou v sondě nesmáli konkrétním slovům, ale tématu a situaci vykreslené slovy. Příklad 68 demonstruje situaci, ve které se smál pouze jeden mluvčí (ve všech 9 případech to byl muž). Takové úseky patřily v sondě k těm kratším (jejich medián je 2,52 vteřin, zatímco medián úseků, ve kterých se smáli oba mluvčí, je 5,02 vteřin). Za příčinu smíchu považujeme v př. 68 vzpomínku na situaci, o níž mluvčí vypráví a z níž reprodukuje slova, u kterých se nyní směje.

(68)

Pavel\_35: nee my jsme hnedka jeli protože Panika řídil a říká <SM pojďme jedem já mám žízeň> ne tak letěli hnedka

V úsecích, v nichž zaznívá smích jediného mluvčího, můžeme o ženině reakci (např. úsměvu) pouze spekulovat. Ve většině z 9 úseků lze vztáhnout absenci ženina smíchu k absenci sdíleného kontextu (úseky s mužovým vyprávěním o včerejším fotbalovém zápase, jehož se účastnil) či špatně podaného vtipu (např. úsek s hraným rozhořčením, že známý, s nímž se má sejít, nevolá). Ve dvou případech slouží smích muži jako prostředek pro překlenutí obtíží – při neschopnosti vzpomenout si a jako obrana před ženinou výčitkou. V jiných dvou případech stojí za mužovým smíchem vtipné prohlášení jeho ženy, ačkoli na jejím hlase není ve chvíli, kdy tak činí, ani poté slyšet náznak úsměvu (př. 69 – tučně je označen úsek, ve kterém zaznívá smích).

(69)

Alena\_36: že to bysme jeli asi zrovna z tama dají [nám aj] oběd ?

Pavel\_35: [tož ale] -

Pavel\_35: (smích)

Pavel\_35: <SM sami nemají myslím co jest> ..

Ve 21 úsecích se smějí oba mluvčí. Jedná se jak o situace, ve kterých se jeden mluvčí směje při mluvení a druhý se jen směje, často v překryvu s prvním mluvčím (př. 70), tak o situace, v nichž se oba mluvčí na chvíli odmlčí kvůli smíchu (př. 67).

(70)

Alena\_36: doufám že nám . u Janásů otevřou .. nějakou zimní zahrádku .. [<SM když teda do prosince>] .. a byl #s tam tedka ?

Pavel\_35: [(smích)]

Nejdelší (sedmnáctivteřinový) úsek rozhovoru, ve kterém se objevuje smích (př. 71), se týká tématu svatby. Oba mluvčí se podílejí na rozvíjení představy, jak bude událost vypadat. Smích v samostatném segmentu se prolíná s mluvou, překrývá se však i smích obou mluvčích.

(71)

Alena\_36: takže tam to bude aj s pokojema že si najme chůvu .. <NH to ti [bude akce století]>

Pavel\_35: [(smích)]

Pavel\_35: <SM no .. kdo ví koho tam Jan zas [vyklepe .. aby to nebyla nevěsta] kurva>

Alena\_36: [(smích)]

Pavel\_35: [(smích)]

Alena\_36: [<NH no to snad ne>] ..

Alena\_36: [(smích)]

Pavel\_35: [(smích)]

Alena\_36: i když u Jana se není <SM čemu divit> že ?

Alena\_36: (smích)

### Shrnutí

V této studii jsme představili způsob, jakým je v korpusu Ortofon anotován smích. Jeho rozdělení na úseky, ve kterých je slyšet pouze smích, nebo řeč modifikovaná smíchem, neumožňuje jednoduše vybrat tematické úseky, k nimž se smích vztahuje. Proto jsme pro analýzu vybrali jednu nahrávku a v ní tyto úseky ručně označili. Ačkoli se veškeré závěry vztahují pouze k jediné nahrávce, domníváme se, že některá pozorování lze formulovat obecně.

Jedním z těchto pozorování je skutečnost, že smích nevzniká náhodně. Jeho příčinu lze vysledovat v předchozí řeči komunikantů, příp. v prostředí, ve kterém se komunikace odehrává. Smích je v neformální spontánní komunikaci většinou sdílený. V analyzované nahrávce se mluvčí střídali v zahájení úseku se smíchem, tj. řekli něco, co vyvolalo smích. V tomto případě je nutno zmínit tři případy, ve kterých žena vyvolala mužův smích, aniž by se sama zasmála (nebo aniž by se smích projevil při její řeči). V případě muže jsme nic takového nezaznamenali. Není překvapivé, že délka úseků, v nichž byl zaznamenán pouze smích jednotlivce, byla v celkovém srovnání kratší než délka úseků se smíchem obou mluvčích.

V téměř jedenáctiminutové nahrávce se objevilo 27 úseků se smíchem, které celkově pokrývají cca 2 minuty. Tematická náplň těchto úseků se týkala běžného života, a to jak osob, tak činností. Ukázalo se, že pro vznik smíchu u obou mluvčích je důležité sdílení kontextu, resp. vzpomínky na určitou situaci. V úsecích vztahujících se k fotbalu byl slyšet pouze smích jednotlivce –

muže, který se fotbalového zápasu zúčastnil. Absenci slyšitelného smíchu druhé mluvčí můžeme v těchto úsecích přisoudit jak různému smyslu pro humor, tak špatnému způsobu podání. Nesmíme ani zapomínat na to, že mluvčí mohla reagovat úsměvem, což z analyzovaných dat nezjistíme.

O působení smíchu na průběh rozhovoru můžeme říci pouze to, že v analyzované nahrávce se jednalo o pozitivní vliv na rozvoj komunikace. Smích byl pravidelně rozložen v celé nahrávce a nikdy nedošlo kvůli výskytu smíchu k úplnému zastavení výměny replik. Smích byl pro rozvoj témat spíše podporou.

### 3.6.2 Emotikony

Emotikony jsou jedním z prostředků, které jsou v obecném povědomí výrazně spjaty s elektronickou komunikací probíhající v neformálním prostředí. Z tohoto prostředí se však rozšířily i do psané neelektronické komunikace. Vzhledem k jejich rozšířenosti a vysoké variantnosti jsou označovány různými termíny, kromě emotikonu smajlík nebo emoji. Buď lze tyto tři termíny považovat za synonyma, nebo je možné jimi rozlišovat různé podtypy jednoho grafického prostředku. V této práci budeme používat termíny emotikon a emoji následovně: v případě obecného, souhrnného označení pro dané grafické prostředky užíváme pojem emotikon. Termíny emotikon a emoji rozlišujeme tam, kde je nutné specifikovat, zda se jedná o řetězec typografických znaků, nebo piktogram. Emotikon chápeme jako takový řetězec tisknutelných znaků,<sup>195</sup> který představuje tvář vyjadřující emoce nebo určité chování spojené s emocemi, např. :-) (Komrsková, 2016). Pojmem emoji označujeme emotikon ve formě piktogramu, např. 😊 (McCulloch, 2019).

Kromě typografického kritéria je pro tuto práci relevantní ještě sémantické kritérium.<sup>196</sup> Podle něj rozlišujeme typy emotikonů podle významu, např. smích, smutek, zlost.<sup>197</sup> Za variantu emotikonu pak považujeme takové případy, ve kterých je jeden ze znaků symbolizující oči, nos či ústa nahrazen jiným, podobným znakem, aniž by došlo ke změně významu, např. :-) a :-] (Komrsková, 2019). Za varianty emotikonu považujeme rovněž případy, v nichž je multiplikován znak pro ústa, např. :-))), vynechán znak pro nos, např. :), příp. i pro oči, např.

---

<sup>195</sup> Přesněji ASCII znaků, tj. všech znaků napsatelných na klávesnici (písmena, číslice, interpunkční znaménka, matematické symboly apod.).

<sup>196</sup> Emotikony mohou být dále rozděleny např. podle kulturní oblasti, ze které pocházejí (západní – emotikony vs. východní – kaomoji), míry animace (statické vs. animované/dynamické) nebo zaměřitelnosti s gesty (McCulloch, 2019 uvádí tzv. emblematické emoji, které jsou grafickými ekvivalenty emblematických gest).

<sup>197</sup> Význam daného emotikonu byl vyhledáván v seznamech emotikonů na internetu, nejčastěji na Wikipedii ([https://en.wikipedia.org/wiki/List\\_of\\_emoticons](https://en.wikipedia.org/wiki/List_of_emoticons)).

))))). Z uvedených možností jsou plně ekvivalentní varianty s nosem a bez něj, u varianty s násobeným znakem pro ústa je tato multiplikace chápána jako vyšší intenzita či délka zobrazovaného chování, tj. :-D znamená smích, ale :-DDD představuje smích větší, déletrvajícím, bouřlivějším.

Používání emotikonů v elektronické komunikaci je odborníky dáváno do souvislosti se snahou pisatele včlenit do psaného textu svoje emoce a postoje úsporným a rychlým způsobem (např. Skovholt et al., 2014). Emotikony jsou chápány jako psané (grafické) ekvivalenty parajazykových a neverbálních prostředků při mluvené komunikaci (např. Rezabek & Cochenour, 1998; Provine et al., 2007). Podle Dresnerové a Herringové (2010: 250) indikují ilokuční sílu textu, k němuž jsou připojeny, a přispívají k jeho pragmatickému významu. Sekundárně mohou fungovat jako náhrada interpunkce a členicí prostředek (např. Markman & Oshima, 2007; Danesi, 2017; Pappert, 2017).

Výzkum emotikonů na angličtině (např. Park et al., 2014) i češtině (např. Komrsková, 2019) potvrzuje dominantní užívání emotikonu :-). U emoji je nejpoužívanějším 😊 (McCulloch, 2019: 182). S tím souvisí jejich využití v širokém významu jakýchkoli pozitivních emocí a s nimi spjatého chování – úsměvu či smíchu. Mimoto je tento emotikon používán jako indikátor ironie. V této studii se kromě obecného popisu typů a variant emotikonů ve třech žánrech elektronické komunikace budeme zabývat právě otázkou, do jaké míry lze funkce emotikonu :- ) odvodit z kontextu, v němž se vyskytuje.

## Metodologie

V korpusu Koditex-web jsou pro emotikony a emoji použity zvláštní tagy.<sup>198</sup> Značkování emotikonů probíhalo automaticky; podkladem byl seznam znaků a symbolů, které se mohou objevit na pozicích vnímaných jako oči, nos a ústa.<sup>199</sup> Z výsledků vyhledávání byly poté odfiltrovány sekvence znaků, které ve skutečnosti netvořily emotikon. Nejčastěji se jednalo o sekvence, ve kterých se objevovala písmena a číslice (např. *dx*, *XP*) nebo které byly součástí hypertextového odkazu (např. %). U emoji nebyla ruční korekce výsledků vyhledávání potřeba.

---

<sup>198</sup> Emotikony byly vyhledány dotazem [tag="Z#. \*"], emoji dotazem [tag="Z%. \*"].

<sup>199</sup> Seznam byl vytvořen autorkou této práce na základě excerptce jazykových publikací o emotikonech a seznamu emotikonů na Wikipedii ([https://en.wikipedia.org/wiki/List\\_of\\_emoticons](https://en.wikipedia.org/wiki/List_of_emoticons)).



	emotikony	emoji
diskuse	322 (1 336,91 ipm)	0
fóra	125 (515,41 ipm)	2 (8,25 ipm)
Facebook	1 202 (4 950,78 ipm)	89 (366,57 ipm)

Tab. 34. Frekvence emotikonů a emoji v korpusu Koditex-web

Tabulka 34 ukazuje disproporci v zastoupení emotikonů a emoji v žánrech elektronické komunikace. Vzhledem k způsobu výběru textů do subkorpusu Koditex-web (viz 3.1.1.1) považujeme tuto disproporci spíše za dílo náhody než za obraz užívání emotikonů a emoji v daných žánrech.

Součástí této studie je analýza funkcí emotikonu :-) a :-D, paralelně k tomu i emoji 😊. Pro tyto účely jsme v rámci internetových diskusí a diskusních fór pracovali se všemi výskyty daného emotikonu včetně jeho variant, ale u příspěvků z Facebooku jsme kvůli vysoké frekvenci pracovali s náhodným vzorkem o velikosti 200 výskytů.

### Analýza

V analyzovaných žánrech se celkem vyskytovalo 9 typů emotikonů, z nichž byl v příspěvcích z internetových diskusí a diskusních fór nejčastější emotikon :-), zatímco na Facebooku emotikon :-D (tab. 35). Při rozdělení na typy jsme považovali emotikony :-O a :-[] za varianty označující překvapení a emotikony ;- \ a ;-( za varianty smutného mrkání. V prvním případě se jedná o vizuální podobnost znaků symbolizujících otevřená ústa při překvapení. Ve druhém typu jsou sice pro ústa použity znaky, které odlišují dva typy negativních emotikonů (rozmrzelý a smutný), ale zároveň jsou oba znaky ve spojení s mrkáním; domníváme se tedy, že emotikony symbolizující ústa mají demonstrovat pouze negativní emoci, nikoli její míru.







	diskuse	fóra	Facebook
:-) (úsměv)	275 (85 %)	69 (56 %)	456 (38 %)
;-) (mrkání)	24 (7 %)	24 (19 %)	174 (14 %)
:-D (smích)	12 (4 %)	15 (12 %)	490 (41 %)
:-/ (rozmrzelost)	3 (1 %)	3 (2 %)	16 (1 %)
:-( (smutek)	5 (2 %)	9 (7 %)	44 (4 %)

:-P (vyplazený jazyk)	3 (1 %)	1 (1 %)	17 (1 %)
:-O :-[] (překvapení)	0	2 <sup>200</sup> (2 %)	4 <sup>201</sup> (> 1 %)
;- \ ;-( (smutné mrkání)	0	1 <sup>202</sup> (1 %)	1 <sup>203</sup> (> 1 %)
celkem	322	124	1202

Tab. 35. Frekvence emotikonů v korpusu Koditex-web

V největším počtu variant se objevil emotikon :-), a to ve variantách s vynechanýmnosem, multiplikovanými ústy, číslicí 8 symbolizující oči a s obráceným pořadím očí a úst, tj. (=). U variant .-) a :\_) pravděpodobně došlo vlivem rychlého psaní k překlepu, protože tečka a podtržítka nejsou standardními znaky objevujícími se na daných místech emotikonu.

U emoji bylo rozdělení do typů obtížnější, protože se nemusí lišit pouze grafická podoba zobrazeného chování (viz emoji pro polibek), ale v jedné variantě může být ještě něco přidáno (např. slzy smíchu u třetí varianty pro emoji smíchu) nebo změněno (viz rozdílné oči u emoji s vyplazeným jazykem). Podobně jako u emotikonů patřilo na Facebooku k nejčastějšímu zobrazení smíchu (tab. 36).<sup>204</sup> Na rozdíl od emotikonů nebyly mezi emoji pouze tváře, ale také tři různá gesta a piktogram znázorňující srdce (ve čtyřech variantách). Nalezená gesta patří do skupiny symbolických gest a jedná se o tyto: palec nahoru, ukazováček a prostředníček tvořící písmeno v, kroužek vytvořený spojením palce a ukazováčku (k jejich významu viz např. Kaderka, 2017).


	fóra	Facebook
 (smích)	1	30
 (vyplazený jazyk)	0	12
 (úsměv)	0	12
 (srdce)	1	12
 (d'ábel)	0	6
 (polibek)	0	4








<sup>200</sup> Každá varianta se objevila v jednom výskytu.

<sup>201</sup> Pouze emotikon :-O.

<sup>202</sup> Pouze emotikon ;-).

<sup>203</sup> Pouze emotikon ;-(.

<sup>204</sup> Dürscheidová a Sieverová (2017) ukázaly, že emoji  je nejčastějším emoji v chatové komunikaci přes aplikaci WhatsApp.

 (mrkání)	0	6
 (zamilovanost)	0	1
 (rozmrzlost)	0	1
 (vztek)	0	1
 (smutek)	0	1
	0	1
	0	1
	0	1

Tab. 36. Přehled emoji nalezených v korpusu Koditex-web

Další odlišností mezi nalezenými emotikony a emoji je rozložení mezi autory textu. Emoji byla použita pouze pěti uživateli Facebooku, přičemž 79 % výskytů emoji se objevilo u jedné autorky. Výskyt emotikonů na Facebooku byl naopak rozložen mezi 63 autorů a největší podíl emotikonů u jednoho autora tvořil 12 % všech výskytů.

Funkci emotikonů v textu není snadné určit, i když se jedná o emotikon vyjadřující jednoznačný výraz tváře nebo chování. Podobně je tomu ostatně s interpretací mimiky a chování při mluvené komunikaci tváří v tvář. Interpretaci slovního vyjádření usouvztažňujeme s neverbálním chováním i zvukovými prostředky, k tomu připojujeme znalosti o komunikační situaci a hovořící osobě. V psaných datech se však můžeme při interpretaci opřít pouze o slovní kontext. Výklad toho, k čemu daný emotikon v textu slouží, je proto třeba posuzovat jako záležitost subjektivní. Přesto předpokládáme, že se uživatel emotikonů nesnaží adresáta emotikony zmást, ale používá je ve shodě s územ, na který je zvyklý (a který my jako interpreti nemusíme dostatečně znát). Pro lepší představu rozebereme následující dvě ukázky obsahující emotikony a emoji.

(72)

Je to pro tebe hodně blbý , že k tématu nic nemáš , že jo ? Sžírá tě to ? Tak to je dobře :-)) Tak znovu - Zkus vyvrátit aspoň některé z argumentů , které jsem ti napsal a které jsi označil za lži . Když si mě namažeš na chleba , tak to zkus :-) Zatím tvoje snahy trapně končí u výkřiků a urážek . A doma pod polštářem brečíš , že nemáš , co bys k tématu napsal :-)))

(diskuse, discussions\_660)

(73)

Ahoj 😊 Takže jelikož na mém příběhu jsou jen 4 👍 😞 rozhodla jsem se že začnu psát ( znovu 😞 ) nový příběh , doufám že to těm 4 - 5 děvčatům nevdadí . . 😊 jmenoval by se : 🇵🇷 Para Siempre en mi

corazon ❤️ Už mám trochu vymyšlený děj , tak bych asi zítra přidala prolog 🙌 😊 \* Ruggelaria \* Ps : Promiňte že to dávám bez obrázku , ale je tu kámoška a tak sem nemohla na pc 😞 😞 ^ Don ' t be scared ^

(Facebook, facebook\_5686)

V ukázce 72 se objevují tři varianty emotikonu :-), každá obsahuje jiný počet znaků pro ústa. První výskyt doplňuje slovní vyjádření autorovy spokojenosti nad tím, že jiný diskutující už nemá čím přispět do diskuse. Druhý emotikon je umístěn k explicitní výzvě, aby onen diskutující zkusil rozbít autorovy argumenty. Třetí výskyt je připojen k urážce diskutujícího. První a třetí emotikon by se dal interpretovat jako výsměšný úsměv či zasmání (na něj odkazuje multiplikace znaku pro ústa) doprovázející dané slovní vyjádření. Druhý emotikon vyjadřuje zcela jiný typ úsměvu, má navodit jednak autorovu sebejistotu, jednak vážnost výzvy.

V ukázce 73 se vyskytuje jedenáct emoji, z toho se třikrát opakuje stejný emoji – usmívající se. První emoji doprovází pozdrav ve snaze zapůsobit pozitivně na adresáta od prvního slova (k této funkci viz Komrsková, 2016). Druhý emoji nahrazuje slovo lajk, tj. palec nahoru, zatímco třetí emoji přidává k první větě emoční náboj – autorka je velmi nespokojená s počtem lajků. Čtvrtý emoji, přidaný společně se slovem *znovu* jako komentář v závorce, vyjadřuje autorčin smutek, že musí začít nový příběh. Pátý emoji je opět snahou o pozitivní působení, tentokrát skrze omluvný úsměv čtenářkám příběhu, ve kterém se autorka rozhodla nepokračovat. Šestý a sedmý emoji rámuje název nového, milostného příběhu, zřejmě proto vybrala autorka emoji vztahující se k lásce. Osmý emoji přidává k předcházející větě informaci o tom, že prolog bude bezvadný, špičkový a devátým, usmívajícím se emoji se autorka snaží v této souvislosti opět působit pozitivně. Desátý emoji vyjadřuje rozmrzelost, nevíme však, jestli nad tím, že autorka nemohla přidat k příspěvku obrázek, nebo kvůli tomu, že autorka nemohla na počítač kvůli návštěvě kamarádky, příp. kvůli obému. Posledním, jedenáctým emoji se autorka loučí, posílá svým čtenářům polibek.

Stejně jako u obou ukázek jsme postupovali při určování funkcí emotikonů :-)) a :-D včetně jejich variant. Výskyty byly rozděleny do šesti skupin. První skupina představuje emotikony signalizující vtíp, resp. označující, že předchozí kontext je určen k pobavení, např. *No Liko , já původně chtěla napsat , že světu se narodil Spasitel u nás želvička :); Vzala jsem tedy 2 půlky lyží , hůlky a botu tam nechala a šla . Tedy spíš hopsala po jedné noze , protože sníh pěkně studil :-D*. Ve druhé skupině jsou zařazeny výskyty, v nichž emotikon vyjadřuje smích, např. *Ano , ano - na deníček soutěžícího v pojídání chili jsem si také vzpomněla , je to jedna z legráček , která už dlouho " běhá po internetu " :)))*. Rozdíl mezi prvními dvěma skupinami odráží

u emotikonu :-) rozdíl v multiplikaci znaku pro ústa. U emotikonu :-D se naopak multiplikace objevovala ojedinele, častější byla sekvence téhož emotikonu: *ha ha ha . . . se popadám za břicho ! ! ! a můžu se počůrat smíchy ! ! ! ha ha ha . . . :-D :-D :-D.*

Třetí skupina obsahuje emotikony interpretované jako výsměch, např. *Nenaletěl lhářům , je to trapný multinickový troll a blb . . . Až mu smažou tenhle kdovíkolikátý účet , udělá si nějaký ještě s trapnějším nickem a trapnějšími názory :-); Chudák Petr Mach , je vidět , že když se nemůže připravit předem na to co říkat , tak je totálně nahrenej :D.* Ve čtvrté skupině jsou zastoupeny emotikony indikující ironii, např. *Do parlamentu a vlády volby už vyhrají a toho se socialistická internacionála bojí jako čert kříže . Proto se dokáží spojit i s pravici , resp . s republikány . Ó jaká dojemná snaha o blaho Francie :-); Petra mi zkazila celý týden :D doufám , že příští týden bude líp.* Do páté skupiny patří výskyty doprovázející omluvu, prosbu či žádost, např. *Máte pravdu . Tituly sbírá na dlouhých . Trochu se mne to pomotalo :).*

Konečně do poslední, šesté skupiny jsme zařadili výskyty, ve kterých jsme nebyli schopni přesněji určit funkci, jakou emotikony v daném kontextu plní. Nepochybně se jedná o snahu zapůsobit pozitivně na adresáta, ovšem k tomu se připojují další funkce, které demonstrujeme v následujících příkladech. Pro tuto skupinu také platí, že se neobjevuje více týchž emotikonů za sebou.

(74)

- Dobré ráno holky jak se máte ? piju čaj a ležím u TV - pohoda , kterou mám ráda . Pokusím se pročíst , co jste napsaly :-)  
(fóra)

- ona se dcera nedá zlomit . . . já když vezmu světýlko tak mi pusou otevře , ale jakmile je cizí člověk na obzoru tak strne . . . musela by to být známá teta . . . :D dcera byla už dvakrát - je jí 20 měsíců . . . byla v roce a v roce a půl . . . :D pokaždé to dopadlo stejně . . . ona není kámoška s cizíma lidma :D

(Facebook)

V příspěvku z diskusních fór (př. 74) chápeme emotikon jako ekvivalent zdvořilého přátelského úsměvu. Podobně je laděn i text, který je pravděpodobně prvním vstupem autorky do diskuse. Autorka ostatním diskutujícím sděluje, že je online, ale ještě není zcela připravená ke komunikaci, nejprve se chce seznámit s předchozími diskusními příspěvky. Emotikonem tedy zároveň signalizuje prosbu o strpení.

Ve facebookovém příspěvku (př. 74) funguje emotikon :-D podobně. Autorka ho připojuje na konec vět, resp. replik v diskusním vláknu, týkajícím se chování malých dětí u lékaře. Zkušenost s vlastní dcerou a jejím chování při přítomnosti cizích lidí je určitě pro autorku (a adresáty) úsměvná, ovšem v textu se nevyskytuje žádné líčení vtipných situací. Emotikon :-

D proto interpretujeme jako snahu o zprostředkování pozitivního dojmu, možná omluvného úsměvu, kterým autorka komentuje chování své dcery.

(75)

To mělo být popíchnutí redakce . Jen jsem tam nedal smajlíky . :-) (diskuze)

Příklad 75 byl jediným svého druhu v analyzovaných vzorcích. Autor textu vysvětluje interpretaci svého předchozího příspěvku, do kterého nevložil emotikon, takže byl dalšími diskutujícími pochopen jinak. Za vysvětlení nyní již emotikon přidává. Můžeme ho tedy chápat jako ilustraci slova smajlík nebo jako pokus navodit v adresátech pozitivní dojem z celého procesu vysvětlování. Zároveň lze tento emotikon považovat za omluvu, tj. emotikon je ekvivalentem omluvného úsměvu.

(76)

- Nicméně já nejsem troll , protože tu píšu své autentické názory . A na to ta diskuze je , s tím se budeš muset smířit :-) Bohouši , můj příspěvek není otázka , ale tvrzení :-) (diskuze)

- PS : nejsem zaměstnanec , jsem svým pánem ty troubo :D (Facebook)

V příspěvku z internetových diskusí (př. 76) ukazujeme dva totožné emotikony v téže funkci. Text, který emotikonům předchází, obsahuje vždy něco, co je pro adresáta negativní nebo nezdvořilé. Připojením usmívajícího se emotikonu dochází ke zmírnění, oslabení tohoto negativního náboje. Autor textu se pravděpodobně snaží adresáta Bohouše uklidnit poté, co mu něco vytkl.

Ve stejné funkci se objevuje i emotikon :D ve facebookovém příspěvku (př. 76). Jeho přítomnost těsně vedle nadávky *ty troubo* působí jako snaha o zmírnění jejího účinku.

Tyto příklady ukazují složení šesté skupiny. Přestože do ní spadá nejvíce výskytů emotikonu :- ) (tab. 37), rádi bychom zdůraznili, že jednotlivé funkční skupiny nejsou zcela srovnatelné. Právě šestá skupina ukazuje problematičnost celé interpretace, neboť sdružuje – na rozdíl od ostatních skupin – těžko popsatelnou a interpretovatelnou funkci, resp. multifunkci. Přítomnost této skupiny i u emotikonu :-D indikuje vzájemné propojení obou pozitivních emotikonů. Je otázkou, nakolik je v daných kontextech relevantní určitý emotikon, zda nestačí vybrat nějaký s podobným významem, s obecně pozitivním účinkem.

funkce	diskuse		fóra		Facebook	
	: -)	: -D	: -)	: -D	: -)	: -D
vtip	38 % (26)	50 % (6)	31,6 % (87)	33,3 % (5)	12 % (24)	56 % (112)
smích	1 % (1)	16,7 % (2)	8 % (23)	60 % (9)	0,5 % (1)	16 % (32)
výsměch	4 % (3)	33,3 % (4)	19 % (51)	0	2 % (4)	9,5 % (19)
ironie	6 % (4)	0	5 % (15)	0	3,5 % (7)	5 % (10)
omluva, prosba, žádost	0	0	: -D	0	0	0
nespecifikované pozitivní působení (úsměv)	51 % (35)	0	36 % (98)	6,7 % (1)	82 % (164)	13,5 % (27)

Tab. 37. Zastoupení funkcí emotikonů :-) a :-D

Velmi často se ve vzorcích objevovaly emotikony signalizující vtip, tj. emotikonem bylo naznačeno, že text je určen pro zasmání. Na rozdíl od emotikonu :-) byl v textu, k němuž se pojil emotikon :-D, často přítomen ironický nádech. Další specifickou funkcí byl výsměch, který se u emotikonu :-) nejvíce projevil v příspěvcích z diskusních fór a u emotikonu :-D v internetových diskusích. Konečně funkce označená prostě jako smích (označující smích, kterému nejsme schopni přisoudit další specifikaci), kterou bychom očekávali jako majoritní u emotikonu :-D, byla sice výrazněji zastoupena právě u tohoto emotikonu (v porovnání s emotikonem :-)), ale nejčastější byla pouze v diskusních fórech.

Další skupiny byly v elektronické komunikaci zastoupeny okrajově, což může být dáno jednak velikostí vzorku (viz s. 136–137), jednak rozlišováním velmi specifických funkcí.

Stejné rozdělení do skupin jsme provedli u nejfrekventovanějšího emoji, tj. u 😊, 😄, 😂. Na rozdíl od obou blíže analyzovaných emotikonů se tento emoji objevoval většinou v sekvenci, např. *Body 1 + 2 taky neřeším - mám co na sebe i kam to dát . . . protože nosím převážně . . . montérky 😂 😄 😂*. Těmto sekvencím jsme nejčastěji přiřadili funkci smíchu nebo reakce na vtip (viz tab. 38). Zbylé tři funkce se vyskytovaly jednotlivě.

funkce	výskyty	sekvence
vtip	11	3
výsměch	5	1
ironie	3	1
smích	11	4
nespecifikované pozitivní působení	1	1

Tab. 38. Zastoupení funkcí u emoji 😊

Ve všech žánrech se v těsné blízkosti emotikonu vyskytovala koncová či pokračovací interpunkce<sup>205</sup> v méně než polovině výskytů. Nejčastějším případem bylo – opět ve všech třech žánrech – umístění emotikonu za koncovou interpunkcí (př. 77): v diskusích a diskusních fórech ve 35 %, na Facebooku v 29 % ze všech výskytů emotikonů v daném žánru. Opačné pořadí se vyskytovalo nejčastěji v internetových diskusích (př. 78), a to v 19 % případů.

(77)

Kdybyste si se stejnou pozorovností a nadšením vyslechla projev Kima , vyšlo by to nastejno . :)

(78)

Není třeba věřit rusům , stačí použít mozek a mapu :) .

Co se týče souvýskytu pokračovací interpunkce a emotikonu, jednalo s výjimkou Facebooku o marginální jev (v diskusích a diskusních fórech se jednalo o jednotky výskytů). Na rozdíl od koncové interpunkce se ta pokračovací objevila častěji za emotikonem (př. 79); na Facebooku byl tento případ zastoupen ve 13 %, zatímco opačné pořadí v méně než 0,2 %.

(79)

Tenkrát byl výběr velmi omezeny a tím pádem i podstatně jednodušší ;-), a ja si vybrala takovou tu placatou obdelnikovou v koženkovém provedení

Souvýskyt interpunkce a emoji v datech z Facebooku byl také zastoupen okrajově (v 19 % ze všech výskytů). Nejčastějším případem pak bylo umístění koncové interpunkce po emoji (př. 80).

(80)

Klid , pohoda a nakrmený pupek je to nejlepší , co může být , a nejen o vánocích 😊😊😊 .

### **Shrnutí studie**

V této studii jsme se zabývali popisem emotikonů a emoji v korpusu Koditex-web. Zjistili jsme, že emotikony jsou zde zastoupeny ve všech třech žánrech, zatímco emoji pouze ve dvou (v diskusních fórech zcela marginálně). Nejvíce se v datech objevovaly emotikony vyjadřující úsměv :-)) a smích :-D; u emoji byl nejčastějším 😊, také vyjadřující smích.

Kromě frekvenčního porovnání jsme se zaměřili i na analýzu pozice emotikonu v případě použití interpunkce. Mezi emotikony a emoji se u souvýskytu s koncovou interpunkcí projevila

---

<sup>205</sup> Za koncovou interpunkcí považujeme následující interpunkční znaménka (včetně jejich multiplikací a kombinací): . ! ? Pokračovací interpunkci představují , - :.



odlišná tendence: emotikon byl umístován za koncovou interpunkci, zatímco emoji před ní. Postavení emotikonu před interpunkci se ukázalo jako frekventované v případě pokračovací interpunkce, a to zejména v příspěvcích z internetových diskusí. I přes tato pozorování je ve všech žánrech nejčastější vynechání koncové interpunkce, pokud je na konci věty či celého textu použit emotikon – ten ji tak do značné míry supluje.

Součástí popisu emotikonů byla dále analýza funkcí emotikonů :-) a :-D a emoji 😊 včetně variant. Při analýze jsme identifikovali šest funkčních skupin (vtip, smích, výsměch, ironie, omluva/prosba/žádost, nespecifikované pozitivní působení), přestože je jednoznačné určení v řadě případů značně diskutabilní. To demonstruje skutečnost, že jsme většinu výskytů emotikonu :-) (bráno s ohledem na všechny tři žánry dohromady) přiřadili k poslední jmenované skupině, jež sdružuje všechny možné interpretace úsměvu. Interpretace tohoto emotikonu jako úsměvu, kterou uvádí různé seznamy emotikonů, se tedy zdá být na místě.

Podobně je tomu u emotikonu :-D a emoji 😄, jejichž význam je popisován jako smích. U obou byla funkce smíchu (včetně skupiny negativního smíchu – výsměchu) výrazně často zastoupena. Těmto dvěma funkčním skupinám konkurovala u všech tří analyzovaných jednotek skupina, v níž byl emotikon či emoji signálem vtipu. K detailnějšímu rozlišení funkcí či posouzení adekvátnosti dané interpretace v daném kontextu by bylo vhodné pracovat s jiným typem dat.

### 3.6.3 Shrnutí kapitoly

Tato kapitola se liší od předchozích v několika ohledech. Zatímco předchozí kapitoly sdružovaly několik případových studií, v nichž byl na základě mluvených i psaných elektronických dat podrobně analyzován a porovnáván jeden jazykový jev, tato kapitola je zaměřena pouze na jeden typ neverbálního chování – smích. Tímto vymezením se tedy podobá spíše případovým studiím z předchozích kapitol, ale i od těch se liší, a to zejména rozdělením analýzy smíchu v mluvené a psané komunikaci do separátních studií a rozdílným zaměřením analýzy v daném komunikačním modu.

V první studii této kapitoly jsme pozorovaly vznik a průběh smíchu v mluvené komunikaci. Zajímala nás zejména příčina smíchu a podíl mluvčích na jeho iniciaci. Druhá studie měla širší záběr: kromě polyfunkčnosti emotikonů jsme se v obecné rovině zabývali popisem všech typů emotikonů v datech a jejich souvýskytem s interpunkcí. U emotikonů :-) a :-D a emoji 😊, která byla v datech nejfrekventovanější, jsme pomocí kvalitativní analýzy ověřovali, zda

a v jaké míře se u nich vyskytuje význam, který je u nich uváděn v různých seznamech či slovnících.

U všech detailněji analyzovaných emotikonů se ukázalo, že jsou používány pro vyjádření smíchu. U emotikonu :-D a emoji 😄 se to může zdát jako samozřejmé, avšak pro emotikon :- ) je to překvapivé, neboť za jeho význam bývá považován úsměv. Emotikon :-) byl interpretován jako smích pouze v sekvenci (tj. :-) :-) atd.), byť představitelnou interpretací je zde i značení, že úsměv trvá delší dobu. Adekvátnost popisu emotikonu :-) jako úsměvu či signálu pro pozitivní pocity či chování byla v analýze také potvrzena.

Mezi funkcemi jsme dále odlišovali negativní smích (výsměch) od smíchu bez jiného sémantického příznaku. Tyto dva typy se vztahují pouze k těm výskytům emotikonů, které nelze považovat za reakci na vyjádření, jehož účelem je pobavit (signál vtipu). U emotikonu :-D, za jehož význam se uvádí smích, je právě tato funkce nejčtenější. Ačkoli se z našeho opatrného popisu může zdát, že tato funkce není v souladu s tradovaným významem emotikonu :-D, je tomu právě naopak.

Přítomnost smíchu i emotikonu vyjadřujícího smích je v konkrétním úseku textu možné považovat za jednoznačný signál, kterým jeho autor označuje adresátovi místo, které sám považuje za vtipné, směšné atd. Zdali bude adresát sdílet toto přesvědčení vedoucí ke smíchu (či minimálně úsměvu) závisí nejen na autorových verbálních schopnostech, ale i na míře sdíleného kontextu, s nímž se denotát pojí.

Závěrem se vraťme k otázce, která se objevila na začátku této kapitoly a od té doby zůstávala v pozadí: lze určité emotikony skutečně považovat za ekvivalenty smíchu? Domníváme se, že nejlépe na ni odpověděla právě funkční analýza vybraných emotikonů, byť se na první pohled mohla zdát zbytečná, protože ke každému emotikonu lze jednoduše na internetu nalézt jeho význam. Jak se ukazuje u všech frekventovaných jazykových prostředků, s rostoucí frekvencí dochází k diferenciaci funkcí, a tedy i významu. Emotikony nejsou výjimkou, přesto si ve většině užití podržují původní (tj. nejčastěji uváděný) význam či funkci. Smích (i úsměv) je v neformální spontánní komunikaci velmi častým neverbálním chováním, které nabývá v komunikaci různého významu včetně schopnosti suplovat verbální vyjádření (např. souhlas). Totéž platí pro emotikony, odpověď na otázku zní tedy kladně. Obrazně řečeno lze emotikony vyjadřující smích (a úsměv) brát jako výzvu, aby se adresát zasmál (pousmál apod.) na místě, které takto pisatel vyznačil.

## 4. Závěr

Tato disertační práce byla zaměřena na porovnání jazyka v neformálních mluvených interakcích s neformální psanou komunikací realizovanou v diskusních fórech, internetových diskusích a facebookových statusech. Mluvenou i psanou komunikaci přitom spojuje následující charakteristika: jedná se o dialogy, v nichž mluvčí/pisatelé participují dobrovolně, délka ani tematika není předem dána. Komunikační prostředí nemá napříč oblastmi zcela totožné vlastnosti, protože elektronická komunikace je vždy omezena konkrétními možnostmi komunikační platformy (např. regulace urážlivých či vulgárních textů, délka jednoho příspěvku), zatímco prostředí, ve kterém byly pořizovány mluvené dialogy, žádná taková omezení neměla. Pro oba typy dále platí, že mluvčí a pisatelé se v daném prostředí scházejí pravidelně, takže komunikace v něm není neobvyklá – naopak v elektronickém prostředí je hlavním účelem jeho existence. Pro zjednodušení jsme tyto mluvené a psané dialogy nazývali neformálními, čímž odkazujeme k chování mluvčích/pisatelů v popsaných podmínkách: participanti vědí, že nekomunikují v žádné oficiální situaci, ani po nich není vyžadována žádná příprava, aby se mohli zapojit, takže se mohou chovat přirozeně, reagovat a vyjadřovat se spontánně. Udávané charakteristiky se promítají do pojmu konceptuální mluvenost (Koch & Oesterreicher, 1985), dalo by se tedy říci, že se v této práci zabýváme konceptuálně mluvenými texty, jež se liší pouze mediálně.

Právě z uvedených charakteristik jsme vycházeli při srovnání mluvených a psaných elektronických textů. Vybrali jsme jejich čtyři společné aspekty (dialogičnost, spontánnost, naplňování pragmatických potřeb mluvčích/pisatelů, výskyt smíchu) a ty podrobili analýze na příkladech jazykových rysů, které jsou považovány za jejich typické projevy. Při výběru rysů jsme se inspirovali předchozími lingvistickými výzkumy zaměřenými na porovnávání textů i teoretickými pracemi o námi zkoumaných žánrech. Ověření, zda vybrané jazykové rysy skutečně reprezentují daný aspekt, bylo dalším cílem této práce.

Obecně lze říci, že mezi zkoumanými neformální mluvenými a psanými elektronickými texty jsou pouze malé rozdíly (viz dále). Tento závěr je nutno nazírat jako dílčí, protože jsme neworkovali s neformálními neelektronickými psanými texty, např. dopisy. Ucelený obraz o neformální komunikaci by samozřejmě bylo vhodné doplnit o srovnání s komunikací vykazující vyšší míru formálnosti. Zájemce o takovou komparaci odkazujeme do odborné literatury, zejména do publikací o MDA v češtině (Cvrček et al., 2018a, 2018b, v tisku), na něž jsme v této práci navázali. Na rozdíl od MDA nebylo našim cílem obecné srovnání textů

a popsání vlivů na registrovou variabilitu (v tom jsme naopak vyšli z předchozích výzkumů), ale hlubší analýza konkrétních jazykových rysů v daném žánru.

Vzájemné srovnání obou existenčních forem jazyka shrňme po jednotlivých oblastech analýzy. Rozdíly mezi mluvenými a psanými texty se projevují již na úrovni frekvence slovních druhů. Zatímco poměr sloves je v obou modech i v psaných žánrech téměř stejný, v dalších plnovýznamových slovních druzích se s výjimkou číslovek ukazují výrazné rozdíly. Zastoupení těchto slovních druhů rozděluje zkoumané žánry elektronické komunikace na dvě skupiny: diskusní fóra, jež se v tomto blíží neformálním mluveným textům, a internetové diskuse s facebookovými statusy, které se chovají stejně jako neelektronické psané texty. Pro mluvené texty a diskusní fóra platí nižší zastoupení podstatných a přídavných jmen společně s vyšším podílem zájmen a příslovcí. Mezi dvěma zbylými psanými žánry jsou podíly těchto slovních druhů odlišné pouze minimálně (viz tab. 1). U zbylých slovních druhů převládá rozdíl mezi mluveným a psaným modelem: nižší frekvenční podíl předložek v kombinaci s vyšším podílem spojek, částic a citoslovcí pro mluvený jazyk a přesně opačný poměr pro psaný. Vzhledem k tomu, že jsme dále nepracovali s jednotlivými slovními druhy, zůstali jsme pouze u těchto obecných konstatování.

Při porovnání nejčastějších 20 lemmat se opět projeví rozdíly mezi jazykovými módy. V mluvených textech jsou zastoupeny výrazy odkazující k dialogické povaze komunikace, např. responzní slova *hmm*, *jo*, *no*, zatímco u psaných textů se dá vystopovat důraz na specifikaci vztahů mezi slovy a větami, vyjádřený např. vztažným výrazem *kteřý* a širším repertoárem předložek než v mluveném jazyce. V tomto hrubém srovnání je však většina lemmat společná bez ohledu na komunikační modus a stupeň formálnosti.

Aspekt dialogičnosti jsme zkoumali prostřednictvím otázek, zájmen pro 2. osobu a nominálního oslovení. Uvedené rysy představují tři různé způsoby, kterými je při výměně komunikačních rolí v psaném i mluveném dialogu akcentován rozdíl mezi mluvčím/pisatelem a adresátem. Otázky a nominální oslovení představují explicitní výzvu mířenou na konkrétního adresáta, čímž dochází k výměně rolí, zatímco zájmena pro 2. osobu nejsou na první pohled přímou výzvou, přestože tak mnohdy fungují. Tato adresnost je s ohledem na frekvenci zkoumaných rysů nejvyšší v mluvených textech; v žánrech psané elektronické komunikace se frekvenčně nevyjevuje žádná tendence směřující k označení jednoho žánru za bližší k mluvenému jazyku. Funkční užití daných rysů poukazuje na to, že srovnáváme texty s jiným nastavením podmínek produkce (konkrétněji dále).

Největší funkční rozdíl mezi jazykovými módy ukazují zájmena pro 2. osobu. Pomineme-li nepřekvapivou odlišnost ve využití singulárních a plurálních forem, která odráží potenciální počet adresátů a vztah komunikujících subjektů (tj. častější užití plurálních forem v psaných textech je dáno tím, že adresátem je zpravidla cizí člověk, kterému se vyká, nebo je text určen skupině), jeví se žánrem nejbližším mluveným textům diskusní fóra: v obou se výrazně často objevují zájmena v singuláru. Tuto blízkost však nereflektuje funkční analýza: ve všech psaných žánrech se zájmena pro 2. osobu bez ohledu na číslo uplatňují především jako prostředek pro vyjádření názoru či jeho podpory, což poukazuje na diskusní charakter těchto žánrů. Naopak v mluvených textech převažuje využití v obratech, kterými se mluvčí vztahuje k adresátovi nebo kterými dává najevo své emoce.

Také u nominálního oslovení se projevuje odlišný vztah mezi mluvčím/pisatelem a adresátem. V psaných textech se častěji než v mluvených uplatňuje oslovení titulem či funkcí (včetně výrazu *pane/paní*) a výrazně větší je i počet různých propriet ve vokativu. Vyšší zastoupení vokativních forem u apelativ indikujících blízký vztah (okruh rodiny, přátel) poukazuje v mluvených textech rovněž na jinou tematiku, protože řada oslovení se vyskytuje v reprodukované řeči.

U posledního z dialogických rysů – otázek – se ukázalo, jakým způsobem vstupuje do analýzy způsob identifikace daného rysu (zde pomocí otazníku) a skutečnost, že pracujeme s mluveným jazykem ve formě transkriptu. Samostatné vyčlenění doplňovacích otázek upozornilo na širší využití otazníku v mluvených textech – jednalo se o případy, které se v psaných textech právě kvůli odlišnému jazykovému modu neobjevují nebo u kterých je zápis otazníku fakultativní, např. žádost o zopakování, vyjádření překvapení, výzva k reakci aj.

Aspekt spontánnosti byl zkoumán prostřednictvím dvou jazykových rysů, které lze chápat jako dysfluency na úrovni větné (opakování stejných slov) a úrovni slovní (přítomnost slovních fragmentů, překlápů a přeroknutí). Jejich výskyt nebyl v žádném jazykovém modu natolik výrazný, aby došlo k narušení souvislosti a plynulosti textu. Nalezení těchto jevů v obou jazykových módech skutečně prokazuje neformální charakter zkoumaných textů, v nichž není nutné trvat na bezchybné podobě výsledného textu. Stejně jako u předchozích tří rysů mají i rysy spontánnosti vyšší frekvenci v mluvených než v psaných textech. Pro psané texty bychom za výraznější indikátor spontánnosti označili fragmenty slov a překlápů, které jsou častější než opakování slov a které vykazují řadu shodných projevů v obou jazykových módech, např. nejčastějším druhem odchylky od standardní podoby slova byla záměna hlásky/písmena, která ze slova činí neexistující nebo zcela jiný lexém. Opakování slov se bez ohledu na jazykový

modus nejčastěji týká neplnovýznamových slov, tento rys se proto dá považovat za projev neformálnosti obecně. Nejvíce se mluveným textům podobaly facebookové příspěvky, v nichž se na rozdíl od zbylých psaných žánrů opakovala hlavně responzní a kontaktní slova. Rozdíl mezi mody se u tohoto jevu projevil v častějším opakování substantiv a celkově delších slov (měřeno počtem slabik).

Analýza dalších tří jazykových rysů (diskurzivních markerů, pronoun non-dropping, iterace grafémů) ukázala, že jsou tyto prostředky užívány v neformálních mluvených i psaných textech totožným způsobem pro přizpůsobení komunikace individuálním potřebám mluvčího/pisatele v aktuální situaci. Konkrétně se to týkalo kontaktu mluvčího/pisatele s adresátem, modifikace komunikačního procesu a modifikace textu. Dílčí výsledky u jednotlivých rysů opět ukazují rozdílnou frekvenci mezi jazykovými mody, avšak stejné funkční využití, např. u vybraných diskurzivních markerů (*fakt, jo, třeba, víš, že jo*) se liší pouze podílem konkrétních diskurzivněmarkerových funkcí, který může být ovlivněn velikostí zkoumaných vzorků. Výraznou shodu mezi mluveným a psaným jazykem ukazuje funkční a kontextová analýza rysu pronoun non-dropping. Takové tvrzení nemůžeme u iterace grafémů poskytnout, protože jsme její výskyt v mluveném jazyce neanalyzovali funkčně, avšak z dotčených lexémů a jejich kontextu jasně vyplývá, že iterace je v psaném textu motivována snahou znázornit zvuk, resp. výslovnost v konkrétním kontextu, pomocí níž mluvčí/pisatel znázorňuje např. délku objektu, vzdálenost, ale i váhání. Právě v tomto bodě by bylo vhodné výzkum v budoucnu rozšířit a zodpovědět otázku, které diskurzivněmarkerové funkce identifikované v mluveném jazyce jsou realizovány v psaném textu.

Konečně i poslední část analýzy, týkající se smíchu a emotikonů jakožto jeho projevu v psaných textech, potvrdila blízkost neformální mluvené a psané elektronické komunikace. Emotikon :-D si i přes postupné rozšiřování funkcí udržuje svůj status indikátoru smíchu. Naopak emotikon :-), jehož hlavním významem by měl být úsměv, je při své iteraci (zopakování) funkčním ekvivalentem emotikonu :-D. Míra, s jakou jsou oba pozitivní emotikony schopny reprezentovat smích, je však, uvažíme-li škálu pousmání – úsměv – zasmání – smích, jejich inherentní vlastností. Co se naopak jeví jako důležité, je skutečnost, že pisatelé cítí potřebu označit emotikonem adresátovi část textu, kterou považují za hodnou zasmání. V rámci žánrů elektronické komunikace se nejvíce emotikonů i emoji objevovalo ve facebookových příspěvcích.

Zkoumané jazykové rysy tedy ukázaly, že rozdíly mezi mluvenými a psanými elektronickými texty jsou většinou ve frekvenci rysu, nikoli v jeho funkci. Ačkoli považujeme frekvenci za

důležitou, zejména proto, že je často prvním a nápadným indikátorem rozdílu, teprve funkční analýza může tuto změnu řádně postihnout. Funkční využití rysu nám ještě sloužilo pro ověření, zda lze rys považovat za reprezentanta charakteristiky, k níž byl na základě odborné literatury přiřazen. Mluvená komunikace realizovaná mezi lidmi v blízkém vztahu a v soukromém prostředí se potvrdila jako primární zdroj jazykových projevů neformálnosti a internetové diskuse, diskusní fóra a facebookové příspěvky jako oblast psané komunikace, jíž je vlastní neformální jazykový projev odvozený ze způsobu, jakým pisatelé mluví. Náš výzkum u iterace grafémů a emotikonů zároveň prokazuje tvrzení Herringové (2012), že v psané elektronické komunikaci přebírá grafická rovina funkce prozódie.

Hlavním přínosem této práce je podle našeho názoru rozšíření a doplnění výzkumu mluveného jazyka a psané elektronické komunikace, pro které jsme využili kombinaci metod korpusové a interakční lingvistiky. Jak je z našich výsledků zřejmé, je při jazykovém bádání důležité specifikovat komunikaci s ohledem na míru formálnosti, neboť její projevy nejsou spjaty pouze s jazykovým modelem, ale s dalšími komunikačními charakteristikami. Kromě toho jsme se snažili ukázat na možnosti využití jazykových korpusů, včetně jejich výhod a nevýhod pro analýzu dialogičnosti, spontánnosti, pragmatických prostředků i neverbální komunikace. Jak už bylo zmíněno výše, srovnání mluvené a psané elektronické komunikace by bylo vhodné doplnit ještě o psanou neelektronickou komunikaci.

## Seznam použité literatury

- Adamec, P. (1962). K některým otázkám srovnávací syntaxe běžně mluvených projevů. In *Otázky slovanské syntaxe: Sborník brněnské syntaktické konference 17.–21. IV. 1961* (s. 339–341). Praha: SPN.
- Adamzik, K. (2016). *Textlinguistik: Grundlagen, Kontroversen, Perspektiven*. Berlin/Boston: De Gruyter.
- Adolphs, S., & Carter, R. (2013). *Spoken Corpus Linguistics: From monomodal to multimodal*. Routledge.
- Ágel, V., & Hennig, M. (2006). *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähetexten 1650-2000*. Tübingen: Niemeyer.
- Aijmer, K. (1996). *Conversational Routines in English*. London/New York: Longman.
- Aijmer, K. (2013). *Understanding Pragmatic Markers: A Variational Pragmatic Approach*. Edinburgh: Edinburgh University Press.
- Andersen, G. (2016). Semi-lexical features in corpus transcription: Consistency, comparability, standardisation. *International Journal of Corpus Linguistics*, 21(3), 323–347.
- Androutsopoulos, J. (2006). Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics*, 10(4), 419–438.
- Androutsopoulos, J. (2011). Language change and digital media: A review of conceptions and evidence. In N. Coupland & T. Kristiansen (Eds.), *Standard Languages and Language Standards in a Changing Europe*. Novus Press.
- Anolli, L., Ciceri, R., & Riva, G. (Eds.). (2002). *Say Not to Say: New Perspectives on Miscommunication*. Amsterdam: IOS Press.
- Auer, P. (2009). On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences*, 31, 1–13.
- Balhar, J. (2005). *Český jazykový atlas*. Praha: Academia.
- Balhar, J., & Jančák, P. (1992). *Český jazykový atlas*. Praha: Academia.
- Baron, N. S. (1984). Computer mediated communication as a force behind language change. *Visible Language*, 18(2), 118–141.
- Baron, N. S. (1998). Letters by phone or speech by other means: The linguistics of email. *Language and Communication*, 18, 133–170.
- Baron, N. S. (2004). See you online: Gender issues in college student use of instant messaging. *Journal of Language and Social Psychology*, 23, 397–423.
- Baron, N. S. (2010a). *Always On: Language in an Online and Mobile World*. Oxford: Oxford University Press.
- Baron, N. S. (2010b). Discourse structures in instant messaging: The case of utterance breaks. *Language@Internet*, 7, article 4.



- Benešová, L., Křen, M., & Waclawičová, M. (2015). Korpus spontánní mluvené češtiny ORAL2013. *Časopis pro moderní filologii*, 1, 42–50.
- Benko, V. (2015). *Srovnatelné webové korpusy Aranea*. ÚČNK FF UK. Dostupný z [www.korpus.cz](http://www.korpus.cz)
- Berger, T. (2002). Specifický poměr zdvořilosti a oslovení v češtině. *Čeština – univerzália a specifika IV*, 189–197.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Biber, D., & Egbert, J. (2016). Register variation on the searchable web: A multi-dimensional analysis. *Journal of English Linguistics*, 44(2), 95–137.
- Bieswanger, M. (2013). Micro-linguistic structural features of computer-mediated communication. In S. Herring, D. Stein, & T. Virtanen (Eds.), *Pragmatics of Computer-Mediated Communication* (s. 463–485). Berlin/Boston: De Gruyter.
- Blinka, L., Subrahmanyam, K., Šmahel, D., & Seganti, F. R. (2012). Differences in the teen blogosphere: Insights from a content analysis of English- and Czech- language weblogs. *Young*, 20(3), 277–29.
- Bogoczová, I., Fic, K., Chloupek, J., Jandová, E., Krčmová, M., & Müllerová, O. (2000). *Tváře češtiny*. Ostrava: Ostravská univerzita v Ostravě.
- Brazil, D. (1995). *A Grammar of Speech*. Oxford: Oxford University Press.
- Brinton, L. J. (1996). *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/New York: Mouton de Gruyter.
- Burgstahler, S., & Doyle, A. (2005). Gender differences in computer-mediated communication among adolescents with disabilities: A case study. *Disability Studies Quarterly*, 25(2). Dostupné z <http://dsq-sds.org/article/view/552/729>
- Clancy, B. (2016). *Investigating Intimate Discourse*. Milton Park/New York: Routledge.
- Coseriu, E. (n.d.). *Sprachtheorie und allgemeine Sprachwissenschaft*. München: Fink.
- Coseriu, E. (1971). *Sprache. Strukturen und Funktionen. 12 Aufträge zur allgemeinen und romanischen Sprachwissenschaft*. Tübingen: Narr.
- Cvrček, V. (2017). Zipfovy zákony. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/ZIPFOVY%20Z%C3%81KONY>
- Cvrček, V., & et al. (2015). *Mluvnice současné češtiny I* (2. vyd.). Praha: Karolinum.
- Cvrček, V., Komrsková, Z., Lukeš, D., Poukarová, P., Řehořková, A., & Zasina, A. J. (2018a). From extra- to intratextual characteristics: Charting the space of variation in Czech

- through MDA. *Corpus Linguistics and Linguistic Theory*. Dostupné z <https://doi.org/10.1515/cllt-2018-0020>
- Cvrček, V., Komrsková, Z., Lukeš, D., Poukarová, P., Řehořková, A., & Zasina, A. J. (2018b). Variabilita češtiny: Multidimenzionální analýza. *Slovo a slovesnost*, 79(4), 293–321.
- Cvrček, V., Laubeová, Z., Lukeš, D., Poukarová, P., Řehořková, A., & Zasina, A. J. (v tisku). *Registry v češtině*.
- Čech, R. (2016). *Tematická koncentrace textu v češtině*. Praha: ÚFAL MFF UK. Dostupné z [http://ufal.mff.cuni.cz/books/preview/2016-cech\\_preview.pdf](http://ufal.mff.cuni.cz/books/preview/2016-cech_preview.pdf)
- Čechová, M., Krčmová, M., & Minářová, E. (2008). *Současná stylistika*. Praha: NLN.
- Čermák, F. (1993). Spoken Czech. In E. Eckert (Ed.), *Varieties of Czech. Studies in Czech Sociolinguistics* (s. 27–41). Amsterdam/Atlanta: Rodopi.
- Čermák, F. (2009). Spoken corpora design: Their constitutive parameters. *International Journal of Corpus Linguistics*, 14(1), 113–123.
- Čermák, F., & et al. (2007). *Frekvenční slovník mluvené češtiny*. Praha: Karolinum.
- Čermák, F., & Křen, M. (2004). *Frekvenční slovník češtiny*. Praha: NLN.
- Čermáková, A., Jílková, L., Komrsková, Z., Kopřivová, M., & Poukarová, P. (2019). Diskurzivní markery. In J. Hoffmannová, J. Homoláč, & K. Mrázková (Eds.), *Syntax mluvené češtiny* (s. 244–351). Praha: Academia.
- Čermáková, A., & Kopřivová, M. (2018). Korpusový výzkum mluveného jazyka na příkladu češtiny a angličtiny: Současný stav. *Slovo a slovesnost*, 79(3), 217–240.
- Čmejrková, S. (1997). Čeština v síti: Psnost či mluvenost? (O stylu e-mailového dialogu). *Naše řeč*, 80(5), 225–247.
- Čmejrková, S. (2008). Jazyk médií a jeho konverzacionalizace. *Jazykovědné aktuality*, 45, 87–100.
- Čmejrková, S., & Hoffmannová, J. (Eds.). (2011). *Mluvená čeština: Hledání funkčního rozpětí*. Praha: Academia.
- Čmejrková, S., Hoffmannová, J., Müllerová, O., Havlík, M., & Zeman, J. (2013). *Styl mediálních dialogů*. Praha: Academia.
- Danesi, M. (2017). *The Semiotics of Emoji. The Rise of Visual Language in the Age of the Internet*. London/New York: Bloomsbury.
- Daneš, F. (1949). Intonace otázky (Kapitola z větší práce). *Naše řeč*, 33(3–4), 62–68.
- Daneš, F. (1988). *Dynamika současné češtiny z hlediska lingvistické teorie a školské praxe*. Praha: Univerzita Karlova.
- Daneš, F. (1999a). Dialektické tendence ve vývoji spisovných jazyků. Příspěvek sociolingvistický. In *Jazyk a text. Výbor z lingvistického díla Františka Daneše I* (s. 249–263). Praha: FF UK.

- Daneš, F. (1999b). Pojem „spisovného jazyka“ v dnešních společenských podmínkách. In *Jazyk a text. Výbor z lingvistického díla Františka Daneše I, část 2* (s. 289–296). Praha: FF UK.
- Derks, D., Bos, A. E. R., & von Grumbkow, J. (2007). Emoticons and social interaction on the Internet: The importance of social context. *Computers in Human Behavior*, 23, 842–849.
- Deumert, A. (2014). *Sociolinguistics and Mobile Communication*. Edinburgh: Edinburgh University Press.
- Dresner, E., & Herring, S. (2010). Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory*, 20, 249–268.
- Duchenne, G. (1990). *The Mechanism of Human Facial Expression*. Originální vydání: *Mecanisme de la Physionomie Humaine*, (1862). (R. Andrew, Přel.). Cambridge: Cambridge University Press.
- Dürscheid, C. (2003). Medienkommunikation im Kontinuum von Mündlichkeit und Schriftlichkeit. Theoretische und empirische Probleme. *Zeitschrift für angewandte Linguistik*, 38, 37–56.
- Dürscheid, C. (2016). Nähe, Distanz und neue Medien. In H. Feilke & M. Hennig (Eds.), *Zur Karriere von ‚Nähe und Distanz‘. Rezeption und Diskussion des Koch-Oesterreicher-Modells* (s. 357–385). Berlin/Boston: de Gruyter.
- Dürscheid, C., & Siever, C. M. (2017). Jenseits des Alphabets—Kommunikation mit Emojis. *Zeitschrift für germanistische Linguistik*, 45(2), 256–285.
- Eimler, S. C., Ganster, T., & Krämer, N. C. (2012). Als Chef hat man nichts zu lachen: Eine Studie zur Wirkung von durch Führungspersonen genutzten Emoticons in berufsbezogenen Emails. In T. Köhler & N. Kahnwald (Eds.), *GeNeMe '12* (s. 145–153). Dresden: Technische Universität Dresden.
- Emonds, J., Newson, M., & Veselovská, L. (2017). Teorie řízení a vázání. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/TEORIE%20%C5%98%C3%8DZEN%C3%8D%20A%20V%C3%81Z%C3%81N%C3%8D>
- Erjavec, T., Cibej, J., Holdt, Š. A., Ljubešic, N., & Fišer, D. (2016). Gold-standard datasets for annotation of Slovene computer-mediated communication. In A. Horák, P. Rychlý, & A. Rambousek (Eds.), *Proceedings of Recent Advances in Slavonic NaturalLanguage Processing, RASLAN 2016*, (s. 29–40). Brno: Tribun EU. Dostupné z [https://nlp.fi.muni.cz/raslan/2016/paper06-Erjavec\\_etal.pdf](https://nlp.fi.muni.cz/raslan/2016/paper06-Erjavec_etal.pdf)
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 1, 29–62.
- Fairclough, N. (1992). *Discourse and Social Change* (Roč. 1992). Cambridge: Polity Press.
- Fic, K. (2000). K charakteristice lexikálního výraziva v dnešních rozhovorech. In I. Bogoczová, K. Fic, J. Chloupek, E. Jandová, M. Krčmová, & O. Müllerová, *Tváře češtiny* (s. 91–102). Ostrava: Ostravská univerzita v Ostravě.

- Filipec, J., Daneš, F., Machač, J., & Mejstřík, V. (1994). *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia.
- Fischer, O. (2011). Cognitive iconic grounding of reduplication in language. In P. Michelucci, O. Fischer, & C. Ljungberg (Eds.), *Semblance and Signification* (s. 55–82). Amsterdam/Philadelphia: John Benjamins.
- Fox, A. B., Bukatko, D., Hallahan, M., & Crawford, M. (2007). The medium makes a difference: gender similarities and differences in instant messaging. *Journal of Language and Social Psychology*, 26, 389–397.
- Fraser, B. (2006). Towards a theory of discourse markers. In K. Fischer (Ed.), *Approaches to Discourse Particles* (s. 1–20). Amsterdam: Elsevier.
- Giles, D., Stommel, W., Paulus, T., Lester, J., & Reed, D. (2015). The Microanalysis of Online Data: The Methodological Development of „Digital CA“. *Discourse, Context & Media*, 45–51.
- Glenn, P. (2003). *Laughter in Interaction*. Cambridge: Cambridge University Press.
- Glenn, P., & Holt, E. (2016). Introduction. In P. Glenn & E. Holt (Eds.), *Studies of Laughter in Interaction* (s. 1–22). London: Bloomsbury Academic.
- Goláňová, H. (2015). A new dialect corpus: DIALEKT. In K. Gajdošová & A. Žáková (Eds.), *Proceedings of the Eight International Conference Slovko 2015 (Natural Language Processing, Corpus Linguistics, Lexicography)* (s. 36–44). Lüdenscheid: RAM-Verlag.
- Grepl, M. (1962). K podstatě a povaze rozdílů mezi projevy mluvenými a psanými. In *Otázky slovanské syntaxe: Sborník brněnské syntaktické konference 17.-21.IV.1961* (s. 342–345). Praha: SPN.
- Grepl, M., & Karlík, P. (1998). *Skladba češtiny*. Olomouc: Votobia.
- Gromnica, R. (2017). Projevy vulgarity a zdvořilosti v diskusních fórech k domácí politice. *Časopis pro moderní filologii*, 99(2), 214–224.
- Gumperz, J. (1972). Introduction. In J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication* (s. 1–25). New York: Holt, Rhinehart & Winston.
- Haller, J. (1933). Spisovná čeština a jazyková kultura II. *Naše řeč*, 17(3), 77–87.
- Hansen, B. (2017). Gramatikalizace. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/GRAMATIKALIZACE>
- Hård af Segerstad, Y. (2002). *Use and Adaptation of Written Language to the Conditions of Computer-Mediated Communication* [Disertační práce]. Department of Linguistics University of Gothenburg.
- Hausenblas, K. (1962). O studiu syntaxe běžně mluvených projevů. In *Otázky slovanské syntaxe: Sborník brněnské syntaktické konference 17.–21. IV. 1961* (s. 313–323). Praha: SPN.

- Hausenblas, K. (1971). *Výstavba jazykových projevů a styl*. Praha: Univerzita Karlova.
- Hausenblas, K. (1984). Text, komunikáty a jejich komplexy (Zamyšlení pojmoslovné). *Slovo a slovesnost*, 45, 1–7.
- Havlík, M. (2017). Interakční lingvistika. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/INTERAK%C4%8CN%C3%8D%20LINGVISTIKA>
- Havránek, B. (1932). Úkoly spisovného jazyka a jeho kultura. In B. Havránek & M. Weingart (Eds.), *Spisovná čeština a jazyková kultura* (s. 32–84). Praha: Melantrich.
- Havránek, B., Bělič, J., Helcl, M., & Jedlička, A. (Eds.). (1960). *Slovník spisovného jazyka českého*. Praha: Academia.
- Hennig, M. (2006). *Grammatik der gesprochenen Sprache in Theorie und Praxis*. Kassel: Kassel University Press.
- Hepburn, A., & Bolden, G. (2013). The conversation analytic approach to transcription. In J. M. Atkins & J. Heritage (Eds.), *The Handbook of Conversation Analysis* (s. 57–76). Chichester: Wiley-Blackwell.
- Herring, S. (2000). Gender differences in CMC: Findings and implications. *CPSR Newsletter*, 18(1). Dostupné z [https://www.researchgate.net/publication/246291970\\_Gender\\_Differences\\_in\\_CMC\\_Findings\\_and\\_Implications](https://www.researchgate.net/publication/246291970_Gender_Differences_in_CMC_Findings_and_Implications)
- Herring, S. (2007). A faceted classification scheme for computer-mediated discourse. *Language@Internet*, 4(article 1).
- Herring, S. (2008). Language and the Internet. In W. Donsbach (Ed.), *International Encyclopedia of Communication* (s. 2640–2645). Blackwell Publishers. Dostupné z <https://ella.sice.indiana.edu/~herring/lg.inet.pdf>
- Herring, S. (2011). Computer-mediated conversation: Introduction and overview. *Language@Internet*, 7(2).
- Herring, S. (2012). Grammar and electronic communication. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Dostupné z <http://info.ils.indiana.edu/~herring/e-grammar.pdf>
- Heyd, T. (2015). Digital genres and process of remediation. In A. Georgakopoulou & T. Spilioti (Eds.), *The Routledge Handbook of Language and Digital Communication* (s. 87–102). Taylor & Francis Ltd.
- Hirschová, M. (1984). K jednomu typu výpovědi běžně mluvené češtiny. *Slovo a slovesnost*, 45(4), 273–276.
- Hladká, Z., & et al. (2005). *Čeština v současné soukromé korespondenci. Dopisy, e-maily, SMS*. Brno: Masarykova univerzita.

- Hladká, Z., & Šmilauer, I. (2006). K specifickým rysům elektronické korespondence. In T. Hoskovec, O. Ševčík, & R. Sova (Eds.), *Teorie a empirie. Bichla pro Krčmovó* (s. 151–162). Brno: Masarykova univerzita.
- Hoffmannová, J. (1997). *Stylistika a ...* Praha: Trizonia.
- Hoffmannová, J. (2006). Intertextualita v současné elektronické komunikaci. In T. Hoskovec, O. Ševčík, & R. Sova (Eds.), *Teorie a empirie. Bichla pro Krčmovó* (s. 163–174). Brno: Masarykova univerzita.
- Hoffmannová, J. (2008). Jazyková neformálnost a „kolokvializace“ v českých médiích. *Jazykovědné aktuality*, 45, 100–111.
- Hoffmannová, J. (2011). Mluvená čeština v zrcadle „psané konverzace“ na chatu. In S. Čmejrková & J. Hoffmannová (Eds.), *Mluvená čeština: Hledání funkčního rozpětí* (s. 393–408). Praha: Academia.
- Hoffmannová, J. (2016a). Internetová čeština: Ekonomie a redukce vyjadřování, explicitnost a implicitnost. In B. Tošovic & A. Wonisch (Eds.), *Interaktion von Internet und Stilistik, Internet und Stil* (s. 55–63). Graz: Institut für Slavistik der Karls-Franz-Universität Graz.
- Hoffmannová, J. (2016b). Každodenní mluvené dialogy. In J. Hoffmannová, J. Homoláč, L. Jílková, K. Mrázková, E. Chvalovská, P. Kaderka, & P. Mareš, *Stylistika mluvené a psané češtiny* (s. 30–78). Praha: Academia.
- Hoffmannová, J. (2017a). Konverzace. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/KONVERZACE>
- Hoffmannová, J. (2017b). Projevy mluvené a psané. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z [https://www.czechency.org/slovník/PROJEVY\\_MLUVENÉ\\_A\\_PSANÉ](https://www.czechency.org/slovník/PROJEVY_MLUVENÉ_A_PSANÉ)
- Hoffmannová, J., & Havlík, M. (2017). Hledání jednotky mluveného projevu: Replika, výpověď, promluva. In J. Chocholová, L. Molnár Satinská, & G. Múcsková (Eds.), *Jazyk a jazykoveda v pohybe II. Na počesť Slavomíra Ondrejoviča* (s. 433–441). Bratislava: Veda.
- Hoffmannová, J., Homoláč, J., Jílková, L., Mrázková, K., Chvalovská, E., Kaderka, P., & Mareš, P. (2016). *Stylistika mluvené a psané češtiny*. Praha: Academia.
- Hoffmannová, J., Homoláč, J., & Mrázková, K. (Eds.). (2019). *Syntax mluvené češtiny*. Praha: Academia.
- Hoffmannová, J., Komrsková, Z., & Poukarová, P. (2017). Reprodukce řeči/myšlení v mluvených projevech jako předmět korpusového výzkumu. *Korpus – gramatika – axiologie*, 15, 3–25.
- Hoffmannová, J., & Müllerová, O. (2000). *Jak vedeme dialog s institucemi*. Praha: Academia.
- Hoffmannová, J., & Müllerová, O. (Eds.). (2007). *Čeština v dialogu generací*. Praha: Academia.

- Hoffmannová, J., & Richterová, O. (2015). Si představ, by mě zajímalo, ti to ukážu: Jednoslabičné začátky syntaktických segmentů v mluvené češtině. *Časopis pro moderní filologii*, 97(1), 9–20.
- Hoffmannová, J., & Zeman, J. (2017). Výzkum syntaxe mluvené češtiny: Vstupní inventarizace problémů. *Slovo a slovesnost*, 78, 45–66.
- Homoláč, J., & Mrázková, K. (2014). K stylistickému hodnocení jazykových prostředků, zvláště lexikálních. *Slovo a slovesnost*, 75(1), 3–38.
- Huffaker, D., & Calvert, S. (2005). Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication*, 10(2). Dostupné z <http://dx.doi.org/10.1111/j.1083-6101.2005.tb00238.x>
- Hunston, S., & Francis, G. (2000). *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.
- Hymes, D. (1972). Models of the interaction of language and social life. In J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication* (s. 35–71). New York: Holt, Rhinehart & Winston.
- Chafe, W., & Danielowicz, J. (1987). Properties of spoken and written language. In M. Ghadessy (Ed.), *Thematic Development in English Text* (s. 83–113). London: Pinter Publisher.
- Chloupek, J. (1962). O výzkumu nářeční syntaxe v terénu. In *Otázky slovanské syntaxe: Sborník brněnské syntaktické konference 17.-21.IV.1961* (s. 416–418). Praha: SPN.
- Chloupek, J. (2000). Slohová výstavba textů. In I. Bogoczová, K. Fic, J. Chloupek, E. Jandová, M. Krčmová, & O. Müllerová, *Tváře češtiny*. Ostrava: Ostravská univerzita v Ostravě.
- Chromý, J. (2017). *Protetické V- v češtině*. Praha: Univerzita Karlova.
- Jakobson, R. (1995). Lingvistika a poetika. In *Poetická funkce* (s. 74–105). Jinočany: H&H.
- Jandová, E., David, J., Hoffmannová, J., Svobodová, D., & Müllerová, O. (2006). *Čeština na WWW chatu*. Ostrava: Ostravská univerzita v Ostravě.
- Jefferson, G. (1984). On the organization of laughter in talk about troubles. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (s. 346–369). Cambridge: Cambridge University Press.
- Jefferson, G., Sacks, H., & Schegloff, E. (1977). Preliminary notes on the sequential organization of smile and laughter. In *Pragmatics Microfiche*. Cambridge University, Department of Linguistics.
- Jeziorský, T. (2019). *NET: korpus polooficiální internetové komunikace*. ÚČNK FF UK. Dostupný z <https://www.korpus.cz>
- Jílková, L. (2016). Elektronická komunikace. In J. Hoffmannová, J. Homoláč, L. Jílková, P. Mareš, K. Mrázková, & E. Chvalovská, *Stylistika psané a mluvené češtiny* (s. 105–143). Praha: Academia.

- Jílková, L. (2017). Elektronická komunikace. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/ELEKTRONICK%C3%81%20KOMUNIKACE>
- Joinson, A., McKenna, K., Postmes, T., & Reips, U.-D. (Eds.). (2007). *Oxford Handbook of Internet Psychology*. Oxford: Oxford University Press.
- Kaderka, P. (2013). Situace a komunikace. In O. Uličný (Ed.), *Studie k moderní mluvnici češtiny. I, Pragmatické aspekty češtiny* (s. 57–82). Olomouc: UPOL.
- Kaderka, P. (2017). Symbolické gesto. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/SYMBOLICK%C3%89%20GESTO>
- Kapidzic, S. (2010). *Nonstandard Features of English used in teen chatrooms* [Diplomová práce]. English Department, Univerzita Sarajevo.
- Karlík, P. (2017). Opakování výrazů. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/OPAKOV%C3%81N%C3%8D%20V%C3%9DRAZ%C5%AE>
- Karlík, P., Nekula, M., & Rusínová, Z. (Eds.). (2003). *Příruční mluvnice češtiny*. Praha: NLN.
- Kim, D.-Y., Lehto, X., & Morrison, A. M. (2007). Gender differences in online travel information search: Implications for marketing communications on the internet. *Tourism Management*, 28(2), 423–433.
- Klimešová, P., Komrsková, Z., Kopřivová, M., & Lukeš, D. (2015). Slovo to v mluvených korpusech ČNK, jeho prefixace a reduplikace. *Časopis pro moderní filologii*, 97(1), 21–30.
- Klimešová, P., Komrsková, Z., Kopřivová, M., & Lukeš, D. (2017). Avenues for research on informal spoken Czech based on available corpora. In P. Pezik & J. T. Walinski (Eds.), *Language, Corpora and Cognition* (s. 145–162). Frankfurt am Main: Peter Lang.
- Kodýtek, V. (2007). *On the replicability of the Biber model: The case of Czech*.
- Kohler, K. (2006). ‘Speech-smile’, ‘speech-laugh’, ‘laughter’ and their sequencing in dialogic interaction. In J. Trouvain & N. Campbell (Eds.), *Proceedings of the Interdisciplinary Workshop on The Phonetics of Laughter* (s. 21–26). Saarbrücken: Saarland University.
- Koch, P. (2004). Sprachwandel, Mündlichkeit und Schriftlichkeit. *Zeitschrift für romanische Philologie*, 120, 605–630.
- Koch, P., & Oesterreicher, W. (1985). Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld und Sprachgeschichte. *Romanistisches Jahrbuch*, 36, 15–43.
- Koch, P., & Oesterreicher, W. (1994). Schriftlichkeit und Sprache. In H. Günther & O. Ludwig (Eds.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung* (s. 587–604). Berlin/New York: Gruyter.



- Koch, P., & Oesterreicher, W. (2007). *Lengua hablada en la Romania: Español – francés – italiano. Versión española revisada, actualizada y ampliada por los autores*. Madrid: Gredos.
- Koch, P., & Oesterreicher, W. (2011). *Gesprochene Sprache in der Romania: Französisch – Italienisch – Spanisch*. Berlin/New York: De Gruyter.
- Kolářová, I., & Štícha, F. (2019). *Slovník univerbizátů (zjednoslovněných názvů)*. Praha: Academia.
- Komrsková, Z. (2014). Emotikony a lingvistika. *Lingvistika Praha 2014*. Dostupné z <http://lingvistikapraha.ff.cuni.cz/sbornik>
- Komrsková, Z. (2015). The use of acronyms in different communication modes (a corpus-based study). In K. Gajdošová & A. Žáková (Eds.), *Natural Language Processing, Corpus Linguistics, Lexicography* (s. 63–71). Lüdenscheid: RAM-Verlag.
- Komrsková, Z. (2016). Komunikace prostřednictvím emotikonů. In M. Veselka, D. Nowicka, N. Czemplik, & C. Neumann (Eds.), *Vícejazyčnost. Sborník z konference InterFaces* (s. 17–29). Praha: Karolinum.
- Komrsková, Z. (2017). What does *že jo* (and *že ne*) means in spoken dialogue. *Jazykovedný časopis*, 68(2), 229–237.
- Komrsková, Z. (2019). Užívání emotikonů v soukromé e-mailové komunikaci. *Jazykovědné aktuality*, LVI(1–2), 26–32.
- Komrsková, Z., Kopřivová, M., Lukeš, D., & Poukarová, P. (2017). New spoken corpora of Czech: ORTOFON and DIALEKT. *Jazykovedný časopis*, 68, 219–228.
- Komrsková, Z., & Škrabal, M. (2018). *The role of paralinguistic behaviour, especially laughter in written interview vs. Spoken dialogue. A corpus-based study*. Second International Conference on Sociolinguistics, Budapešť.
- Kopecký, K. (2007). *Moderní trendy v e-komunikaci*. Olomouc: Hanex.
- Kopřivová, M., Goláňová, H., Klimešová, P., Komrsková, Z., & Lukeš, D. (2014). Multi-tier transcription of informal spoken Czech: The ORTOFON corpus approach. In L. Veselovská & M. Janebová (Eds.), *Complex Visibles Out There* (s. 529–544). Olomouc: UPOL.
- Kopřivová, M., Komrsková, Z., Lukeš, D., & Poukarová, P. (2017). Korpus ORAL: sestavení, lematizace a morfologické značkování. *Korpus – Gramatika – Axiologie*, 15, 47–67.
- Kopřivová, M., Komrsková, Z., Lukeš, D., Poukarová, P., & Škarpová, M. (2017). *ORTOFON: Korpus neformální mluvené češtiny s víceúrovňovým přepisem* (Verze 1). ÚČNK FF UK. Dostupný z <http://www.korpus.cz>
- Kosík, K. (1966). Každodennost a dějiny. In *Dialektika konkrétního* (s. 52–61). Praha: Academia.

- Krčmová, M. (1995). K pramenům specifika mluveného projevu. In D. Davidová (Ed.), *K diferenciaci současného mluveného jazyka* (s. 26–32). Ostrava: Ostravská univerzita v Ostravě.
- Krčmová, M. (2000). Zvuková stavba textů. In I. Bogoczová, K. Fic, J. Chloupek, E. Jandová, M. Krčmová, & O. Müllerová, *Tváře češtiny* (s. 119–138). Ostrava: Ostravská univerzita v Ostravě.
- Krčmová, M. (2017). Kontoid. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/KONTOID>
- Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Jelínek, T., Kovářiková, D., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Vondříčka, P., & Zasina, A. J. (2015). *SYN2015: Reprezentativní korpus psané češtiny*. ÚČNK FF UK. Dostupný z <http://www.korpus.cz>
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19(3), 273–309.
- Labov, W. (2006). *The Social Stratification of English in New York City*. Cambridge: Cambridge University Press.
- Lee, C. (2011). Micro-blogging and status updates on Facebook: Texts and practices. In C. Thurlow & K. Mroczek (Eds.), *Digital Discourse: Language in the New Media* (s. 110–130). Oxford: Oxford University Press.
- Leech, G. (2000). Grammars of spoken English: New outcomes of corpus-oriented research. *Language Learning*, 50(4), 675–724.
- Lišková, M. (2006). Některé komunikační a mluvnické aspekty počítačového slangu. In P. Pořízka & V. Polách (Eds.), *Tzv. Základní výzkum v lingvistice—Desideratum, nebo realis?* (s. 46–52). Olomouc: UPOL.
- Machač, P., & Zíková, M. (2013). Redukční procesy v řeči z hlediska fonetických rysů. In M. Beneš, P. Machač, Z. Palková, M. Prošek, K. Smejkalová, O. Uličný, & M. Zíková, *Studie k moderní mluvnici češtiny. 5, K české fonetice a pravopisu* (s. 17–44). Olomouc: UPOL.
- Mareš, P. (2003). Spisovnost a nespisovnost, formálnost a neformálnost. In *Přednášky z 46. běhu Letní školy slovanských studií* (Roč. 2003, s. 99–108). Praha: FF UK.
- Mareš, P. (2016). Úvod. In J. Hoffmannová, J. Homoláč, E. Chvalovská, L. Jílková, P. Mareš, K. Mrázková, & P. Kaderka, *Stylistika mluvené a psané češtiny* (s. 11–21). Praha: Academia.
- Markman, K. M., & Oshima, S. (2007). *Pragmatic play? Some possible functions of English emoticons and Japanese kaomoji in computer-mediated discourse*. Association of Internet Researchers Annual Conference 8.0: Let's Play!, Vancouver, B. C., Kanada.
- Mathesius, V. (1947). Pronominální podmět v hovorové češtině. In *Čeština a obecný jazykozpyt* (s. 286–293). Praha: Melantrich.

- McCarthy, M. (1998). *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.
- McCulloch, G. (2019). *Because Internet. Understanding the New Rules of Language*. New York: Riverhead Books.
- McMahon, A. M. S. (1994). *Understanding Language Change*. Cambridge: Cambridge University Press.
- Meyer, R. (2017). Otázka. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/OTÁZKA>
- Mikulová, M., & Urešová, Z. (2008). Rekonstrukce standardizovaného textu z mluvené řeči. In M. Kopřivová & M. Waclawičová (Eds.), *Čeština v mluveném korpusu* (s. 167–176). Praha: NLN.
- Miller, J., & Weinert, R. (1998). *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Oxford University Press.
- Mlčoch, M. (2005). Anglicismy v chatové komunikaci mládeže aneb Hod'te nějakéj link na dobrej house. In O. Uličný (Ed.), *Eurolitteraria & Eurolingua* (s. 305–311). Liberec: Technická univerzita.
- Mrázková, K. (2017). Rejstřík. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/REJSTRÍK>
- Müllerová, O. (1979). *Komunikativní složky výstavby dialogického textu*. Praha: Karlova Univerzita.
- Müllerová, O. (1989). Ke vztahu mluveného a psaného textu (srovnání mluveného a psaného vypravování). *Slovo a slovesnost*, 50(3), 205–216.
- Müllerová, O. (1994). *Mluvený text a jeho syntaktická výstavba*. Praha: Academia.
- Müllerová, O. (2004). Opakování a syntaktický paralelizmus v rozhovorech s muži a ženami v jednom z druhů institucionální komunikace. *Naše řeč*, 87(3), 124–135.
- Müllerová, O. (2007). Postpozitivní žejo, jo ve vzpomínkovém vyprávění. In J. Hoffmannová & O. Müllerová (Eds.), *Čeština v dialogu generací* (s. 291–297). Praha: Academia.
- Müllerová, O. (2011). Výstavba mluvených projevů, jejich členění a syntax. In S. Čmejrková & J. Hoffmannová (Eds.), *Mluvená čeština: Hledání funkčního rozpětí* (s. 129–172). Praha: Academia.
- Müllerová, O., & Hoffmannová, J. (1994). *Kapitoly o dialogu*. Praha: Pansofia.
- Nekula, M. (2003a). Aktuální členění. In P. Karlík, M. Nekula, & Z. Rusínová (Eds.), *Příruční mluvnice češtiny* (s. 633–645). Brno: NLN.
- Nekula, M. (2003b). Text. In P. Karlík, M. Nekula, & Z. Rusínová (Eds.), *Příruční mluvnice češtiny* (s. 652–698). Brno: NLN.

- Nekula, M. (2010). Navazování a udržování kontaktu a oslovení. In S. Čmejrková & J. Hoffmannová (Eds.), *Užívání a prožívání jazyka. K 90. Narozeninám Františka Daneše* (s. 239–244). Praha: Karolinum.
- Nekula, M. (2017). Text. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy — Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/TEXT>
- Nekvapil, J. (2017). Konverzační analýza. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/KONVERZA%C4%8CN%C3%8D%20ANAL%C3%9DZA>
- Nekvapil, J., & Havlík, M. (2017). Replika v rozhovoru. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/REPLIKA%20V%20ROZHOVORU>
- Nöth, W. (2000). *Handbuch der Semiotik*. Stuttgart, Weimar: Verlag J. B. Metzler.
- Orton-Johnson, K., & Prior, N. (Eds.). (2013). *Digital Sociology: Critical Perspectives*. London/New York: Palgrave Macmillan.
- Palková, Z. (1997). *Fonetika a fonologie češtiny*. Praha: Karolinum.
- Palková, Z. (2017). Větná intonace. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/V%C4%9ATN%C3%81%20INTONACE>
- Pappert, S. (2017). Zu kommunikativen Funktionen von Emojis in der WhatsApp-Kommunikation. In M. Beißwenger (Ed.), *Empirische Erforschung internetbasierter Kommunikation* (s. 175–211). Berlin/New York: De Gruyter.
- Park, J., Baek, Y. M., & Cha, M. (2014). Cross-cultural comparison of nonverbal cues in emoticons on Twitter: Evidence from big data analysis. *Journal of Communication*, 64(2), 333–354.
- Paulus, T., & Phipps, G. (b.r.). Approaches to case analyses in synchronous and asynchronous environments. *Journal of Computer-Mediated Communication*, 13(459–484), 2008.
- Paulus, T., Warren, A., & Lester, J. (2016). Applying conversation analysis methods to online talk: A literature review. *Discourse, Context & Media*, 12, 1–10.
- Pérez-Sabater, C. (2012). The linguistics of social networking: A study of writing conventions on Facebook. *Linguistik Online*, 56(6). Dostupné z [https://pdfs.semanticscholar.org/2182/1dd810f055541f5ecf9f2b44d88cfb0bcdfe.pdf?\\_ga=2.80353879.1579329132.1585062055-493910997.1576674711](https://pdfs.semanticscholar.org/2182/1dd810f055541f5ecf9f2b44d88cfb0bcdfe.pdf?_ga=2.80353879.1579329132.1585062055-493910997.1576674711)
- Peters, P., & Wong, D. (2015). Turn management and backchannels. In K. Aijmer & C. Rühlemann (Eds.), *Corpus Pragmatics: A Handbook* (s. 408–429). Cambridge: Cambridge University Press.

- Petkevič, V. (2017). Lemmatizace. In P. Karlík, M. Nekula, & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny*. Dostupné z <https://www.czechency.org/slovník/LEMMATIZACE>
- Petrusek, M. (2017). Každodennost. In Z. J. Nešpor (Ed.), *Sociologická encyklopedie*. Sociologický ústav AV ČR. Dostupné z <https://encyklopedie.soc.cas.cz/w/Ka%C5%BEdodennost>
- Petrusek, Miloslav. (1992). *Alternativní sociologie: Úvahy o smyslu sociologie v nealternativní společnosti*. Praha: Klub osvobozeného samizdatu.
- Potter, J., & Hepburn, A. (2010). Putting aspiration into words: ‘Laugh particles’, managing descriptive trouble and modulating action. *Journal of Pragmatics*, 42(6), 1543–1555.
- Pravidla českého pravopisu*. (1999). Praha: Fortuna.
- Prokešová, S. (1962). Jak se odráží improvizovanost v jazykové výstavbě mluvených projevů. In *Otázky slovanské syntaxe: Sborník brněnské syntaktické konference 17.–21. IV. 1961* (s. 348–351). Praha: SPN.
- Provine, R. R., Spencer, R. J., & Mandell, D. L. (2007). Emotional expression online: Emoticons punctuate website text messages. *Journal of Language and Social Psychology*, 26, 299–307.
- Reifová, I., & et al. (2004). *Slovník mediální komunikace*. Praha: Portál.
- Rezabek, L., & Cochenour, J. (1998). Visual cues in computer-mediated communication: Supplementing text with emoticons. *Journal of Visual Literacy*, 18(2), 201–215.
- Rovai, A., & Baker, J. (2005). Gender differences in online learning. *Quarterly Review of Distance Education*, 6(1), 31–44.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Scott, C. R., & Timmerman, C. E. (2005). Relating Computer, Communication, and Computer-Mediated Communication Apprehensions to New Communication Technology Use in the Workplace. *Communication Research*, 32(6), 683–725.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Schnitzer, C.-V. (2012). *Linguistische Aspekte der Kommunikation in den neueren elektronischen Medien SMS – E-Mail – Facebook*. München: Ludwig-Maximilians-Universität München.
- Sinclair, J. (1991). *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.
- Sinclair, J., & Mauranen, A. (2006). *Linear Unit Grammar: Integrating Speech and Writing*. Amsterdam/Philadelphia: Benjamins.
- Skovholt, K., Grønnig, A., & Kankaanranta, A. (2014). The communicative functions of emoticons in workplace e-mails: :-). *Journal of Computer-Mediated Communication*, 19, 780–797.

- Straková, J., Straka, M., & Hajič, J. (2014). Open-source tools for morphology, lemmatization, POS tagging and named entity recognition. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 13–18.
- Suchomel, V. (2012). Recent Czech web corpora. In A. Horák & P. Rychlý (Eds.), *Proceedings of Recent Advances in Slavonic Natural Language Processing* (s. 77–83). Brno: Tribun EU.
- Svobodová, D. (2006). Slovní zásoba na www chatu. In E. Jandová, J. David, J. Hoffmannová, D. Svobodová, & O. Müllerová, *Čeština na WWW chatu* (s. 101–158). Ostrava: Ostravská univerzita v Ostravě.
- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Škrabal, M., & Komrsková, Z. (2019). *Top frequency words in written/spoken Czech and English – can they be used to measure a rate of diglossia?* *Corpus Linguistics 2019*, Cardiff.
- Šmilauer, V. (1966). *Novočeská skladba*. Praha: SPN.
- Štícha, F. (2010). Nerozumím tomu vs. Já tomu neorzumím (K užívání zájmena já z hlediska korpusových nálezů). In S. Čmejrková, J. Hoffmannová, & E. Havlová (Eds.), *Užívání a prožívání jazyka* (s. 227–232). Praha: Karolinum.
- Štícha, F. (2013). *Akademická gramatika spisovné češtiny*. Praha: Academia.
- Tagliamonte, S., & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen Language. *American Speech*, 83(1), 3–34.
- Thurlow, C., Lengel, L., & Tomic, A. (2004). *Computer Mediated Communication: Social Interaction and the Internet*. London/Thousand Oaks/New Delhi: Sage.
- Tossell, C., Kortum, P., & Shepard, C. (2012). A longitudinal study of emoticon use in text messaging from smartphones. *Computers in Human Behavior*, 28(2), 659–663.
- Urbanová, L. (2003). *On expressing meaning in English conversation*. Brno: Masarykova univerzita.
- Vachek, J. (1942). Psaný jazyk a pravopis. In B. Havránek & J. Mukařovský (Eds.), *Čtení o jazyce a poesii* (s. 231–306). Praha: Družstevní práce.
- Valček, P. (2011). *Slovník teórie médií*. Bratislava: Literárne informačné centrum.
- Varnhagen, C. K., McFall, G. P., Pugh, N., Routledge, L., Sumida-MacDonald, H., & Kwong, T. E. (2010). lol: New language and spelling in instant messaging. *Reading and Writing*, 23(6), 719–733.
- Vepřek, J. (2018). *Individuální stylové dominanty v postojových diskusích na sociální síti Facebook. Originalita vyjadřování ve veřejné psané komunikaci na internetu a její příčiny* [Disertační práce], Masarykova univerzita, Filozofická fakulta.
- Wilson, J. (2010). *Moravians in Prague*. Frankfurt am Main: Peter Lang.

- Wolf, A. (2000). Emotional expression online: Gender differences in emoticon use. *Cyberpsychology and Behaviour*, 3, 827–833.
- Zasina, A. J., Lukeš, D., Komrsková, Z., Poukarová, P., & Řehořková, A. (2018). *Koditex: Korpus diverzifikovaných textů*. ÚČNK FF UK, Praha. Dostupný z <http://www.korpus.cz>
- Zeman, J. (2013). *Verbální a neverbální prostředky v komunikaci mládeže*. Hradec Králové: Gaudeamus.
- Zimová, L. (1988). Vyjadřování podmětu osobními zájmeny 1. a 2. osoby. *Naše řeč*, 71(1), 22–32.