Masarykova univerzita
Filozofická fakulta

Ústav českého jazyka

Český jazyk se specializací počítačová lingvistika

Daniela Ryšavá

Automatická tvorba imperativů z infinitivu

Magisterská diplomová práce

Vedoucí práce: doc. PhDr. Klára Osolsobě, Dr.

2017

Prohlašuji, že jsem magisterskou diplomovou práci (část projektu, kterou v ní popisuji) vypracovala samostatně a pouze na základě pramenů a zdrojů, které uvádím v seznamu použité literatury.

_____

Podpis autora

Velice ráda bych poděkovala doc. PhDr. Kláře Osolsobě, Dr., za možnost psát diplomovou práci pod jejím vedením, za velmi milé, trpělivé a přínosné konzultace a za spoustu cenných rad. Děkuji také Mgr. Marku Grácovi, PhD., za přátelské pracovní prostředí a za to, že mi umožnil se na projektu podílet, a získávat tak cenné zkušenosti.

Velké díky patří také mé kolegyni a kamarádce Nikol Volkové za spolupráci a sdílení všech studijních i pracovních problémů a radostí.

Nejvíce však vděčím svému manželovi Jakubovi a našemu malému synovi Štěpánovi za trpělivost, prostor a nekonečnou motivaci a celé své rodině za povzbuzení, pomoc a podporu. Děkuji své mamince, která mě naučila nikdy se nevzdávat.

Cením si toho a mám radost, že jsem díky Vám všem dokázala dokončit diplomovou práci a s ní snad i zdárně uzavřít jednu krásnou kapitolu mého života.

Obsah

Úvod **6**

1 Projekt tvorby dat pro morfologický analyzátor **8**

1.1 O systému vzorů 9

1.1.1 Syntax kódu 13

1.2 Nová řešení v popisu vzorů 14

1.2.1 Parciální vzory sloves 14

1.2.2 Záplaty 15

1.2.3 Charakteristika hlásek a automatické alternace 16

1.2.4 Přidané operace 16

1.3 Guesser 18

2 Využití korpusových dat během výzkumu **19**

3 Infinitiv a imperativ v odborné literatuře **20**

3.1 Slovo 20

3.1.1 Stavba slova 20

3.2 Sloveso 22

3.2.1 Gramatické kategorie sloves 22

3.2.1.1 Kategorie slovesného způsobu (modus) 23

3.3 Imperativ 23

3.3.1 Imperativ a jeho způsoby vyjádření 24

3.3.1.1 Užití imperativních tvarů 25

3.3.1.2 Jiné jazykové konstrukce (jako alternativa ke tvarům imperativu) 26

3.3.2 O imperativu v (odborné) literatuře 30

3.3.3 Imperativ v rámci mluvnického systému vzorů 33

3.3.3.1 Slovesné tvary podle kmene prézentního 33

3.3.3.2 Slovesné tvary podle kmene infinitivního/minulého 38

3.4 Infinitiv 41

4 Hláskové alternace ... 45
4.1 Samohláskové alternace ... 45
4.2 Souhláskové alternace ... 46
4.3 Popis pravidel alternací při tvoření tvarů sloves ... 46
4.3.1 Pravidla samohláskových alternací ... 47
4.3.2 Pravidla souhláskových alternací ... 48
4.3.3 Grafické alternace ... 49
4.4 Hláskové alternace v mluvnických vzorech slovesných tříd dle prézentního kmene ... 50
4.5 Shrnutí k tématu alternací ... 52
5 Homonymie imperativních tvarů ... 53
6 Analýza tvarů imperativu ... 55
6.1 Postup a výsledky prohledávání korpusových dat ... 55
6.1.1 Tvary imperativu ve 2. osobě singuláru ... 55
6.1.2 Plurálové tvary imperativu ... 60
7 Tvorba vzorů pro imperativy ... 62
7.1 Postup při popisu vzorů ... 63
7.1.1 Imperativy od infinitivů na *-at* ... 64
8 Analýza vzorů ... 71
8.1 Imperativ ve vztahu k infinitivu ... 77
8.2 Imperativ vzhledem k hláskovým změnám ... 79
8.3 Přiřazení priority ... 83
8.4 Anotace v číslech ... 83
8.4.1 Vzory nabízené během anotací ... 83
8.4.2 Přiřazené vzory ... 84
Závěr ... 88
Seznam použité literatury ... 90
Tištěné zdroje ... 90
Elektronické zdroje ... 92
Přílohy ... 93

Úvod

Český jazyk je morfologickými analyzátory a jejich databázemi popsán různými způsoby. Jejich kvalitu nepopíráme, ačkoliv máme v rámci celého projektu (nikoliv této diplomové práce) ambice je překonat. Především upozorňujeme na jejich hlavní nedostatek – nepřístupnost zdrojových dat. Vznikla proto myšlenka vytvořit novou databázi, která by byla plně veřejně dostupná a ke které by mohl svými připomínkami či vlastní prací přispět kdokoliv, kdo by se chtěl na její tvorbě podílet a zdokonalovat ji.

Marek Grác (2006) tuto myšlenku poprvé zrealizoval ve své magisterské práci *Tvorba morfologické databáze z neoznačkovaného korpusu*. Pod jeho vedením pak českou morfologii zpracovala v diplomové práci *Určování tvarotvorných vzorů neznámých slov* Dagmar Husárová (2015), která využila příležitosti tvořit vzory pro českou flexi jinak, než je tomu v české mluvnici zvykem. Vzory mají zachovat určitou systematičnost a přehlednost. Počítačová lingvistika však navíc nabízí možnosti oprostit se od mluvnického systému vzorů a vytvořit vzory nové, postavené na korpusových datech, které umožňují hledat skupiny slov s podobnými rysy chování při flexi.

Po úspěšném dokončení celého systému vzorů pro českou deklinaci i konjugaci Dagmar Husárové se Marku Grácovi podařilo realizovat projekt pod záštitou firmy Red Hat, v níž dnes vedle své akademické činnosti působí. Spolu s ním jakožto koordinátorem projektu a s Nikol Volkovou jsme začali na projektu pracovat. Upravili jsme přístup, podle kterého jsme přepracovali nový systém vzorů, vylepšili technickou stránku zápisu a tvoření pravidel a na základě nových vzorů jsme anotovali slova, na nichž se posléze pokusíme postavit automatické řazení správných vzorů k dalším slovům.

Zaměření diplomové práce je zúženo na problém tvorby slovesných imperativů ze základního tvaru – infinitivu. V práci přihlížíme k jazykovým mluvnicím, zabýváme se mluvnickými vzory a hledáme v nich důkazy toho, že nejsou uzpůsobeny tvorbě imperativů. Mluvnické vzory jsou odvozeny většinou od prézentního kmene. Tvary jiné než indikativ prézentu podle stanovených tříd a vzorů jsou proto často nepravidelné. Jelikož jsme výchozím tvarem pro tvoru imperativů stanovili infinitiv (nikoliv 3. osobu singuláru indikativu prézentu, jak je v mluvnicích běžné), je nám bližší dělení sloves podle kmene infinitivního. Ani v tomto případě však mluvnice zdaleka nenabízí konzistentní řešení pro imperativ. Stejný problém nacházíme i u dalších slovesných tvarů, což nás přivádí k myšlence vytvořit parciální vzory, které budou řešit jednotlivé slovesné tvary izolovaně. Imperativ je tak jedním ze sedmi dílčích skupin, které dohromady tvoří celé slovesné paradigma. Je zároveň jedním ze sedmi slovesných vzorů, které jsou při anotacích slovesům přiřazovány.

Snahou našeho přístupu je zůstat na půli cesty mezi jazykovou schopností a strojovým zpracováním. Vzory tvoříme tak, aby byly intuitivní, do systému řazené v rámci logické hierarchie. Je však třeba přesně strojově popsat slovesa jako řetězce znaků, charakterizovat skupiny sloves na základě jejich infinitivního ukončení a definovat jejich společné rysy. Systém vzorů by měl být dostatečně konkrétní,

aby zpracoval běžná i okrajová slova české slovní zásoby. Naopak by měl být také dostatečně obecný, aby tvořil správné tvary od sloves nových, dosud v databázi nezavedených. Dostojíme-li v projektu všemu zmíněnému, pak se domníváme, že jsme schopni vytvořit i jinak než počítačově užitečný systém vzorů, který je zároveň funkční a úspěšný i v oblasti automatické morfologické analýzy a mohl by být návodem i pro databáze dalších jazyků.

1 Projekt tvorby dat pro morfologický analyzátor

V rámci projektu se zabýváme tvorbou vzorů pro dílčí subparadigmata slovesných tvarů z infinitivního základu. Hlavním cílem je vytvořit systém takových vzorů, které pokryjí celé formální spektrum sloves, jak pravidelné jevy, tak výjimky. V závislosti na vzorech vytváříme slovník, který odráží výsledky nového systému.

Samotná myšlenka projektu vznikla s magisterskou prací Marka Gráce (2006), tehdy studenta Aplikované informatiky na Masarykově univerzitě, který vytvořil morfologickou databázi pro slovenský jazyk. Na jejím základě a pod jeho odborným vedením zpracovala Dagmar Husárová (2015) český ekvivalent, v němž je popsána celá deklinace a konjugace českých substantiv, adjektiv a sloves. Diplomová práce Dagmar Husárové je pro náš projekt kvalitním východiskem. Klademe si za cíl proces tvorby databáze vylepšit, přepracovat a zefektivnit ho tak, abychom získali co nejpřesnější výsledek, minimalizovali nutnost ručních anotací, a tím snížili pravděpodobnost vzniku chyb. Projekt, v rámci kterého má být kompletní morfologická databáze češtiny vytvořena, zaštiťuje firma Red Hat. Marek Grác, dnes výzkumný a vývojový pracovník na Filozofické fakultě Masarykovy univerzity a zároveň softwarový inženýr ve firmě Red Hat, v projektu figuruje jako koordinátor a věnuje se jeho technické stránce.

Morfologická databáze jakožto data pro morfologický analyzátor, kterou ve práci představíme, není zdaleka první a jistě ani poslední, která vznikla. Nepopiratelnou výhodou pro její využití však je, že firma Red Hat nabízí své produkty jako tzv. open source. Díky tomu jsou data našeho projektu neomezeně plně přístupná každému, kdo by s nimi chtěl jakkoli dále pracovat, upravovat je, či je přímo využít v libovolných systémech nebo aplikacích. Většina morfologických analyzátorů a jejich databází vzniká na akademické půdě (např. brněnské *ajka* a *majka* z Fakulty informatiky Masarykovy Univerzity, *Morfologický analyzátor pomocí konečných automatů* z Vysokého učení technického, pražský Hajičův morfologický analyzátor či ve firmách (např. Google, Lingea), jiné firmy používají akademické databáze (např. Seznam.cz, Heureka). Dostupné jsou tak často pouze výstupy analyzátorů (prostřednictvím uživatelských rozhraní) nebo jsou zpoplatněny. Příčinou menší poptávky po zpřístupnění zdrojových dat však může být postavení češtiny mezi ostatními jazyky. Aplikace pro češtinu jakožto malý jazyk nejsou tržně masivně žádané. Tendence směřují spíše k nástrojům pracujícím se světovými jazyky, v současnosti zejména s jazykem anglickým.

Druhou a z funkčního hlediska zásadní výhodou je schopnost automatického doplnění o neznámá slova. Znamená to, že aplikace, v nichž bude databáze použita, budou umět zpracovat i slova, která v databázi nejsou obsažena, a zapamatovat si je. Za účelem obohacení databáze o další skupinu slov nebude třeba vytvářet nové verze.

Kód, ve kterém jsou pravidla popsána, je jednoduchý, uživatelsky přívětivý. Jeho podobě se budu podrobněji věnovat v kapitole 1.1.1#. Tuto skutečnost ocení zejména lingvisté a další humanitně zaměření uživatelé, kteří by chtěli s našimi vzory pracovat.

V neposlední řadě si klademe za cíl vytvořit systém, který bude aplikovatelný na další (alespoň typově podobné) jazyky. Tuto schopnost dokazuje samotný vznik české databáze, která byla tvořena podle slovenského vzoru. Předpokládáme, že snadno bude celý systém převoditelný zejména na slovanské jazyky, které mají společný prajazyk a u nichž jsou gramatické kategorie vyjadřovány formálně podobně. Stejně tak bychom se chtěli pokusit aplikovat systém na jazyky, jako je například francouzština nebo němčina, které jsou tržně zajímavé.

Diplomová práce Marka Gráce slouží projektu jako „návod“, podle kterého vznikala práce Dagmar Husárové a kterého se nadále držíme i my. Metodologický postup popisu pravidel pro generování slovních tvarů jsme zachovali, měníme spíše jazykový přístup. Jinak řečeno – pravidla píšeme stejným způsobem, liší se pouze postoj k tomu, co popisují. V první řadě v následujícím textu popíšeme systém pravidel, který je základem tvorby dat. Poté budeme charakterizovat problémy, na které jsme v původních pravidlech narazili. Rozebereme změny, které jsme provedli, abychom systém pravidel vylepšili.

1.1 O systému vzorů

Ačkoliv se diplomová práce zaměřuje primárně na generování imperativních slovesných tvarů, dovolujeme si věnovat kapitolu alespoň obecnému popisu celého systému vzorů, jeho struktuře a snaze o nový pohled na deklinaci a konjugaci českých slov.

Nový systém vzorů má představovat kompromis mezi počítačovým zpracováním a klasickým mluvnickým popisem. Strojový popis musí být velmi exaktní, protože počítač pochopitelně žádné předchozí znalosti o jazyce, kterými disponuje běžný uživatel jazyka, nemá. Vzniká tak velké množství pokynů, se kterými operuje počítačový program, a dobrá čitelnost pro běžného uživatele není vyžadována. V rámci počítačového zpracování přirozeného jazyka lze pracovat na základě velkých dat, snadno využívat různé složité vzorce a statistické údaje. Mluvnice má nevyvratitelnou výhodu, že je napsána v přirozeném jazyce pro adresáta, který má s fungováním přirozeného jazyka předchozí zkušenosti, o něž se opírá, takže mu stačí poměrně jednoduché rozdělení ohebných slov mezi deklinační a konjugační vzory, které zná již ze základní školy. Není třeba vytvářet speciální podvzory kvůli různým alternacím, se kterými si počítač bez popisu neporadí (*kluk* > *kluci*, nikoliv **kluki*). Přesto se v jazyce vyskytuje nespočet výjimek a dublet, které se musí naučit i rodilý mluvčí. Ačkoliv popisujeme jazyk pro strojové využití, snažíme se, aby exaktní popis vzorů byl co nejjednodušší, nejčitelnější i pro ty, kteří jej netvořili. Postupujeme systematicky pokud možno podle známých mluvnických

vzorů: naznačujeme hierarchii vzorů, jednotlivé úrovně tvoříme podle stejných kritérií, jako jsou např. různé druhy alternací.

Jak jsme již zmínili úvodem, projekt se primárně zaměřuje na substantiva, adjektiva a slovesa. Tyto tři skupiny české slovní zásoby zahrnují (co se formy týče) i část zájmen a číslovek, které v mnoha případech formálně odpovídají jmenným slovním druhům. Jednak se shodují v koncovkách (př. *pětka, který, čí, první, pátý*), jednak se jedná o substantiva, a to buď tzv. číslovky nepravé, tj. substantiva, která vlastnosti číslovek nabyla v rámci kontextu (př. *řada, hromada*), nebo číslovky základní (*desítka*, *sto, tisíc, milion, miliarda*). Nedržíme se striktně hranic mezi lingvisty definovanými slovními druhy. Ostatní jednotky[1], typicky neohebné slovní druhy a slova pravidly nezachytitelná, v této fázi projektu ponecháváme stranou s tím, že se do databáze zanesou jednotlivě.[2]

Při tvorbě pravidel přihlížíme k morfologické charakteristice příslušné skupiny slov, částečně z ní vycházíme, plně na ni ovšem nespoléháme a využíváme ji především k ověření pravidel. K tomu nám také slouží data z Internetové jazykové příručky (dále jen IJP), slovníků (on-line prostřednictvím platformy *DebDict*) a korpusů (korpusy řady SYN). Lingvisté by nám mohli vyčítat, že ačkoliv existuje kvalitní popis české morfologie, nevyužíváme znalostí o morfematice a alternacích v rámci slovních ukončení, jako je tomu u jiných dat morfologických analyzátorů, viz *ajka*, Osolsobě (1996). Nový přístup umožňuje nahlížet na slova jako na neznámé jednotky a hledat mezi nimi spojitosti, které bychom bez použití počítačových nástrojů a práce s velkým množstvím testovacích dat nebyli schopni objevit.

Ke slovu přistupujeme jako k řetězci znaků abecedy. Při popisu pravidel se zaměřujeme především na zakončení řetězce, a to tak, že postupně čteme jednotlivé znaky od poslední pozice ve slově a hledáme společné rysy pro charakteristickou skupinu slov – co nejdelší, aby byl popis co nejvíce exaktní, ale dost krátký na to, aby zahrnul co nejvíce slov s popsanými rysy. Například každé pravidlo pro slovesa popisuje slovo jako řetězec, který končí na *-t*, *-ci*. Dále se pravidla krystalizují na základně tvarů, které z nich generujeme. Jedno pravidlo tak například popisuje skupinu sloves končících na *-ovat*, která jsou v imperativu zakončena na *-uj* (př. *kupovat* > *kupuj*, *studovat* > *studuj*, *upgradovat* > *upgraduj*, nikoliv *chovat* > **chuj*).

Vstupem pro systém určování vzorů je slovo v základním tvaru, tzv. lemmatu. Lemma je charakterizováno jako „reprezentativní slovníková podoba“ slova (NESČ, 2016, s. 962). Pro substantiva a adjektiva je to tvar nominativu singuláru, pro slovesa pak tvar infinitivu. Systém zpracovává pouze jeden řetězec znaků, proto neodpovídá vždy významové charakteristice slova, po formální stránce víceslovnému výrazu (př. *být* > *byl jsem*). Přesněji řečeno je vstupem tzv. textové slovo, jak je

[1] kromě stupňování příslovcí

[2] Pokud se podaří domluvit další spolupráci s firmou Red Hat, projekt by mohl být rozšířen o vzory řešící derivaci slov napříč slovními druhy, př. substantiva slovesná (*plavání*) od sloves, adverbia od adjektiv (*krásně*) apod.

charakterizuje NESČ (2016, s. 1694), a to jako jednotku grafickou (řetězec znaků oddělený mezerami/oddělovači).

Pravidla si lze představit jako několikavrstvé síto, jehož jednotlivá patra představují vždy jeden znak od konce slova. Čím přesnější je popis vzoru, tím dále slovo propadne a získá tím pravděpodobnější charakteristiky. V rámci pravidel lze definovat také negativní podmínky. Např. infinitivy charakterizované ukončením *.*c[ti]* často v imperativu podléhají alternaci *c* > *č* (*péct/péci* > *peč*). Tímto „otvorem" v sítu však projde i infinitiv *říc[ti]*, který nikdy na stejném principu imperativ netvoří. Zavedeme proto negativní charakteristiku „vyjma vstup *.*říc[ti]*" a definujeme pro něj jiný, vlastní vzor. Slovu *říci/říct* tak bude automaticky přiřazen pouze jeden vzor a nebude třeba spoléhat na ruční anotaci či guesser, který by mezi těmito dvěma vzory rozhodoval, který je správný. Na stejném principu lze samozřejmě definovat mnohem větší skupiny slov, což ve výsledku umožní přiřazovat vzory ke slovům plně automaticky nebo vybírat jen mezi malým počtem potenciálních vzorů.

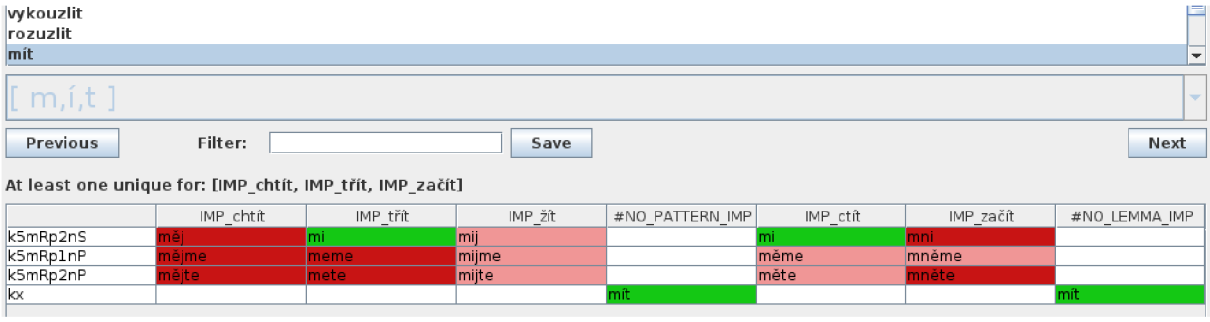
Vzory jsme testovali na sadě o velikosti zhruba 100 000 slov, které jsme ručně anotovali. Mezi vzory, které systém nabídl jako odpovídající danému řetězci, jsme označili ty, které generovaly všechny existující správné tvary. Záměrně se vyhýbáme tvrzení, že jsou vybrané tvary spisovné, jelikož jsme zvolili benevolentnější hodnocení a zařazovali i výrazy a jejich tvary nespisovné i jinak příznakové (archaismy atd.). Vycházíme z faktu, že systém nemá za úkol kontrolovat spisovnost. Nespisovné prvky uvnitř slova mimo popsané ukončení může pravidly projít stejně jako slovo spisovné. Náš systém zohledňuje nespisovné tvary v těchto případech:

a) Nespisovnost je způsobena prvky uvnitř slovesa, tj. mimo ukončení popsané vzorem. Formálně vzory tvořit imperativy od těchto sloves umí. Jedná se o protetické *v-* (*opakovat* i *vopakovat*), slova s variantami s/z (*kompenzovat* i *kompensovat*), *-ej-* z původního *-ý-* uvnitř slovního základu, které není zahrnuto do definic vzoru (*vydýchávat* i *vydejchávat*). Všeobecně platí, že variantní imperativ tvoříme od variantního infinitivu (př. *opakovat* > *opakuj*, *vopakovat* > *vopakuj*).
b) Vzory jsou tvořeny tak, aby zpracovaly i slovesa s určitým druhem nespisovných jevů, které se týkají charakteristiky popsaného infinitivního ukončení. Jedná se o slovesa s nespisovným *-ej-* (pův. *-ý-*, př. *týrat* i *tejrat*) v ukončení, opět rozlišujeme variantní infinitivy (*týrat* > *týrej*, *tejrat* > *tejrej*)[3]. Dva základy tvoří stejný imperativ u variantních infinitivů typu *píct* ke spisovným *péct* (*péci/péct* > *peč*). Naopak z jednoho tvaru generujeme nepříznakový a variantní imperativ (př. *vzít* > *vezmi/vem*, slovesa typu *vyhlásit* > *vyhlas/vyhlaš*, *skákat* > *skákej/skač*).

[3] V popisu vzoru je charakterizováno ukončení *.*V+rat/.*ejrat*.

c) Nespisovné výrazy se formálně nijak neliší od tvarů spisovných. Otázka, zda mají či nemají být ve výsledném slovníku zařazeny, záleží na jeho konkrétním využití. Nám jde především o co nejširší pokrytí úzu, proto anotujeme i slovesa nespisovná a vulgární jako např. *bonzovat*, *camrat*, *srát* apod.

Anotace provádíme ve speciálním rozhraní, které graficky přehledně znázorňuje jednotlivé vzory a jimi generované tvary. Navíc barevně odlišuje tvary, které jsou doložené v korpusu (sytě zeleně zbarvené) či nikoliv (bledě zelené) a zda jsou mezi nabízenou skupinou tvarů jedinečné (sytě červená značí tvar, který je doložený v korpusu a je formálně odlišný od všech tvarů všech nabízených vzorů, bledě červená značí to samé u tvarů nedoložených).



Obrázek 1: Ukázka rozhraní pro ruční anotace

Během anotací jsme se opírali o neoznačkovaná korpusová data získaná z korpusu czTenTen, pracovali jsme pouze s informací „je/není doloženo v korpusu“, což nemuselo být s ohledem na homonymii vždy určující (*tvarohové* jako potenciální maskulinní substantivum v nominativu plurálu se zakončením *-ové*, *vstřebanou* jako špatně vytvořená 3. osoba plurálu slovesa *vstřebat* apod.). Ve slovnících jsme ověřovali správnost generovaných tvarů. Primárně jsme využívali IJP, která sice ne vždy plně koresponduje se slovníky češtiny a není kodifikovaná, ale je přehledná a aktuální.[4]

Výstupem systému po anotacích je textový dokument obsahující dvojice *výraz a jemu přiřazené vzory* ve formě:

lemma:vzor1;vzor2;…;vzorN (pro libovolný počet přiřazených vzorů).

Celkem jsme vytvořili 68 vzorů. Důležitou roli však nehraje jejich celkové množství, ale počet vzorů, přes které výraz prošel – čím méně, tím větší pravděpodobnost, že pravidla vzoru generují správné tvary zadaného výrazu. Celkové množství popsaných vzorů není nijak určující.

[4] Nejnovější výkladový *Slovník spisovné češtiny* byl poprvé vydán roku 1978 a naposledy upraven a doplněn roku 2003 (4. vydání z roku 2003 je obsahově totožné s předchozím.) IJP vznikla roku 2008 a její poslední verze je z roku 2017.

1.1.1 Syntax kódu

V této podkapitole se pokusíme stručně popsat vzorové pravidlo psané v XML kódu. Podobný popis se nachází i v práci Dagmar Husárové (2015, s. 35–73).

Jazyk XML je velice intuitivní, základem psaní vzorů podle naší metodologie je pochopení syntaxe a jednotlivých operací. Díky nepovinnému odsazování je kód velmi přehledný, je možné si do něj psát i poznámky pomocí speciálního zápisu *<!--poznámka-->*.

Celý vzor je uzavřen do kořenového elementu *pattern*. Ukázka níže představuje vzor pro sloveso *být*, jehož název zavádíme pomocí atributu *id*. Můžeme použít také atribut *priority*, kterému po přiřazení hodnoty udáváme informaci o tom, jaká je pravděpodobnost, že slovo patří právě k tomuto vzoru a žádnému jinému (podrobněji viz kap. 8.3).

Uvnitř elementu *condition-list* definujeme ukončení slov, které ke vzoru přísluší. Pomocí atributu *charno* udáváme pozici (*charno="-1"* znamená první znak od konce). Atribut *value* přiřadí dané pozici zadané znaky. Atribut *type* rozhoduje, zda v hodnotě atributu *value* budou hodnotou konkrétní znaky (*type="equal"*), nebo definované třídy znaků (*type="class"*). V rámci jednoho vzoru je možné definovat několik základů (prázdný element *basechange*), ke kterým chceme připojovat zakončení. Jednotlivé základy/báze označujeme číslem (*no="1"*). Báze z lemmatu upravujeme pomocí atributu *operation*, jehož hodnota může např. přidávat (*append*) či ubírat (*chop*) (více viz kap. 1.2.4). I v tomto atributu je možné použít atributy *charno* a *value*, a na konkrétní pozice tak přidávat definované znaky.

Element *form* přidává k jedné z bází (*base="1"*) ukončení, které je zapsáno uvnitř vnořeného elementu *suffix*, a příslušnou morfologickou značku – vnořený element *tag*.

```
<pattern pos="verb" id="IMP_být" priority="100">
<!--vzor BÝT-->
    <condition-list lemma="yes">
    <!--imp: -Ø-->
        <condition charno="-1" type="equal" value="t" />
        <condition charno="-2" type="equal" value="ý" />
        <condition charno="-3" type="equal" value="b" />
        <basechange no="1" operation="chop" />
        <basechange no="1" operation="chop" />
        <basechange no="1" operation="append" value="u"/>
        <basechange no="1" operation="append" value="ď"/>
    </condition-list>
    <form base="1"><suffix></suffix><tag>k5mRp2nS</tag></form>
    <form base="1"><suffix>m,e</suffix><tag>k5mRp1nP</tag></form>
    <form base="1"><suffix>t,e</suffix><tag>k5mRp2nP</tag></form>
</pattern>
```

K popisu morfologických vlastností používáme brněnský atributivní tagset[5], je však možné jej převézt na libovolný jiný. U sloves zadáváme informace o slovním druhu (*k5*), módu (*mR*), osobě (*p2*) a čísle (*nS*). Nerozlišujeme kategorii vidu, protože je těžce (až nemožně) odvoditelná z ukončení infinitivu. Dále ve značce neinformujeme o spisovnosti či příznakovosti.

1.2 Nová řešení v popisu vzorů

Jak již bylo naznačeno, na základě vzorů, které vytvořila Dagmar Husárová, a při práci na vzorech nových jsme provedli v popisu vzorů několik změn. V této kapitole se budeme věnovat obecné charakteristice těch nejzásadnějších z nich.

1.2.1 Parciální vzory sloves

Na rozdíl od substantiv a adjektiv jsme ke slovesům zvolili ne zcela běžný přístup. Rozhodli jsme se upustit od tradičních vzorů, které popisují celé paradigma, a rozdělit je podle kategorie času, slovesného způsobu a rodu na 7 nezávislých skupin, tzv. parciálních vzorů pro tvary:[6]

a) přítomný čas (PRIT)
b) minulý čas/l-ové příčestí (MIN)
c) trpný rod/nt-ové příčestí (TRP)
d) Imperativ (IMP)
e) budoucí čas (BUD)
f) podmiňovací způsob (PODM)
g) přechodníky[7]

Budoucí čas skýtá dvě možnosti: Zaprvé se jedná o tvary slovesa být v budoucím čase (jediný vzor), které tvoří část analytického futura. Druhou variantou je syntetické futurum, které je tvořeno prefixací (*jít* > *půjdu*, *psát* > *napíšu, běžet* > *poběžím*) nebo se shoduje s prézentem dokonavých sloves (*kýchnout* > *kýchne*, *leknout* > *lekne*). Podmiňovací způsob je složený tvar (*spal bych*), který se skládá vždy z příslušných tvarů slovesa *být* (*bych*, *bys*, *aby*, *kdyby*…) a tvaru sloves v l-ovém příčestí. Je tedy tvořen pouze jedním vzorem pro sloveso *být,* druhá složka je označena samostatně jako MIN (*spal*). První čtyři skupiny jsou mnohem členitější a pro náš projekt stěžejní.

Vstup parciálních vzorů u všech skupin zůstává stejný, lemmatem je infinitiv. Výstupem pak bude opět zadaný výraz a k němu přiřazený alespoň jeden vzor z každé skupiny:

lemma:PRIT_vzor;MIN_vzor;TRP_vzor;IMP_vzor;BUD_vzor;PODM_vzor;

[5] Tagset je určitým způsobem definovaná a realizovaná sada značek, které nesou informace o (v tomto případě) gramatických a morfologických vlastnostech výrazů.
[6] Rozdělení na 7 subparadigmat je i v koncepci popisu vzorů Osolsobě (1996).
[7] Přechodníky jsme prozatím do systému nezavedli, uvádím je pouze pro úplnost slovesného paradigmatu.

Důvodem, pro zavedení parciálních vzorů namísto jednoho vzoru pro celé paradigma, bylo velké množství vzorů, které se od sebe v mnoha případech lišily jen několika tvary (často korespondujícími s tvary jednoho z parciálních vzorů). Už v mluvnici jsou si některé vzory velmi podobné. Například vzory *trpět* a *sázet* jsou ve všech tvarech stejné, výjimkou jsou tvary imperativu (*trp-*∅, *sáz-ej*). Tvary, které spadají do ostatních výše stanovených skupin, jsou totožné. Původní systém bez parciálních vzorů obsahoval různá pravidla pro vzory *objevit* a *obvinit*, které se od sebe liší pouze tvary trpného rodu (*objev-en*, *obvin-ěn*), vzory *končit* a *objevit* jsou různé v imperativu (*konč-i*, *objev-*∅) a mnohé další.

Tímto krokem jsme celkové množství vzorů nesnížili, naopak, ke každému lemmatu jich přiřazujeme 6krát více. Ovšem snížil se počet vzorů, které se během přiřazování ke vstupním výrazům nabízí. Navíc jsme tím získali možnost kombinovat jednotlivé parciální vzory a předpokládáme, že výsledky guesseru naznačí určité vztahy mezi nimi. Slovesné parciální vzory se často daří popsat v takové míře, že se k jednomu výrazu nabídne právě jeden dílčí vzor – přiřazení vzorů je v takových případech plně automatické.

1.2.2 Záplaty

Novinkou v projektu jsou také tzv. záplaty. Řeší malé vychýlení od pravidelného jevu zachyceného pravidly. Namísto toho, abychom pro nepravidelná slova, která jsou vyjma několika tvarů totožná s jedním ze vzorů, psali vzor nový, vytvoříme záplatu, která má dvojí vlastnost:

a) Přidává tvary, které nejsou v příslušném vzoru obsaženy.
b) Odstraňuje tvary, které v příslušném vzoru obsaženy jsou.

V každém případě záplata obsahuje informaci o tom, které vzory doplňuje či upravuje. O tom, zda záplata ubírá či přidává tvary, rozhoduje znaménko +/- před značkou daného tvaru.

Formálně zapisujeme záplatu podobně jako jakýkoliv vzor:

```
<pattern pos="noun" id="duch_Z" parent="archeolog">
    <condition-list lemma="yes">
        <condition charno="-1" type="equal" value="ch" />
        <condition charno="-2" type="equal" value="u" />
        <condition charno="-3" type="equal" value="d" />
    </condition-list>
    <form base="2"><suffix>á,c,h</suffix><tag>+k1gMnPc6</tag></form>
</pattern>
```

Název pravidla se od vzoru odlišuje pomocí označení *záplata_Z* za názvem (*id*). Záplata *duch_Z* upravuje vzor *archeolog*, který je zde uveden jako nadřazené pravidlo (*parent*). V charakteristice ukončení (*condition-list*) je uveden řetězec znaků „duch“ a pomocí značky (*+k1gMnPc6*) je uvedeno,

že pro všechna maskulinní životná substantiva s tímto ukončením se v lokálu plurálu přidá další tvar s novým sufixem (*-ách*).

1.2.3 Charakteristika hlásek a automatické alternace

Z diplomové práce Dagmar Husárové jsme převzali soubor s charakteristikou všech hlásek/znaků. Ty obsahují tyto základní informace:

- velké písmeno (malá písmena jsou výchozí, žádnou charakteristiku nemají)
- samohláska
  - krátká/dlouhá
- souhláska
  - měkká/tvrdá/obojetná

  + nevyužité:
  - znělá/neznělá + její neznělý/znělý protějšek
  - slabikotvorná
  - retná/sykavka/velára + laryngála

Tuto charakteristiku jsme ponechali i přesto, že obsahuje informace, které zatím nejsou v pravidlech využívány. Upravili jsme případné chyby a přidali další charakteristiky, které nám usnadnily popis pravidel.

Charakteristika *měkčící* usnadňuje rozlišit, zda po daném znaku může následovat *ě* (*b*, *d*, *f*, *m*, *n*, *p*, *t*, *v*). Tímto lze jednoduše odlišit např. imperativní vzory *chtít* (> *cht-ěj*) a *sít* (> *s-ej*).

Celá charakteristika je zapsána v souboru *.xml* a přiložena jako příloha B v elektronické podobě této práce.

1.2.4 Přidané operace

V určitých případech bylo třeba velmi rozsáhlého popisu, abychom pokryli veškeré alternace. Nejproblematičtější bylo popsat vzor *prosit*. U sloves, která podle něj tvoří imperativ, může dojít hned ke dvěma alternacím – krácení kořenného vokálu/diftongu (*á*, *é*, *í*, *ó*, *ů*, *ý*, *ou*) a měkčení (*d*, *t*, *n* graficky v infinitivu na *ď, ť, ň* před mezerou či nulovou koncovkou a koncovkami *-me, -te*). Kombinací těchto alternací vzniká nutnost popsat všech 21 možností. Díky novým operacím S*horten* a *SoftenA* lze celý problém shrnout do pouhých dvou definic, pro dlouhé vokály a pro dvojhlásku *ou*, kterou operace nepokrývá.

```
<condition-list lemma="yes">
    <!--ou.it-->
    <condition charno="-1" type="equal" value="t" />
    <condition charno="-2" type="equal" value="i" />
    <condition charno="-4" type="equal" value="u" />
    <condition charno="-5" type="equal" value="o" />
    <basechange no="1" operation="chop" />
    <basechange no="1" operation="chop" />
    <basechange no="1" operation="chop" charno="-3"/>
    <basechange no="1" operation="softenA" charno="-1"/>
</condition-list>

<condition-list lemma="yes">
    <!--VKit-->
    <condition charno="-1" type="equal" value="t" />
    <condition charno="-2" type="equal" value="i" />
    <condition charno="-3" type="class" value="souhláska" />
    <condition charno="-4" type="class" value="samohláska" />
    <basechange no="1" operation="chop"/>
    <basechange no="1" operation="chop"/>
    <basechange no="1" operation="softenA" charno="-1"/>
    <basechange no="1" operation="shorten" charno="-2"/>
</condition-list>
```

Jelikož je na příslušné pozici definována samohláska, vzorem projdou i slova s krátkým vokálem, operace v těchto případech neproběhne. Stejně tak je tomu u souhlásek, které se neměkčí.

Shrneme-li použité operace, základními jsou:

- *chop* (odebrání znaku)
- *append* (přidání znaku)

Nově přidané operace:

- *Shorten* (krácení dlouhé samohlásky)
- *SoftenA* (měkčení hlásek: *d* > *ď*, *t* > *ť*, *n* > *ň*)
- *SoftenB* (měkčení velár: k > *c*, *g* > *z*, *h* > *z*, *ch* > *š*)

Seznam alternací by bylo vhodné rozšířit o další alternace – alespoň z lingvistického hlediska. Prozatím jsme zavedli pouze takové alternace, které nám komplikovaly snahu o jednoduchý a přehledný popis vzorů.

1.3 Guesser

Dalo by se říci, že guesser je počítačový protějšek člověka – anotátora. Ruční anotace mají tu výhodu, že rodilý mluvčí jazyk zná a ve slovnících či příručkách si může správnost přiřazení vzorů ke slovu ověřit. Je-li však třeba anotovat statisíce slov, je jednak vysoce pravděpodobné, že dojde k chybám z nepozornosti, jednak je to časově velmi náročné i pro větší počet anotátorů. V této chvíli znamená guesser snadné a rychlé řešení. Bez znalostí a jazykové schopnosti je mu ale potřeba definovat nápovědu, navést jej, podle čeho se při rozhodování mezi vzory řídit. V době, kdy tato diplomová práce vznikala ještě nebyl guesser hotový, jeho vznik je závislý na našich ručně anotovaných datech. Není proto dosud zřejmé, na jakém principu bude pracovat. Nastíním alespoň prozatímní návrhy.

V první řadě počítáme s využitím neoznačkovaného korpusu. Získáme slovní tvary bez jakýchkoliv gramatických a morfologických informací a použijeme je jako kontrolu, zda se daný tvar v korpusu vyskytuje, či nikoliv.

Bez ohledu na značkování také můžeme získat informace pomocí kontextu, ačkoliv nelze počítat s tím, že budou stoprocentně spolehlivé. S velkou pravděpodobností úspěchu můžeme využít předpokladu, že po předložce bezprostředně následuje slovo obsahující informaci o pádu. Na základě předložek, které se pojí pouze s jedním konkrétním pádem, získáme seznam slov, kterým přisuzujeme příslušný pád. Další možností je využití shody mezi větnými členy. Stejně tak by bylo možné hledat i jiné pravidelné tendence v textu (např. genitivní vazba) a zkoumat jejich spolehlivost.

2 Využití korpusových dat během výzkumu

Cílem projektu je vytvořit morfologický slovník, nový, nezávislý na dalších aplikacích, postavený na neoznačkovaných datech, která nejsou ovlivněna případnými chybami. Spoléháme se proto jen na vlastní práci a k ověření svých tezí a výsledků využíváme především korpusová data, a to na dvou rovinách.

První rovina se týká samotného výzkumu. Při tvorbě vzorů využíváme méně radikální formu metody corpus-driven, což je přístup řízený korpusovými daty. Rozsáhlé texty v korpusech umožňují hledat skupiny slov, jejichž ukončení se chovají stejně (nebo podobně), a naopak výjimky. Počítáme však s možnými chybami ve značkování daného korpusu i s tím, že zkoumaný imperativ není v textech dostatečně doložen. Proto vzory postavené na informacích získaných čistě z korpusových dat v případě potřeby doplňujeme (na základě vlastních znalostí či během anotací) o chybějící části.

Na rovině anotací pracujeme s daty z brněnského korpusu czTenTen. Jak již bylo zmíněno, rozhraní pro anotace na základě těchto informací zvýrazní slovní tvary, které jsou v tomto korpusu obsaženy. V případě imperativů není tato nápověda zcela spolehlivá. Zaprvé proto, že imperativ se v psaných textech vyskytuje velmi zřídka (viz kap. 3.3.1). Zadruhé kvůli tvarové homonymii, jelikož se nápověda neopírá o značkování v korpusu (více o homonymii viz kap. 5).

Předpokládáme, že budou korpusová data využívána i při tvorbě guesseru, bude-li jej možné opřít o kontextová pravidla.

3 Infinitiv a imperativ v odborné literatuře

3.1 Slovo

Termín *slovo* je v lingvistice mnohdy obtížně definovatelný. V různých rovinách jazyka může označovat jinou (jinak velkou, obecnou) jednotku jazyka. Čechová (2011, s. 75) charakterizuje slovo jako základní, centrální jazykovou jednotku, a to jak z hlediska systému slovních druhů, tak z hlediska slovotvorby a gramatiky. Se slovem a jeho tvary operuje tvarosloví (morfologie) i skladba (syntax). Dle definice (Čechová, 2011, s. 75) je „základním prostředkem pro vyjadřování pojmenování a označování vůbec".

Čechová (2011, s. 75) ve své knize o češtině vymezuje slovo ze šesti hledisek:

a) **sémantické**: Slovo je jednotka slovní zásoby, která má věcný, lexikální význam.
b) **grafické**: Slovo představuje celek tvořený několika spojenými grafémy. Tento celek je oddělený od ostatních slov po obou stranách mezerou.
c) **fonetické**: Slovo je útvar složený z jednoho a více fonémů, sjednocený přízvukem a oddělený od ostatních slov potenciálními pauzami. Upozorňuje však i na slova bezpřízvučná (tzv. předklonky a příklonky), která sdílí přízvuk se slovem přízvučným, k němuž se kloní.
d) **morfologické**: Tvar slova může obsahovat i více slov grafických.
e) **větné**: Plnovýznamová slova mají funkci větných členů nebo jsou větnými ekvivalenty. Ta ostatní se pouze podílejí na výstavbě a na pragmatickém významu výpovědi.
f) **kontextu**: Ať už slova jednoznačná, nebo mnohoznačná, v konkrétním kontextu mívají jeden určitý význam.

Pro naše účely vymezíme *slovo* jako slovo *grafické*, jehož detekce je strojově nejsnazší. Také v případě analytických složených tvarů pracujeme s jeho jednotlivými částmi samostatně, např. *byl bych býval* zpracujeme jako trojici, v níž se vyskytuje tvar od lemmatu *být* v minulém čase, tvar lemmatu *být* v kondicionálu a tvar lemmatu *bývat* v minulém čase.

Další nástroje, které morfologické slovníky využívají, jako například morfologický analyzátor, pracuje s pojmem *kontextové slovo*. Obecně lze popsat způsob morfologické analýzy (běžně užívaný v českém prostředí) tak, že z více interpretací nabízených slovníkem vybere na základně konkrétního kontextu právě jednu správnou, která danému kontextu odpovídá.

3.1.1 Stavba slova

V následujících několika odstavcích se budeme věnovat slovu a jeho vnitřní struktuře, která je z lingvistického hlediska pro tvorbu odvozených tvarů a hledání zákonitostí velmi důležitá.

Náš popis zúžíme na formální tvarosloví sloves. Budeme přitom vycházet z popisu uvedeného v příručce M. Čechové (Čechová, 2011, s. 185–87). Tvarosloví sloves se od ostatních slovních druhů

liší způsobem flexe. Jako jediný slovní druh vytvářejí paradigmata konjugací, nikoliv deklinací. Z toho vyplývá, že pro tvorbu slovesných tvarů používáme odlišnou, pro slovesa specifickou sadu zakončení[8].

Jádrem slovního tvaru je tzv. *tvarotvorný základ*. Ten nese věcný význam slova, jedná se o jeho lexikální složku. *Tvarotvorný formant* je pak složka gramatická. Třetí důležitou částí slova je *kmenotvorná přípona*, která spolu s tvarotvornou příponou tvoří *kmen* slova. Na konci slova se nachází *tvaroslovná přípona,* často označovaná spíše termínem *koncovka*. Kmenotvorná přípona a koncovka společně tvoří *tvarotvorný formant*. Pro přehlednost přikládám tabulku (tab. 1), která jednotlivé části slova názorně ilustruje.

Tabulka 1: Stavba slova *kryješ*

Tvarotvorný základ	**tvarotvorný formant**	
	Kmenotvorná přípona	**Konj. formant/koncovka**
kry-	-je-	-š
kmen		

Ve většině mluvnic se dočteme o třech různých kmenech, dle kterých tvoříme slovesné tvary. Jsou to kmen prézentní, minulý a infinitivní. Zdenka Rusínová (1982, s. 76) však k této problematice přistupuje podrobněji. Na základě množiny všech tvarů slovesných a množiny všech typů proměn kmene při tvorbě tvarů vymezuje pět slovesných kmenů:

a) **kmen imperativní** pro tvary imperativu aktivního,
b) **kmen prézentní aktivní** pro tvary indikativu prézentu aktivního,
c) **kmen pasivní** pro tvary opisného pasíva,
d) **kmen minulý** pro tvary s l-ovým základem a
e) **kmen infinitivní** pro tvary s infinitivem.

Rusínová (1982, s. 79) svoji teorii o slovesných kmenech završuje takto: „Tato soustava kmenů je maximální, od většího počtu kmenů své tvary netvoří žádné sloveso, naopak v řadě případů je počet kmenů menší.“ Některé různé kmeny se totiž mohou a nemusí formálně lišit (např. inf. *prosi-t* a min. *prosi-l*, ind. *ber*-u a imp. *ber*-Ø). Její konkrétnější dělení sloves se blíží našemu pojetí, kdy členíme slovesné tvary do sedmi parciálních skupin. Nepovažujeme však za stěžejní kritérium kmen ani žádnou další konkrétní formální složku slovesa. Zajímáme se především o infinitivní ukončení a hláskové/grafické alternace.

Vedle jednoduchých slovesných tvarů, které pro svoji úplnost nepotřebují žádná doplnění, uvádí Čechová (2011, s. 235) tvary složené a zvratné.

[8] Žádné jiné slovní tvary nemají totožná ukončení s imperativy.

Složené tvary jsou zcela závislé na tvarech jednoduchých, kterými jsou tvořeny. Jedná se o spojení jmenného tvaru plnovýznamového slovesa (příčestí nebo infinitiv) a jednoduchého či složeného tvaru pomocného slovesa *být*. Tato skutečnost umožňuje strojově přistupovat ke slovesným tvarům jako ke grafickým slovům a jednotlivé části složených tvarů analyzovat odděleně, aniž by pro morfologickou analýzu chyběly jakékoliv zásadní informace o slovu, na rozdíl například od syntaktické roviny jazyka, v níž je třeba určit složený slovesný tvar jako celek.

Zvratné tvary jsou jednoduché nebo složené tvary doplněné o zájmeno *se* nebo *si*. Nikoliv všechna slova takové doplnění dovolují. Jsou slovesa, která nejsou zvratná nikdy (*mít, být, žít*) a naopak tzv. reflexiva tantum, což jsou slovesa, která existují pouze ve spojení se zvratným zájmenem *se*/*si* (*bát se, snažit se, troufat si*). U velké skupiny sloves existují obě možnosti, často však mají vliv na význam celé výpovědi (*trápit* × *trápit se, hrát* × *hrát si*).

3.2 Sloveso

„Mluvnickým i významovým centrem české věty je sloveso, neboť určuje obsazení určitých pozic ve větě a organizuje tím její stavbu. Sloveso je také z hlediska soustavy tvarů a mluvnických kategorií nejbohatší. Avšak z hlediska sémantického často vyžaduje doplnění“ (Čechová, 2011, s. 170).

Komárek (1986, s. 129) svoji kapitolu o slovesech uvádí těmito slovy: „Základní slovní druh sloveso (verbum) vyjadřuje dynamický příznak substančních jevů probíhající v čase, tj. především děje, a to děje nemutační a děje mutační (tzv. události), ale také tzv. neděje (statické významy).“

Jak Komárek (1986), tak Čechová (2011, s. 235) vyzdvihují sloveso jako jediný slovní druh, jehož tvary mohou být tvořeny i více než jednou výrazovou jednotkou. Složené tvary se tedy nevyskytují u žádných tvarů jiných slovních druhů.

O jedinečnosti slovesa mezi slovními druhy se Komárek (1986, s. 129) zmiňuje ještě jednou v souvislosti s jeho schopností funkčně i tvarově zaujímat pozice dalších základních slovních druhů – substantiv, adjektiv a adverbií. Díky tomu má nejvýznamnější komplex větotvorných funkcí. Z toho vyplývá, že sloveso obecně je v textech hojně zastoupeno, a v korpusových textech je tak možné získat velmi rozsáhlá data pro jeho zkoumání.

3.2.1 Gramatické kategorie sloves

Slovesné tvary mají schopnost vyjadřovat mluvnické významy, gramatické kategorie. Na základě schopnosti vyjádření těchto kategorií dělíme slovesné tvary na určité neboli finitní a neurčité neboli infinitní – infinitiv činný, přechodník přítomný činný, přechodník minulý činný, příčestí cinné (l-ové) a příčestí trpné (t/n-ové) (Čechová, 2011, s. 235). Finitní tvary vyjadřují všechny slovesné mluvnické kategorie, jimiž jsou slovesná osoba, číslo, způsob, čas, slovesný rod a vid. Neurčité tvary vyjadřují pouze slovesný rod a vid (např. c*hválit* × *pochválit* × *být (po)chválen*). Čechová (2011, s. 170)

za základní, nepříznakový tvar sloves považujeme tvar infinitivu, který nenese příznak ani jedné ze dvou zmíněných kategorií.

3.2.1.1 Kategorie slovesného způsobu (modus)

Rusínová (1982) ve své *Současné české morfologii* charakterizuje jednotlivé gramatické kategorie sloves. Slovesný způsob neboli modus vyjadřuje, jaký je vztah autora promluvy k realizaci přísudku nebo větného základu, nebo jinak řečeno je to autorovo hodnocení platnosti predikačního příznaku" (Rusínová, 1982, s. 71). Rozlišuje tři základní významy a každý je zastoupen jedním slovesným modem. Způsob oznamovací neboli *indikativ*, který je nepříznakový, vyjadřuje reálnost děje, způsob přací neboli *kondicionál* vyjadřuje možnost děje a konečně způsob rozkazovací neboli *imperativ* vyjadřuje nutnost děje. Čechová (2011, s. 173–175) řadí kategorii způsobu (na rozdíl od dosavadních mluvnic) před slovesný čas. Kategorii času chápe jako závislou na slovesném způsobu, tedy jako její podkategorii. Čas se totiž u všech určitých tvarů neprojevuje, například u imperativu kategorii času neurčujeme. Děj vyjádřený imperativem nastává až po okamžiku promluvy.

3.3 Imperativ

Imperativ je jeden ze slovesných způsobů, který vyjadřuje rozkaz, zákaz či přání. „Je to jediný modus, který je v češtině v aktivu, pro vyjádření všech rysů, které jsou pro něj dostupné, vybaven sufixem ve vnitřní struktuře slovesa, tedy pro strukturalistickou jazykovědu typická morfologická kategorie: piš, dělej, mysli" (NESČ, 2016, s. 675). Formálně se vyjadřuje speciálními koncovkami, zpravidla není homonymní s dalšími slovesnými tvary (až na výjimky, jako je např. *zvát* > *zvěte*, ale i *zvete*, což je tvar shodný s indikativem).

Česká morfologie a korpusy (Osolsobě, 2014, s. 225) užívá imperativ dokonce jako ukazatel v algoritmu pro určování slovesných tříd. U řady (především frekventovaných) sloves je jeho tvoření intuitivní. Určení imperativních tvarů však nemusí být vždy jednoznačné. Možná i proto bývá tento tvar uveden ve výkladových slovnících. Sami rodilí mluvčí dělají v imperativu chyby. Děje se tak jednak kvůli nesprávně určenému základu (záměna vidu, př. *přiblížit* > *přibližuj* namísto *přibliž*), jednak kvůli hláskovým alternacím (*přihlaš* namísto *přihlas*) a jednak kvůli různým kolísáním a přechodům mezi vzory (*shromáždi* i *shromažď*). Výjimky jsou naopak zřejmé, protože jsou často užívané (*být* > *buď*, *jíst* > *jez*).

Tvorba imperativu se v některých případech může zdát jako nesnadný úkol i běžnému uživateli českého jazyka. Dalším důvodem je jistě také to, v jakém poměru se vyskytuje v jazyce/textech. Ke zjištění takové informace použijeme korpusy. Nahlédli jsme do korpusů SYN2000, SYN2005, SYN2010 a SYN2015. Ve všech případech jsme se s malými odchylkami dopočítali velmi podobných čísel. V každém korpusu je průměrně asi 121 milionů pozic celkem, z toho je asi 17 milionů pozic zastoupeno slovesy, imperativ pak tvoří skupinu okolo 240 000 pozic. Převedeno na procenta tvoří slovesa v těchto

psaných korpusech kolem 14 %, imperativy pouhých 0,2 % ze všech slov korpusu a asi 1,4 % ze slovesných pozic. Mezi mluvenými korpusy Českého národního korpusu je označkovaný pouze *Korpus prezidentských projevů Speeches*. Nejedná se o běžnou mluvu, ale zmiňujeme jej alespoň pro zajímavost a srovnání s psanými texty. Slovesné tvary jsou zde zastoupeny 13 %, imperativy opět 0,2 %, v poměru ke všem slovesným tvarům 1,7 %. Objektivnější by pro tuto sondu byly korpusy řady ORAL, které však nejsou opatřeny morfologickým značkováním.

Uvedená čísla předznamenávají komplikace při zkoumání imperativů a pokusech o jejich automatickou tvorbu. Zaprvé nabízí malé množství dat k učení a hledání pravidelných znaků, zadruhé nenabízí velký prostor pro detekci chyb. Během výzkumu je třeba na tento fakt myslet a hledat i další opěrné zdroje v odborné literatuře a slovnících.

3.3.1 Imperativ a jeho způsoby vyjádření

Důvodem, proč se v jazyce imperativ vyskytuje v menší míře, je také možnost užití alternativních způsobů vyjádření. Vedle klasických syntetických tvarů imperativu máme k dispozici několik dalších konstrukcí, které zachovávají stejnou komunikační funkci, ale v rámci komunikační strategie slouží například ke zjemnění apelu, jako zdvořilejší varianta. Díky těmto opisným imperativům je možné apelovat také na 1. osobu singuláru a 3. osobu singuláru i plurálu, pro které neexistuje speciální imperativní tvar. V Novém encyklopedickém slovníku (2016, s. 2067–2068) se v hesle *Výzvová výpověď* dočteme, že „všechny druhy výzev kromě ‚povelu‘ lze realizovat a indikovat jazykově“.

Štícha věnuje pozornost imperativu/apelovým komunikačním funkcím ve svých knihách *Akademická gramatika spisovné češtiny* (2013, s. 768–775) a *Kapitoly z české gramatiky* (2011, s. 552–565). V první uvedené knize vymezuje, jakým komunikačním funkcím je apel nadřazen. Jsou to výzva, žádost, prosba, zákaz, příkaz, rozkaz, hrozba, varování aj. Štícha (2013, s. 765) však také poukazuje na fakt, že ačkoliv „existuje mnoho různých jazykových prostředků (morfologických, slovosledných, syntaktických, lexikálních a intonačních), jimiž jsou různé komunikační funkce signalizovány, neexistuje jednoznačný vztah mezi jazykovou formou věty a její specifickou aktuální komunikační funkcí. Naopak: jde tu o mimořádně spletitou asymetrickou síť formálně-obsahových vztahů, v níž se jazykové formy a jejich komunikační funkce nejrůznějším způsobem kříží a přesahují.“

Jeho výpis vyjádření apelu přerozdělíme na dvě části, s užitím tvarů imperativu a s užitím jiných jazykových konstrukcí, případně doplníme o informace čerpané z další literatury. Přehled alternativních konstrukcí omezíme pouze na takové konstrukce, které lze převézt na věty s imperativní formou, aniž by byl pozměněn její komunikační záměr. Právě tato možnost užití jiných jazykových prostředků je totiž jedním z hlavních důvodů, proč je výskyt imperativních tvarů redukován, a jeho zkoumání je tak z velké části omezeno, ačkoliv výzva jako taková je v komunikaci velmi častá.

3.3.1.1 Užití imperativních tvarů

1) Věty s imperativem

„Obecnou větnou formu apelovou tvoří v jazykovém systému češtiny především přísudek v gramatickém modu imperativu. Jím někoho k něčemu (ve velmi obecném smyslu) vyzýváme.“ (Štícha, 2013, s. 770)

Sloveso v imperativu může mít jak dokonavý, tak nedokonavý vid. V záporném imperativu se vid pravidelně převádí na vid nedokonavý.

Př.: Dej mu to.

Př.: Nedávej mu to.

2) Syntetický imperativ + infinitiv

Jedná se o konstrukci, která formu klasického (tj. syntetického) imperativu obsahuje, získává ale charakter imperativu analytického.

Základní syntetický imperativ (*Uklid'!*) je zde zastoupen dvojicí, v níž je prvním článkem imperativní tvar slovesa. To částečně ztrácí svůj původní význam a plní funkci ukazatele apelu. Druhou část představuje infinitiv, ve kterém se koncentruje diktum a jehož význam je pro tuto frázi stěžejní. Jedná se o kombinace imperativu slov *pojď(te)*, *běž(te)*, *koukej(te)*, *hleď(te)*, *chraň(te)*, *opovaž(te) se*, *rač(te)* a libovolného infinitivu. (Štícha, 2011, s. 562–564)

Př.: Běž spát.

Př.: Opovaž se promluvit!

Př.: Račte vstoupit.

Př.: Koukejte vypadnout!

3) Konstrukce *Pojď / Pojďte* + vedlejší věta

Souvětné konstrukce, kde hlavní věta obsahuje imperativní tvary *pojď / pojďte* a věta vedlejší je založena na přísudku v podobě 1. osoby plurálu indikativu prézentu.

Př.: Pojďte, pustíme se do práce.

4) Varování či hrozby: Věty s imperativem a částicí *jen*

V případě vět tohoto typu se vyjadřuje pomocí částice *jen* a afirmativním tvarem imperativu negativní význam přísudku.

Př.: Jen toho psa drážděte!

5) **Analytický imperativ *buď(iž)/buďte(ž)* + příčestí trpné**

Tento typ opisného imperativu se objevuje ve 2. a 3. osobě singuláru i plurálu. Je však velmi vzácný a dnes má již pouze knižní charakter. Vyskytuje se v náboženských a liturgických textech a archaických pozdravech.

Př.: Buď posvěceno jméno tvé.

Př.: Buď vítán. Buď pochválen pan Ježíš Kristus.

6) **Ustálená spojení**

Existují ustálená spojení, která obsahují původní imperativní tvary, které mnohdy ztratily svůj původní význam.

Př.: Buď jak buď.

Př.: Děj se co děj.

3.3.1.2 Jiné jazykové konstrukce (jako alternativa ke tvarům imperativu)

1) **Užívání performativních sloves**

Jedná se o výpovědi s formou oznamovací věty, v níž je komunikační (ilokuční) funkce věty realizována a určena užitím ilokučních sloves, *performativ*, v indikativu. V případě méně silné výzvy lze využít namísto indikativu tato performativa v kondicionálu (např. *navrhoval bych*, *vytkl bych*). (NESČ, 2016, s. 459–460)

Štícha (2013, s. 768) uvádí následující komunikační funkce s užitím performativ:

a) **Rada a doporučení**

Př.: Doporučuji ti, abys to přijal / to přijmout.

Př.: Radím ti, abys dorazil včas.

b) **Návrh**

Sloveso *navrhovat* bývá mnohdy jediný způsob, jak explicitně vyjádřit, že něco navrhujeme.

Př.: Navrhujeme, aby se jednání odložilo.

c) **Pozvání**

Př.: Zvu tě na večeři.

d) **Výzva v užším smyslu**

Př.: Prosím/Žádám/Vyzývám vás, abyste odešel.

e) Žádost a prosba

Př.: Prosím/Žádám tě, abys mi pomohl / o pomoc.

f) Zákaz

Př.: Zakazuji ti tam chodit! / Tam nesmíš chodit!

g) Příkaz a rozkaz

Př.: Nařizuji/Rozkazuji/Přikazuji/Poroučím ti, abys ji nechal na pokoji.

h) Hrozba a varování

Př.: Varuji vás, posedlost při prosazování sociálněpolitických reforem se nevyplácí.

2) Věty s indikativem nebo kondicionálem větné formy oznamovací

a) Věty s přísudkem ve 2. osobě indikativu prézentu nebo futura

„V imperativním významu se běžně užívá prézentní formy sloves dokonavých nebo futura sloves nedokonavých.“ (Štícha, 2013, s. 770)

Př.: Tady si budeš s babičkou povídat, jo?

Př.: Po obědě si aspoň na hodinu lehneš.

b) Věty s přísudkem v 1. osobě indikativu prézentu nebo kondicionálu

Tvary singuláru se často vyjadřuje jistá dávka hrubosti a skutečné či simulované nadřazenosti. Výrazy tohoto typu často využívají prostředky nespisovné či obecné češtiny.

Př.: Tak makám a neflákám se!

Formulací s 1. osobou plurálu mluvčí s jistou autoritou vůči adresátovi často vyzývá k zahájení či ukončení nějaké činnosti.

Př.: Tak děti, jdeme. Končíme.

Indikativu v 1. osobě plurálu se užívá často také v návodech, receptech a různých instrukcích. (NESČ, 2016, s. 2067)

Př.: Těsto přikryjeme utěrkou a necháme hodinu odpočívat.

c) Věty s přísudkem ve 3. osobě indikativu prézentu

Tyto věty jsou typické v situacích, v nichž je mluvčí autoritou vůči kolektivu a obrací se na jednoho jeho člena.

Př.: K tabuli půjde Novák.

3) **Věty s indikativem nebo kondicionálem větné formy tázací**

Různé modifikace věty tázací s indikativem nebo kondicionálem v přísudku jsou patrně nejfrekventovanější komunikační formou výzvy.

Př.: Otevřel/Neotevřel byste prosím vás okno?

Př.: Můžete/Mohl byste prosím vás otevřít okno?

Př.: Neposadíte se? / Nechcete se posadit?

4) **Infinitiv**

Infinitiv se jako expresivní prostředek užívá v autoritativních výzvách, rozkazech, příkazech apod.

Př.: Vyčurat a spát!

Další význam infinitivu bývá význam výtky, v němž mluvčí vyjadřuje nařčení. Jedná se o hovorový prostředek.

Př.: Takhle mluvit před matkou!

5) **Nominální konstrukce**

Nominálních výzev se (pro jejich strohost) využívá jako prostředku důrazné a expresivní výzvy.

Př.: Ven! Rychle!

Př.: Bez řečí!

6) **Speciální výzvové konstrukce a/nebo funkce**

a) **Věty s obecně výzvovou funkcí: Analytický imperativ *ať/nechť* + indikativ**

O apelativních větách s užitím částic *ať, nechť* píše Štícha (2011, s. 556–559) podrobněji ve své knize *Kapitoly z české gramatiky.*

Tento způsob opisu imperativu je velmi častý, mimo jiné nahrazuje chybějící formu imperativu pro 3. osobu – je v tomto případě základním gramatickým prostředkem pro vyjádření výzvy 3. osoby.

Př.: Ať mi nechodí na oči!

Konstrukce se užívá i namísto syntetického imperativu, ačkoliv není tak frekventovaná jako u 3. osoby. *Ať + 2. osoba indikativu* se dle Štíchy (2011, s. 556) užívá asi čtyřicetkrát méně než konstrukce se 3. osobou.

Př.: Do půlnoci ať jsi/jste doma!

Konstrukce *Ať + 1. osoba plurálu indikativu* má frekvenci užití asi dvěstěkrát menší, než je frekvence konstrukcí se 3. osobou.

Př.: Ať nezmeškáme vlak!

Nejméně užívaná je konstrukce *Ať + 1. osoba singuláru indikativu*. V takovém případě je mluvčí zároveň adresátem děje. Používá se především ve speciálních případech — v pohádkách, anekdotách, vtipech apod.

Př.: Ať to mám z krku. Ať se propadnu, jestli…

Ve všech zmíněných případech podle Štíchy (2011, s. 559) platí, že konstrukce s částicí *nechť* jsou asi stokrát méně frekventované, než jim odpovídající konstrukce s částicí *ať*.

b) Věty s komunikační funkcí návrhu

i) Konstrukce typu *Co + kdyby(ch/chom/ste)/aby(ch/chom/ste)* + příčestí minulé

Obvykle se v úzu vyskytuje varianta kondicionálu s *kdyby*, varianta s *aby* je příznakovější a méně frekventovaná.

Př.: Co kdybychom ho pozvali?

Př.: Co abychom šli spolu na pivo?

ii) Konstrukce typu *co* + infinitiv

V rámci těchto konstrukcí se často vyskytují zesilující částice, např. t*akhle, třeba* apod.

Př.: Co takhle dát si špenát?

c) Některé typy vět s komunikační funkcí námitky

i) Částice *co* v začátkovém větném postavení + kondicionál + klesavá intonace

Významově se jedná o razantní zamítnutí nějaké výzvy, kterého se běžně užívá v hovorovém jazyce.

Př.: Co bys tam chodil!

ii) Konstrukce s tázacím slovem a infinitivem + klesavá intonace

Př.: Proč ho poslouchat!

d) Komunikační funkce výčitky, napomenutí apod.: Větná konstrukce s počátkovou částicí *že*

Př.: Že takhle můžeš mluvit!

e) Některé typy vět s komunikační funkcí varování či hrozby

i) Konstrukce s *ne* + *aby* v počátkovém větném postavení

Př.: Ne aby vás napadlo opisovat!

ii) Větná konstrukce s indikativem a počátkovou částicí *že*

Př.: Že tě ten pes kousne!

3.3.2 O imperativu v (odborné) literatuře

Jak píše v *Mluvnici češtiny 2* Komárek (1986, s. 417): „Imperativ jakožto typická apelová kategorie jazyka má i specifickou morfémovou stavbu. Formálně imperativní tvary existují v 2. osobě singuláru, která je nepříznakový tvar imperativu. Dále pak 1. a 2. osoba plurálu, které tvoříme od základního tvaru přidáním tzv. a*glutinačních osobních morfů*, které se shodují u všech sloves: pro 1. osobu plurálu *-me*, pro 2. osobu plurálu *-te*."

Různé mluvnice přistupují k popisu tvaru imperativu různými způsoby. Komárek (1986, s. 417) uvádí dvě možné varianty imperativního morfému (viz tabulka 2):

Tabulka 2: Varianty imperativu dle *Mluvnice češtiny 2*

ukončení	**lemma**	**pomocný tvar**	**2. os. sg.**	**2. os. pl.**
-Ø	vést	ved-Ø-ou	veď-Ø-Ø	veď-Ø-Ø-te
	házet	ház-ej-í	ház-ej-Ø	ház-ej-Ø-te
-i/ě(e)	třít	tř-Ø-ou	tř-Ø-i	tř-Ø-e-te

V *Příruční mluvnici* (Karlík, Nekula a Rusínová 2012, s. 322) jsou popsány formální prostředky pro tvorbu imperativu pomocí tří druhů koncovek, které se přidávají k prézentnímu kmeni (viz tabulka 3):

Tabulka 3: Varianty imperativu dle *Příruční mluvnice*

ukončení	**lemma**	**pomocný tvar**	**2. os. sg.**	**2. os. pl.**
-Ø	nést	nese	nes-Ø	nes-Ø-te
	minout	mine	miň-Ø	miň-Ø-te
-i	umřít	umře	umř-i	umř-e-te
	řvát	řve	řv-i	řv-ě-te
-ej	volat	volá	vol-ej	vol-ej-te

Nejpodrobněji popisuje ukončení imperativních tvarů Rusínová (1982, s. 76), podle které tvoříme imperativy pomocí kmene imperativního a přidáním šesti druhy alomorfů:

Tabulka 4: Varianty imperativu dle Rusínové (1988, s. 76)

ukončení	**lemma**	**Pomocný tvar**	**2. os. sg.**	**2. os. pl.**
-Ø	nést	nes-	nes-Ø	nes-Ø-te
-j	krýt	kry-	kry-j	kry-j-te
-uj	kupovat	kup-	kup-uj	kup-uj-te
-ˇ	minout	min-	miň	miň-te
-ej	dělat	děl-	děl-ej	děl-ej-te
-i	tisknout	tiskn-	tiskn-i	tiskn-ě-te

V našem systému vzorů vycházíme z vlastní predikce, že lze slovesa na základě ukončení imperativu rozdělit do tří základních skupin: -Ø, *-i*, *-j*. Pracujeme s jediným výchozím tvarem, a to s infinitivem. Nemáme k dispozici žádný další opěrný tvar, jako je tomu v mluvnici, která se opírá o 3. osobou singuláru indikativu prézentu, podle níž imperativ tvoří.

Tvoření imperativu, respektive celá morfologická flexe je strojovému zpracování cizí. Podobně je tomu i u cizinců, osob, které nejsou rodilými mluvčími českého jazyka, jejich jazyková intuice tak nezahrnuje zabudované základy české mluvnice. Proto jsme (alespoň pro zajímavost) nahlédli do tří učebnic pro cizince.

V knize *Čeština pro cizince: základní kurs* Karly Hronové a Milady Turzíkové (1993) je imperativ a jeho tvoření popsáno ve 12. lekci (z 25 lekcí) následovně (Hronová, Turzíková, 1993, s. 164–165):

Formy imperativu se odvozují od 3. osoby plurálu, různé koncovky se přiřazují podle ukončení základu:

a) Jedna souhláska: -Ø, *-me*, *-te*
b) Dvě souhlásky: *-i*, *-ěme*/*-eme*, *-ěte*/*-ete*
c) *-ají*, *-ejí*/*-ějí*: *-ej*, *-ejme*, *-ejte*

Nepravidelně jsou tvořeny imperativy od následujících sloves:

Tabulka 5: Nepravidelné tvoření imperativů v *Češtině pro cizince* (Hronová, Turzíková, 1993)

1. osoba sg.	**3. osoba pl.**	**imperativ (2. osoba sg.)**
jsem	jsou	buď
jím	jedí	jez
odpovím	odpoví	odpověz
stojím	stojí	stůj
mám	mají	měj
jdu	jdou	jdi/pojď

V učebnici *Adam a Eva v Českém ráji* (Froulíková, 2008, s. 103–109) nacházíme teorii o tvorbě imperativu až ve 13. lekci (z celkem 15). Výchozími vzory jsou infinitiv a 3. osoba plurálu. Tvoření imperativů uvádí velmi stručně (Froulíková, 2008, s. 105):

Tabulka 6: Tvoření imperativů v učebnici *Adam a Eva v Českém ráji*

1. osoba sg.	**3. osoba pl.**	**imperativ (2. osoba sg.)**
prosit	pros-í	pros, prosme, proste
číst	čt-ou	čti, čtěme, čtěte
čekat	ček-ají	čekej, čekejme, čekejte

Vedle toho uvádí i tvorbu od sloves nepravidelných (*být, mít, jíst, vědět, vidět, odpovědět, jít, jet*).

Do třetice uvádíme *Základní kurs češtiny pro cizince* (Šlosar a Rusínová, 1994), který byl vydán na půdě Filozofické fakulty Masarykovy univerzity. K infinitivům tato učebnice přistupuje jako běžné mluvnice, jeho tvorbu uvádí v rámci celých paradigmat jednotlivých tříd.

Imperativy jsou v učebnicích pro cizince uvedeny až v pozdějších lekcích. Nerozlišují slovesa na základě slovesných tříd, ale přistupují k jednotlivým tvarům izolovaně (srov. parciální vzory viz kap. 1.2.1). Dalším důkazem, že je tvoření imperativu pro nerodilé mluvčí složité, je velmi stručný popis principu tvorby, učebnice jsou zaměřeny spíše na praktická cvičení.

3.3.3 Imperativ v rámci mluvnického systému vzorů

Různé mluvnice uvádějí různé přístupy k dělení sloves na jednotlivé skupiny zastoupené vzory. Výchozím kritériem je kmen slovesa. Vybrali jsme knihu Františka Štíchy *Akademická gramatika spisovné češtiny* (2013), která uvádí nejčastější a ve školách nejrozšířenější přístup – určování dle kmene prézentního. Druhý nejčastější je pak jako rozřazovací prostředek kmen minulý. Dělení sloves do tříd podle infinitivního/minulého kmene popisuje *Příruční mluvnice češtiny* (Karlík, Nekula a Rusínová, 2012, s. 326–332).

V následujícím textu se s oběma přístupy stručně seznámíme. Nebudeme popisovat zdaleka celou charakteristiku systémů, zaměříme se pouze na tvorbu imperativů podle příslušných vzorů. Nezohledníme tedy případné přechody mezi třídami, hláskové alternace a další jevy, které se odrazí ve tvarech jiných než v imperativu.

Tabulka 7: Přehled slovesných kmenů

Kmen prézentní	**Kmen minulý**	**Kmen infinitivní**
nes- (e)	nes- (l)	nés- (t)
ber-	bra-	brá-
maž-	maza-	maza-
peč-	pek-	péc-
tř-	tře-	tří-
tiskn-	tisk(-nu)-	tisknou-
min-	minu-	minou-
tn- / začn-	ťa- / zača	tnou-
kryj-	kry-	krý-
kupuj-	kupova-	kupova-
pros- (í)	prosi-	prosi-
trp-	trpě-	trpě-
sáz-	sáze-	sáze-
vol- / děl- (á)	volá- / dělá-	vola- / děla-

3.3.3.1 Slovesné tvary podle kmene prézentního

Jak již bylo zmíněno, nejčastější dělení sloves do slovesných tříd je založeno na kmeni prézentním. Ten získáme ze 3. osoby singuláru indikativu prézentu aktivního: *nes-e*, *tiskn-e*, *kupuj-e*, *pros-í*, *děl-á* (Rusínová, 1982, s. 77). Tento kmen i s příslušnou kmenotvornou příponou lze získat také odtržením koncovky 2. osoby singuláru nebo 1. a 2. osoby plurálu indikativu prézentu aktivního: *nes-e-š/-me/-te* atd. (Štícha, 2013, s. 452)

Na základě prézentního kmene dělíme slovesa do pěti slovesných tříd. Kmenotvorné přípony, které jsou základním charakteristickým znakem každé třídy, jsou: *-e*, *-ne*, *-(u)je*, *-í*, *-á*. Základní informace o třídách a imperativu, který je tvořen na jejich základě, uvádíme podle popisu Františka Štíchy (2013, s. 456–496).

První třída (*-e-*)

Třída s kmenotvornou příponou *-e* je zastoupena pěti vzory. Ty se liší dílčími tvary, a to v imperativu, který je tvořen od téhož kmene, dále na základě druhotného dělení podle minulého a infinitivního kmene v dalších příslušných tvarech. Tato třída definuje slova, u kterých je kmenotvorná přípona připojována k základu zakončenému souhláskou.

1) **Slovesa typu *nes-e***

Podle vzoru *nese* časujeme slovesa původní. V tvarech odvozených od minulého kmene je kmenotvorná přípona nulová (*nes-Ø-l*).

Imperativ se tvoří přidáním nulové kmenotvorné přípony (*nes-Ø*, *lez-Ø*, pas-Ø, *spleť-Ø*). U sloves se základem ukončeným na *-d/-t* dochází k měkčení na *-ď/-ť*.

2) **Slovesa typu *peč-e***

Do této podtřídy patří slovesa základová a odvozená předponami (*u-péct, za-tlouct* apod.) Kmenová přípona v minulém a infinitivním kmeni je u těchto sloves nulová (*pek-Ø-l*).

Imperativ je tvořen pomocí koncovek -Ø, *-te* (*peč-Ø*, *tluč-Ø*).

3) **Slovesa typu *ber-e***

Podvzor *bere* seskupuje slovesa, jejichž minulý kmen obsahuje kmenotvornou příponu *-a-* nebo infinitivní kmen *-á-* (*br-á-t* > *br-a-l* > *ber-e*). Patří sem slovesa základová a předponová, která krátí kmenotvornou příponu základu (*prát* > *vyprat*), a slovesa, jejichž základ je považován za citoslovce (*dupat, šňupat, ...*).

Imperativ je v tomto případě dvojí:

a) U některých sloves dochází k hláskové alteraci Ø/e v prézentním kmeni. K takovému základu přidáváme koncovky -Ø, *-te* (*ber-Ø*).
b) U víceslabičných a některých jednoslabičných (s neslabičným základem) přidáváme koncovky *-i -ěte* (*rv-i, rv-ě-te*).

Víceslabičná slovesa, u kterých nedochází k alternaci Ø/e, kolísají se vzorem *dělá* 5. slovesné třídy, získávají koncovku *-ej*, *-ejte* (*nasyp-Ø* x *nasyp-ej*).

4) **Slovesa typu *maž-e***

Pod typ *maže* řadíme slovesa, která mají v infinitivním i minulém kmeni *-a-* (*maz-a-t*). Jedná se o slovesa původní a jejich předponové deriváty (*do-kázat*).

Imperativ je zde velmi rozkolísaný, má dokonce tři možnosti:

a) Slovesa *psát, ukázat, kázat, dokázat, vázat a tázat* mají imperativní tvary pouze s koncovkami -∅, *-te* (*piš-∅*, *ukaž-∅*).
b) Slovesa, jejichž infinitiv končí na *-slát*, slovesa s neslabičným základem (*lhát*) a jejich předponové deriváty tvoří imperativ pomocí koncovek *-i, -ete* (*pošl-i, pošl-ete*).
c) V případě, že je prézentní kmen ukončen na *ž/š/č* a minulý a infinitivní na *-z/s/k*, lze vedle nulové koncovky v 2. os. sg. připojit koncovku imperativu 5. třídy *-ej* (*maž/mazej,* nespisovný tvar *skač-∅/skák-ej*).

5) **Slovesa typu *tř-e***

Tento typ zahrnuje slovesa s kmenotvornou příponou *-e-* v kmeni minulém a *-í-* v infinitivním (*tř-e-l, tř-í-t*). Ta se připojuje k základu končícímu nejčastěji konsonantem *ř,* dále sem řadíme sloveso *mlít* a jeho předponové deriváty.

Imperativ tvoříme dvěma způsoby:

a) Je-li prézentní kmen ukončen na *ř*, přidáváme koncovky *-i, -ete* (*dř-i*, *dř-ete*).
b) U sloves typu *mlít* dochází ke vkladné hláskové alternaci ∅/e a přidávají se koncovky -∅, *-te* (*mel-∅*).

Druhá třída (*-ne-*)

Druhou slovesnou třídu charakterizuje kmenotvorná přípona *-ne-*.

1) **Slovesa typu *tisk-ne***

Podle tohoto vzoru se časují slovesa se základem ukončeným konsonantem. Nepravidelně jsou zde zahrnuta i slovesa bez *-e-,* a to *ukrást, říct*, *nalézt* a *vynalézt*.

Imperativ má dvojí možnost tvorby:

a) Imperativ je tvořen přidáním kmenotvorné přípony -n- a koncovek *-i*, *-ěte* (*tisk-n-i*, *tisk-n-ěte*; *řek-n-i*, *řek-n-ěte*).
b) Slovesa, která kolísají se vzorem *peče*, např. *obléct/obléknout,* mají i dvě možnosti imperativu, a to *-i*, *-ěte* a -∅, *-te* (*obléknl/obleč*).

2) **Slovesa typu *mi-ne***

Vzor *mine* určuje tvary slovesům, která mají základ ukončený vokálem[9] (často na *i, y*) a obsahují kmenotvornou příponu *-nou-* v infinitivním kmeni a *-nu- v* kmeni minulém.

Imperativ tvoříme pravidelně přidáním koncovek -Ø, *-te*.

3) **Slovesa typu *zač-ne***

Do této skupiny patří slovesa se základem končícím na konsonant a s prézentní příponou *-ne/-me* (*zač-ne, sej-me*). Jedná se o skupinu s velkou mírou nepravidelnosti.

Imperativ se však tvoří pravidelně, a to stejně jako u vzoru *tiskne* připojením koncovek *-i, -ěte* (*zač-n-i, zač-n-ěte*).

Třetí třída (*-je/-uje-*)

Třetí třídu odlišuje od ostatních kmenotvorná přípona *-(u)je-*. Je zastoupen dvěma vzory.

1) **Slovesa typu *kry-je***

Základ sloves typu *kryje* je zpravidla ukončen vokálem. v tvarech s infinitivním a minulým kmenem není kmenotvorná přípona. Patří sem slovesa původní a jejich předponové deriváty (*vy-plout*).

Imperativ je tvořen pravidelně přidáním kmenotvorného konsonantu *-j* (*kry-j*) a koncovkami -Ø, *-te*. U slovesa *hrát* navíc existuje varianta *hrej*.

2) **Slovesa typu *kup-uje***

Jako vzor *kupuje* se časují slovesa s rozšířenou kmenotvornou příponou *-uje-*, v infinitivním a minulém kmeni *-ova-,* rozšířenou *-zova-* u přejatých sloves zakončených na *-izovat* (*dramatizovat*, *klimatizovat* apod.). Jedná se o základová slovesa a jejich předponové deriváty (*na-kupovat*), slovesa odvozená od derivačních základů domácích slov (*známk-ova-t*) a slovesa přejatá a sekundární imperfektiva, tedy nedokonavá slovesa utvořená od slov dokonavých (*rozšířit* > *rozšiřovat*).

Tvoření imperativu je pravidelné pomocí kmenotvorné přípony *-uj-* a koncovkami -Ø, *-te* (*kup-uj-*Ø).

[9] Všimli jsme si, že podle vzoru *minout* tvoří tvary i slovesa, která mají v kořeni sonoru: *trnout, hrnout, lnout, mnout*. Ta tvoří imperativy buď stejně jako ta, která mají základ ukončený vokálem (*trň, hrň*), nebo mají v imperativu koncovky *-i, -ěme, -ěte* (*lni, mni*).

Čtvrtá třída (*-í-*)

Typická je pro čtvrtou slovesnou třídu kmenotvorná přípona *-í-* v prézentním kmeni.

1) **Slovesa typu *pros-í***

Tato slovesa mají v minulém i infinitivním kmeni kmenotvornou příponu *-i-* (*mluvil*, *mluvit*). Řadíme sem slovesa základová i odvozená, a to od substantiv (*hřešit, vyčíslit*) a od adjektiv (*červenit*, *omladit*).

Tvary imperativu získáme dvěma způsoby, přičemž u některých sloves mohou být užity oba:

a) Končí-li tvarotvorný základ jedním konsonantem, imperativ tvoříme pomocí koncovek -∅, *-te* (*pros*-∅).
b) Končí-li tvarotvorný základ dvěma konsonanty, použijeme koncovky *-i*, *-(e/ě)te* (*vysvětl-i, vysvětl-ete; zjist-i, zjist-ěte*).
c) U některých sloves platí, že končí-li jejich základ na *-ď, -ť, -ř, -ž, -žď*, mohou mít imperativ typu a i b (*brzď*-∅/*brzd-i, zamhuř*-∅/*zamhouř-i*).

2) **Slovesa typu *trp-í***

Podle vzoru *trpí* časujeme slovesa s kmenotvornou příponou *-e-/-ě-* v kmeni minulém i infinitivním. Jsou to jak slovesa základová (*hledět*), tak slovesa odvozená, a to zejména od adjektiv (*tmavět*) a od citoslovcí (*bzučet*).

Imperativ dělí tato slovesa na 2 části:

a) Nejčastěji je tvořen koncovkami -∅, *-te* (trp-∅).
b) Ke slovesům s neslabičným základem *spát, lpět, znít* přidáváme koncovky *-i*, *-ěte* (*sp-i, sp-ěte*).

3) **Slovesa typu *sází***

V kmeni minulém a imperativním mají slovesa tohoto typu kmenotvorné přípony *-e-* nebo *-ě-*. Jsou to slovesa základová (*házet*), ale i odvozená od substantiv (*koulet*) a především adjektiv (*šedivět*).

Imperativ tvoříme od prézentního kmene s kmenotvornou příponou *-ej-* a koncovkami -∅, *-te* (*sáz-ej*).

Pátá třída (*-á-*)

Pátá třída je zastoupena jediným vzorem *dělá s* kmenotvornou příponou *-á-* v prézentním kmeni a *-a-* v kmeni minulém i infinitivním (*dělat, dělal*). Časují se podle něj slovesa základová (*volat*) i odvozená, a to buď od adjektiv (*červenat*), nebo citoslovcí (*kňourat*).

Imperativ je tvořen pomocí kmenotvorné přípony *-ej-* a koncovek -∅,*-te* (*loup-ej*).

Slovesa nepravidelná

Vedle tříd, ke kterým přiřazujeme slovesa podle daných kritérií, stojí slovesa nepravidelná s vlastním zvláštním typem flexe. Jsou to slovesa *být, mít, chtít, jít, jíst, bát se, spát, stát, vědět*. Ačkoliv se jedná o slovesa frekventovaná, jsou nezařaditelná v důsledku historického vývoje. Imperativy od těchto devíti sloves mají následující podoby (Štícha, 2013, s. 496–503):

- být > buď, buďme, buďte
- mít > měj, mějme, mějte
- jít > jdi, jděme, jděte
- chtít > chtěj, chtějme, chtějte
- jíst > jez, jezme, jezte
- bát > boj, bojme, bojte
- stát > stůj, stůjme, stůjte
- spát > spi, spěme, spěte
- vědět > věz, vězme, vězte

Systém tvarů podle kmene prézentního se dělí na 5 skupin, celkem 14 vzorů a 9 nepravidelných sloves. v rámci jednotlivých vzorů dochází k dalším rozrůzněním. Ta jsou způsobena dvěma hlavními problémy. Prvním jsou hláskové alternace, druhým je nepravidelnost imperativů uvnitř některých vzorů. Je zřejmé, že imperativ není hlavním kritériem pro rozlišení tříd, ty jsou pro jeho konstrukci pouze více či méně spolehlivou nápovědou.

3.3.3.2 Slovesné tvary podle kmene infinitivního/minulého

Druhým a méně častým přístupem určování slovních tvarů je podle kmene minulého nebo infinitivního. Kmen infinitivní získáme odtržením koncového *-t* od tvaru infinitivu aktivního (*děl-a-t*), kmen minulý odtržením koncového *-l* od příčestí minulého, l-ového (*děl-a-l*). Takto dělíme slovesa do 6 skupin. Naznačíme dělení sloves do tříd zapsané v *Příruční mluvnici češtiny* (Karlík, Nekula a Rusínová, 2012).

První třída (-Ø-)

Kmenotvorná přípona před koncovkou infinitivu je nulová, základ slovesa před ní je určující pro další dělení na podtřídy.

Imperativ se tvoří buď nulovým morfem (*nes*, *vez*) nebo měkčením koncového konsonantu (*veď*, *plet*). Končí-li základ imperativu na dva konsonanty, tvoří se imperativ pomocí koncovky *-i* (*čti*):

1) **Slovesa typu *péct***

Do této podskupiny patří slovesa s dvojím možným ukončením na *-ct/-ci*, jako jsou *vléct/vléci, téct/téci, tlouct/tlouci, moct/moci*.

Archaické tvary imperativu (*pec*, *vlec*) nepodléhají měkčení. v některých případech se tyto tvary zachovaly dodnes (*pomoz*), běžně však mají imperativní tvary v současné češtině variantu s měkkým koncovým konsonantem (*peč, tluč,* nespisovné *pomož*).

2) **Slovesa typu *krýt***

Vzor *krýt* zastupuje slovesa s jednoslabičným infinitivem a jejich předponovým derivátům. Patří sem tato slovesa: *bít*, *blít*, *hnít*, *lít*, *mýt*, *šít*, *vít*, *žít*, *zout*, *obout*, *chvět* se, *pět*, *spět*, *kát se*, *hřát*, *přát*, *sít*, *smát se*, *vlát*, *zrát*, *kout*, *plout* a další.

Imperativ tvoříme pomocí koncovky *-j* (*kryj* / regionálně substandardně *(za)krej*).

3) **Slovesa typu *třít***

Pod vzor *třít* řadíme slovesa, která mají v základu infinitivu ukončení na dva konsonanty.

Imperativ je tvořen přidáním koncovky *-i* (*tři*).

Druhá třída (*-nou-*)

Druhou třídu charakterizuje kmenotvorná přípona *-ou-/mou-*. Dále ji dělíme podle zakončení základu před příponou.

1) **Slovesa typu *tisknout***

Pro tuto skupinu sloves je typické, že se před zakončením *-nout* nachází konsonant.

Imperativ je tvořen pomocí zakončení *-i* (*tiskni, utni*).

2) **Slovesa typu *minout***

Před zakončením *-nout/-mout* je obvykle vokál nebo slabikotvorný konsonant *r*, *l*.

Imperativ se tvoří dvěma způsoby:

a) Nulová koncovka s měkčící alternací *n* > *ň* (*zahyň, hrň, viň*),
b) Zakončení na - (*hni*, *lni*, *promni*).

3) **Slovesa typu *tnout***

Tato skupina zahrnuje vedle *tnout* slovesa na *-jmout,* dále *pnout, vzít, začít* a *žnout*.

Imperativ tvoříme pomocí koncovky *-i* (*tni, pojmi, vezmi, začni, žni*).

Třetí třída (*-ě-/-e-*)

Do třetí třídy řadíme slovesa, jejichž infinitiv má kmenotvornou příponu *-e-/-ě-*.

1) **Slovesa typu *trpět***

Vzor *trpět* zastupuje neodvozená slovesa zakončená na *-et/ět* (*bolet*, *bydlet/bydlit*, *klečet*, *ležet*, *myslet*, *sedět*, *slušet* apod.), jednoslabičné infinitivy zakončené na *-ít* a *-ět/ít* (*bdít*, *čnět/čnít*, *skvět/skvít*, *tlít*, *znít*) a sloveso *spát*.

Imperativ má nulovou koncovku (*trp*), toto pravidlo zde však neplatí ve všech případech, některá slovesa mají v imperativu zakončení na *-i* (*bydli, mysli, spi*).

2) **Slovesa typu *sázet***

Skupina sloves, které jsou řazeny k typu *sázet*, jsou mimo jiné i slovesa odvozená od sloves nebo adjektiv (*bolavět, růžovět*).

Imperativ tvoříme pravidelně pomocí koncovky *-j* (*sázej*).

Čtvrtá třída (*-i-*)

Kmenotvorná přípona, charakteristická pro čtvrtou třídu, je *-i-*.

1) **Slovesa typu *prosit***

Imperativ těchto sloves bývá často pod vlivem příčestí trpného, deverbativní adjektiva a substantiva. Objevují se v něm v rámci mluvených forem také pravidelné alternace (*pros/proš*, *přihlas/přihlaš*).

Pátá třída (*-a-/-á-*)

Pátá třída zahrnuje ta slovesa, jejichž kmenotvornou příponou je *-a-/-á-*.

1) **Slovesa typu *volat***

Imperativ je u těchto sloves tvořen pomocí koncovky *-j* (*volej*). Některá slovesa s imperativem s nulovou koncovkou k tomuto typu přecházejí (*klep* > *klepej, oř* > *orej*).

2) **Slovesa typu *mazat***

Imperativ má nulovou koncovku, může docházet k měkčení koncového konsonantu (*maž*).

3) **Slovesa typu *brát***

Imperativ má nulovou koncovku, dochází zde k alternaci Ø/*e* (*ber*).

a) **Slovesa typu *cpát***

Imperativ tvoříme pomocí koncovky *-i* (*cpi*), ostatní tvary se od sloves typu *brát* neliší.

Šestá třída (*-ova-*)

1) **Slovesa typu *kupovat***

Imperativ se v této třídě tvoří pravidelně pomocí zakončení *-uj* (kupuj).

Izolovaná slovesa

Vedle pravidelných sloves, které třídíme do výše popsaných šesti tříd, zůstávají mimo systém slovesa nepravidelná (Karlík, Nekula a Rusínová, 2012, 334–334):

- být > buď, buďme, buďte
- jíst > jez, jezme, jezte
- vědět > věz, vězme, vězte
- vidět > viz, vizme, vizte
- mít > měj, mějme, mějte
- chtít> chtěj, chtějme, chtějte
- jít > jdi, jděme, jděte

Slovesné dělení podle kmenů infinitivního/minulého sestává z 6 tříd a celkem 14 vzorů, do kterých nelze zařadit 7 izolovaných sloves s vlastním způsobem tvorby tvarů imperativu. Ani tento přístup není pro generování imperativů zcela praktický. S naším systémem má společný základ, tedy infinitiv. Jelikož však bere v potaz celé paradigma, nikoliv pouze a primárně imperativní tvary, je jeho členění pro naše účely stále nedostatečné.

3.4 Infinitiv

Infinitiv představuje absolutně bezpříznakový tvar slovesa, reprezentuje sloveso jako lexém, jeho aktivní forma jej zastupuje ve slovnících. Jedná se o jednoduchý slovesný tvar, tvoří se přidáním koncovky *-t* (*-ti*, *-i*) k infinitivnímu kmeni. Ze slovesných kategorií vyjadřuje pouze slovesný vid a rod. Bývá součástí složených slovesných tvarů, jako je např. opisné futurum sloves nedokonavých (*budu zpívat*).

V této kapitole definujeme formální podobu imperativu, jak jsme ji ověřili na korpusových datech (korpus SYN, verze 3)[10], včetně nespisovných a okrajových jevů.

Při popisu infinitivních tvarů sloves jsme vymezili taková zakončení, která v žádném případě infinitivu nepřísluší, případně pouze jeho nespisovným variantám, a naopak taková zakončení, která jsou infinitivům vlastní:

[10] Nevyužili jsme novější verzi korpusu SYN, protože přebíráme vlastní výzkum ze svých dřívějších prací.

1) vokál + *t(i)*:

Nikdy ukončení nepředchází tyto vokály + *t*:

- *-ot(i)*
- *-ůt(i)*
- *-ét(i)*
- *-yt(i)*
- *-[^o]ut(i)* (tj. ukončení *-ut,* jiné než *-out*)

V nespisovné češtině či nářečích se mohou vyskytovat tato ukončení:

- *-out(i)* > *-ót* (př. *zvednout* > *zvednót*)
- *-out* > *-ut* (př. *vyzout* > *vyzut*)
- *-ýt(i)* > *-ét* (př. *být* > *bét*)
- *-ýt(i)* > *-yt* (př. *být* > *bét*)

Naopak vokály v existujících spisovných infinitivech:

- *-at(i)*: *poslat(i)*
- *-át(i)*: *dát(i)*
- *-et(i)*: *přijet(i)*
- *-ět(i)*: *uvádět(i)*
- *-it(i)*: *prosit(i)*
- *-ít(i)*: *žít(i)*
- *-out(i)*: *vyzout(i)*
- *-ýt(i)*: *nabýt(i)*

2) konsonant + *t(i)*:

Ve spisovné češtině se před koncovkou *-t(i)* vyskytují pouze následující konsonanty a dále jim předcházející hlásky:

- *-st(i)*: *třást*
- *-ást(i)*: *třást*
- *-ést(i)*: *nést* (> *-ýst(i)*: *nýst* – nespisovný tvar)
- *-íst(i)*: *číst*
- *-ůst(i)*: *růst*
- *-ást(i)*: *krást*
- *-bst(i)*: *zábst*
- *-zt(i)*: *vézt*
- *-ézt(i)*: *nalézt* (> *-ízt(i)*: *vlízt* – nespisovný tvar)
- *-ýzt(i)*: *hrýzt*
- *-ct(i)*: *tlouct*

V nespisovné/hovorové češtině navíc můžeme narazit na zakončení:

- *-ýt(i)* > *-ejt*: *krýt* > *krejt*

3) vokál + *ct/ci*:

Ve spisovné češtině mohou mít infinitivy omezené množství hlásek před zakončením *-ct/-ci*:

- *-éc[ti]*: *péct/péci*
- *-íc[ti]*: *říct/říci*
- *-ouc[ti]*: *tlouct/tlouci*
- *-oc[ti]*: *pomoct/pomoci*

Nespisová čeština dovoluje navíc tyto varianty:

- *-ec[ti]* > *-íct*: *péc[ti]* > *píct* (nespisovné/hovorové varianty)

Ačkoli jsou tyto definice plně spolehlivé, je třeba počítat s výrazným přegenerováním. Velmi často jsou tato zakončení vlastní i slovním tvarům dalších slovních druhů.

Na základě dat uvedených v korpusu SYN 2015[11] analyzujeme infinitivní zakončení a uvádíme jejich frekvence v číslech v následující tabulce:

Tabulka 8: Analýza infinitivních tvarů

	počet výskytů	**počet lemmat**	**% (lemmat)**	**rozložení v korpusu**
-[sz]t	344 366	237	1,15 %	1,53 %
-c[ti]	651 529	88	0,43 %	0,65 %
-at	4 479 800	12 190	59,31 %	52,90 %
-ovat	1 593 283	6 250	30,41 %	27,49 %
-át	628 271	142	0,69 %	0,95 %
-et	1 010 100	775	3,77 %	4,23 %
-ět	635 797	665	3,24 %	3,03 %
-it	3 130 979	4 149	20,19 %	23,75 %
-ít	1 834 737	353	1,72 %	2,24 %
-out	908 087	1 880	9,15 %	10,39 %
-ýt	4 079 669	54	0,26 %	0,33 %
celkem	17 768 551	20 553	100 %	

[11] Slovesných tvarů je v korpusech velmi velké množství a prohledávání korpusu SYN v5 je velmi zdlouhavé, technicky ne zcela možné. Proto vybíráme tuto alternativu, korpus, který je vyvážený a pro tak běžný lingvistický objekt, jako jsou slovesa, přinejmenším dostačující.

Poslední sloupec tabulky připomíná rozložení sloves podle infinitivního ukončení v korpusu SYN2015. Ačkoliv jsme vzorek k anotování vybrali náhodně, procentuální zastoupení všech skupin sloves zhruba odpovídá korpusovým datům, předpokládáme proto, že odráží rozložení běžně užívaných sloves.

Tabulka napovídá, že některé typy sloves jsou uzavřené, zastupují malé množství sloves (*-ci*, *-ýt*). Naopak ukazuje, že zakončení *-ovat* je nejproduktivnějším zakončením, velmi početné jsou také skupiny na *-it* a za pozornost stojí také ukončení *-out*. Všechna tato ukončení charakterizují typy sloves, jejichž tvarosloví kopírují přejatá cizí slova (*mail-ovat*, *googl-it*, *kill-nout*).

4 Hláskové alternace

„Morfologické alternace lze definovat jako pravidelné střídání určitých fonémů, popř. jejich skupin“ (Osolsobě, 2011, s. 20). Čechová (2011, s. 187–188) uvádí přehled všech hláskových skupin, kterými se v následujícím popisu inspirujeme. Zohledňujeme pouze ty změny, kterými hlásky prochází při tvorbě imperativu z infinitivu, případně v závorce uvádíme mezikrok, ke kterému dochází v indikativu (př. *á (> o) > ů*). Je třeba zdůraznit, že se zaměřujeme na grafickou podobu jednotlivých hlásek. Ke hlásce *d* ve slově *škodit* tedy přistupujeme jako k neměkčenému *d,* nikoliv *ď*, ačkoliv jej tak v daném případě vyslovujeme. Půjde nám tedy o grafické alternace v českém tvarosloví.

4.1 Samohláskové alternace

Tabulka 9: Alternace vokálu

alternace		**příklady**
kvantita: krácení	á > a	pást > pas, vrátit > vrať
	é > e	nést > nes, lézt > lez
	í > i	lízat > liž
	ý > y	krýt > kryj
kvalita	á > ě	vát > věj
	á > e	třást > třes, zábst > zeb, stlát > stel
	á > o	bát > boj, ustát > ustoj
	á (> o) > ů	stát (> stojí) > stůj
	í > ě	pospíšit > pospěš, chtít > chtěj
	í > e	sít > sej
	ů > o	růst > rosť
	ý > u	být > buď
kvantita a kvalita	ou > u	tlouci > tluč, plout > pluj
alternace s nulou	Ø > e	brát > ber, mlít > mel
	Ø > i	psát > piš

Samohláskové alternace dělíme na kvantitativní (při tvorbě imperativu dochází pouze ke krácení vokálů, nikdy k dloužení), kvalitativní, smíšené a alternace s nulou. Charakterizují další stupeň podvzorů. Často k nim dochází v kombinaci s alternacemi samohláskovými. Vzor s nejvíce objevenými kombinacemi je vzor *prosit*, podle kterého může dojít jak ke krácení kořenového vokálu (je-li dlouhý), tak k případnému měkčení koncového konsonantu *d*, *t*, *n* (*sv**í**tit* > *sv**i**t**'**, *br**á**nit* > *bra**ň***).

Alternaci *ou* > u jsme zařadili na stejnou úroveň jako klasické alternace krácení, neboť k ní dochází ve stejných případech (vzor prosit: ***sou**dit* > ***su**d*).

4.2 Souhláskové alternace

Tabulka 10: Alternace konsonantu/konsonantů

alternace	**příklady**
bs > b	zábst > zeb
c > č	péci > peč, tlouci > tluč
c > z (ž)	pomoci > pomoz (pomož)
c **> k (c)**	říci > řekni (rci)
d > ď	chladit > chlaď, hledět > hleď
n > ň	bránit > braň, hanět > haň
r > ř	orat > oř
s > š	česat > češ
t > ť	mlátit > mlať,
z > ž	ukázat > ukaž
d > z	vidět > viz
h > ž	hnát > žeň, lhát > lži
k > č	plakat > plač
r > ř	orat > orej (oř)
s > š	psát > piš
s > d	ukrást > ukradni
s (> d) > ď	dovést > doveď
s (> t) > ť	mást > mať
sl > šl	poslat > pošli
s > sť	růst > rosť

Zapsali jsme 20 souhláskových alternací, se kterými se u tvorby imperativů setkáváme s určitou pravidelností. Jsou jedním z hlavních důvodů štěpení hlavních slovesných vzorů našeho systému. Ve většině případů dochází k alternaci na konci slovního základu, za nímž následuje nulová koncovka (nebo *-me, -te*).

4.3 Popis pravidel alternací při tvoření tvarů sloves

Klára Osolsobě ve své knize *Morfologie českého slovesa a tvoření deverbativ jako problém strojové analýzy* (2011) popisuje alternace, k nimž dochází při časování slovesa, a dělí je do skupin dle typu

morfu, ve kterém se alternující hlásky nacházejí. Pro naši práci je důležité, že si všímá i grafické realizace hlásek.

a) Kmenotvorná přípona (KP) a samohláska/diftong v základu (kořenový vokál, KV)

Alternace KP tvoří velmi bohatý systém. Střídání má pevná pravidla, na základě kterých rozlišuje mluvnice slovesa na třídy a vzory. I v našem případě jsou tyto alternace stěžejní východisko pro tvorbu vzorů. Tyto alternace podléhají vlivu hláskového okolí.

Alternace KV patří mezi alternace volné, takové, které nejsou ovlivněny okolními hláskami. Jejich přítomnost je vázána na jednotlivá slova. Z toho vyplývá, že jejich hledání a zachycení do vzorů může činit obtíže.

b) Finální souhláska/konsonant kořene (KF)

Dle Osolsobě (2011, s. 22) působí v systému sloves tendence k neměnnosti kmene, vyrovnání některých souhláskových alternací. To způsobuje vznik velkého množství dublet (př. *plakat* má ustálené obě varianty imperativu: s měkkou finálou *plač* i bez alternace *plakej*).

c) Samohlásky v prefixu, na švu prefix/základ

Prefixy se v této práci nezabýváme, protože zkoumáme slovní tvary dle koncového řetězce. Přesto jsou při tvorbě imperativu důležité, v některých případech samohláska v prefixu alternuje (alternace *e*/Ø: vehnat > *vžeň, stít* > *setni*). Klára Osolsobě (2011, s. 38–44) alternacím v prefixu věnuje značnou pozornost, my na ni v tuto chvíli pouze odkazujeme.

Osolsobě (2011, s. 23–44) sepsala pečlivý popis pravidel alternací při tvoření tvarů sloves. Nejprve zpracovala alternace vokalické, následně konsonantické.

4.3.1 Pravidla samohláskových alternací

Co se týče alternací vokalických (samohláskových), je pro naši práci zásadní tvrzení, že:

- V imperativu dochází ke kolísání kvantity/kvality hlásek i střídání *e*/Ø pravidelně.

Pomocí pravidel pak Osolsobě popisuje výskyt jednotlivých hláskových alternací. Opět zmíníme pouze ta pravidla, která se týkají imperativu a jeho tvoření z infinitivu:

a) Dlouhý KV ve všech tvarech se v imperativu krátí.
    - Slovesa 1. třídy vzoru *brát* a *mazat* (*lámat* > *lam, vázat* > *važ*), jsou-li zakončení imperativu -Ø, *-me, -te*.
    - Slovesa 4. třídy vzorů *prosit* a *trpět* (*chválit* > *chval, šílet* > *šil*).

b) Dlouhý KV v infinitivu se ve všech tvarech krátí.
   - Slovesa 1. třídy vzorů *nést, péci* (*nést* > *nes*, *péct* > *peč*) a sloves 3. třídy vzoru *krýt* (*í/ý/ou* > *i/ě/y/u: šít* > *šij*, *krýt* > *kryj*, *dít* > *děj*, *plout* > *pluj*).
   - Neplatí u slovesa *moci* a všech, která zahrnuje pravidlo *c*.

c) Dlouhý KV v infinitivu a l-ovém příčestí se v pasivním příčestí a všech tvarech od kmene přítomného krátí.
   - Sloves *mást, pást, třást* a *zábst* 1. třídy vzoru *nést* (*á* > *a/e*: *pást* > *pas*, *třást* > *třes*).
   - Sloveso *tlouci* od vzoru *péci* (*ou* > *u: tlouci* > *tluč*).

d) Dlouhý KV ve všech tvarech od kmene minulého se ve všech tvarech od kmene přítomného krátí.
   - Slovesa 3. třídy vzoru *krýt* (*á* > *a/e/ě/(oj)*: *sát* > *saj*, *přát* > *přej, smát* > *směj*, *bát* > *boj*) – u sloves s infinitivním KV *á*.
   - Sloveso *hrát* zahrnuje jiné pravidlo, které zohledňuje alternace i v deverbativech. Upustíme-li od nich, podléhá tomuto pravidlu s alternací *á* > *e* (*hrej* i *hraj*).

e) Nulový KV ve všech tvarech od kmene minulého se mění ve všech tvarech kmene přítomného v *e*.
   - Slovesa 1. třídy vzoru *brát* (*hnát, brát, drát, prát, srát, žrát, stlát* > *žeň, ber, der* atd.) a vzoru *umřít* (pouze sloveso *mlít* > *mel*).

f) Nulový KV ve všech tvarech od kmene minulého se v indikativu mění na dlouhý a v imperativu na krátký.
   - Sloveso *psát* 1. třídy vzoru *mazat* (*psát* > *píše* > *piš*).

g) Dlouhý KV ve všech tvarech od kmene minulého se ve tvarech indikativu krátí a v imperativu alternuje dlouhý.
   - Sloveso *stát* (> *stojím* > *stůj*).

h) Krátký KV ve všech tvarech od kmene minulého se v indikativu dlouží, v imperativu zůstává krátký.
   - Sloveso *moci,* které nemá imperativní tvar.
   - Slovesa vzoru *mazat* (*stonat* > *stonej* a *plakat* > *plakej/plač*).

Některé alternace se vyskytují ve třídách uzavřených i otevřených. v otevřených (produktivních) třídách mohou alternace KV souviset s přechodem sloves ze tříd uzavřených.

4.3.2 Pravidla souhláskových alternací

Dále Osolsobě (2011, s. 32–37) zkoumá a popisuje pravidla pro střídání souhlásek na švu kořene a KP a alternace souhlásky, která je součástí KP. Nejprve opět uvádíme tvrzení relevantní naší práci:

- K souhláskovým alternacím nedochází u sloves 3. třídy vzorů *krýt, kupovat* a sloves 5. třídy vzoru *dělat*.

- Do alternačních řad vstupují pouze některé souhlásky/skupiny souhlásek na konci kořene. Výjimečně, je-li kořen reprezentován jedinou hláskou nebo skupinou dvou hlásek, může dojít k alternaci celého kořene. Pravidelně dochází k alternaci před *e/ě*. K alternacím dochází pravidelně před koncovkami tvarů imperativu, je-li KP nulová.
- Pravidelně dochází k alternacím *n* v kmenotvorné příponě u sloves 2. třídy vzorů *tisknout* a *minout*, je-li KP bez vokalického komponentu (*tiskni, miň, začni*).

Pravidla souhláskových alternací (s viditelnou grafickou změnou na pozici alternující hlásky/skupiny hlásek):

a) K alternaci KF dochází před KP nulovou a tvarovými koncovkami imperativu.
   - U sloves 1. třídy vzoru *nést* se obligatorně střídá *d* > *ď*, *t* > *ť* (z původního *s*, př.: *vést* > *vede* > *veď, mést* > *mete* > *meť*).
   - U sloves 1. třídy vzoru *péci* se obligatorně střídá k > *č/c* (pekl > *pečl/péci*), *h* > *z/ž* (*mohl* > *pomoz/pomož*).
   - U sloves 4. třídy vzoru *prosit a trpět* se fakultativně střída *s/š* (*přihlásit* > *přihlas/přihlaš*), *z/ž* (*vrazit* > *vraz/vraž*).
b) K alternacím *n*, které je součástí KP, dochází, je-li KP nulová (tedy bez vokálu), před koncovkami imperativu.
   - U sloves 2. třídy vzorů *tisknout* (> *tiskni*) a *minout* (> *miň*) a některých sloves vzoru *začít* (> *začni*).

4.3.3 Grafické alternace

Grafické alternace jsou pro nás velmi důležité, pokud zkoumáme hláskové alternace v psaném jazyce. „Hovoříme-li o grafických alternacích, máme na mysli korespondence mezi některými morfologickými alternacemi a jejich grafickou realizací, které je třeba mít na zřeteli.“ (Osolsobě, 2011, s. 38)

Grafická realizace se liší v následujících případech:

a) realizace *e* pomocí grafému *ě* po grafémech *d, t, n, b, p, f, v, m,*
b) realizace *bj, pj, mj/mň, vj, (fj)* pomocí *b, p, m, v f* před grafémem *ě,*
c) realizace *ď, ť, ň* jako *d, t, n* před grafémy *ě*, *i*, *í*.

Navíc přidáváme grafické alternace, na které jsme narazili během výzkumu. Jedná se o změny, které mají historický původ a jsou ze synchronního hlediska nejasné. Charakterizujeme je jako střídání znaků s nulou (změna Ø > *ď* ve slovesu *být* > *buď*). Velmi často ke grafickým alternacím dochází u sloves, které mají imperativ na *-ď/ň* nebo *-di/ni/mi*, ale v infinitivu chybí (př. *být* > *buď*, *začít* > *začni*, *vzít* > *vezmi*). Jejich podrobnější popis uvádíme v kapitole, ve které analyzujeme výsledky systému vzhledem k hláskovým alternacím (kap. 8.2).

4.4 Hláskové alternace v mluvnických vzorech slovesných tříd dle prézentního kmene

Komárek (1986, s. 427–494) popisuje třídy a paradigmata jednotlivých vzorů velmi podrobně, včetně jejich vývojových tendencí, častých poznámek k jednotlivým slovesům a právě hláskovým alternacím. v této části práce stručně zmíníme hláskové alternace, ke kterým dochází v základu během tvorby imperativu z infinitivu z pohledu systému slovesných tříd a vzorů.

První třída

1) **Typ A: vzor *nese – nesl, nést*, podvzor *peče – pekl, péct***

Pod vzor *nese* spadají slovesa, která zanechávají základ bez alternace, stejně tak slovesa, která alternují s na *d/t*, v imperativu *ť/ď*. Právě tato alternace zapříčinila přidání slovesa *jít, jet* (*jdi, jeď*) k tomuto vzoru.

U sloves dle podvzoru *peče* se základem na veláru *k/h* v minulém kmeni dále dochází k jejich palatalizaci (*pekl* > *péci* > *peč*).

V tvarotvorném základu jsou nejednotně realizované vokalické alternace dloužení v infinitivu, resp. krácení v imperativu: *á/a* (*klást* > *klaď*), *é/e* (*nést* > *nes*, *péct* > *peč*), í/*e* (*hnís*t > *hnět*), *ý/y* (*hrýzt* > *hryz*), *ů/o* (*růst* > *rost*), *ou/u* (*houst* > *huď*).

U sloves *moci* (*mohl*) a *stříci* (*střehl*) dochází ke střídání základu s velárou *h v* minulém kmeni. Jedná se o alternace *h/ž/z/c:* (*stříci (střehl)* > *střeže* > *střez; pomoci (pomohl)* > *pomohu/pomůžu* > *pomoz*).

2) **Typ B: vzor *bere – bral, brát*, podvzor *maže – mazal, mazat***

Charakteristická pro slovesa dle vzoru *bere* je alternace Ø/e (*b-Ø-rát* > *b-e-r*). U slovesa *hnát* se navíc základové *h* mění v *ž* (*hnát* > *žeň*).

Slovesa dlevzoru *maže* podléhají při tvorbě imperativu častým alternacím. Jedná se o měkčení finály prézentního tvarotvorného základu, a to: *z/ž* (*kázat* > *kaž*), *k/č* (*plakat* > *plač*), *s/š* (*psát* > *piš*), *h/ž* (*lhát* > *lži*), *r/ř* (*orat* > *oř*) a *sl/šl* (*poslat* > *pošli*).

Při tvorbě prézentního kmene z infinitivního dochází ke dloužení kmenového vokálu, v imperativu se však opět krátí (*plakat* > *pláče* > *plač*). U slovesa *psát* dochází k alternaci Ø/*í v* prézentním kmeni, v imperativu nezaniká, pouze se krátí (*p-Ø-sát* > *p-í-še* > *p-i-š*).

3) **Typ C: vzor *tře – třel, třít***

U vzoru *tře* Komárek neuvádí žádné alternace až na sloveso *mlít*, u kterého (na rozdíl od ostatních sloves tohoto vzoru) dochází ke vznikové alternaci Ø/*e* (*mlít* > *mel*).

Druhá třída

1) **Typ A: vzor *tiskne – tiskl, tisknout***

Alternace v tomto vzoru jsou pouze fonetické, ve tvaru *tiskni* zůstává grafická podoba *n* neměkčena kvůli koncovému *-i*. Nepravidelně zde však dochází k tvorbě tvaru *řekni* od infinitivu *říci.*

2) **Typ B: vzor *mine – minul, minout***

Naopak u vzoru *mine* přidáváme v imperativu koncovku nulovou, koncové *n* se tedy i graficky měkčí na *ň* (*miň*). Jedná se o uzavřenou skupinu sloves.

3) **Typ C: vzor *začne – začal, začít***

Stejně jako u typu A k alternacím nedochází, stojí za povšimnutí pouze *-n-*, které v infinitivu chybí.

Třetí třída

1) **Typ A: vzor *kryje – kryl, krýt***

Infinitiv má většinou dlouhý kmenový vokál, u kterého dochází ke krácení: *í/i* (*pít* > *pij*), *ý/y* (*krýt* > *kryj*), *á/a* (*sát* > *saj*), *ou/u* (*plout* > *pluj*), nebo změně kvality *í/ě* (*dít* > *děj*).

2) **Typ B: vzor *kupuje – kupoval, kupovat***

Tato skupina slov je velice rozsáhla a otevřená. Nepodléhá žádným hláskovým alternacím v základu. Mění se pouze kmenotvorná přípona *-ova-* na *-uj*.

Čtvrtá třída

1) **Typ A: vzor *prosí – prosil, prosit***

Vzor *prosí* podléhá při tvorbě imperativu krácení kvantity ve finální slabice, jedná se o alternace: *á/a* (*bránit* > *braň*), *é/e* (*léčit* > *leč*), *í/i* (*mířit* > *miř*), *ý/y* (*zvýšit* > *zvyš*), *ů(ú)/u* (*rozpůlit* > *rozpul*), *ou/u* (*bloudit* > *blud*).

Dále dochází k měkčení primárně hlásek *d, t* a *n*. v úzu proniká také do sloves se základem na *s, z*: *s/š* (*přihlásit* > *přihlaš*) a *z/ž* (*kazit* > *kaž*).

2) **Typ B: *trpí – trpěl, trpět*, podvzor *sází – sázel, sázet***

U sloves typu B nedochází k alternacím často. Výjimečně se vyskytují alternace *d/z* (*vidět* > *viz*) a u kmenotvorné přípony *í/ě* (*lpít* > *lpěj*).

Pátá třída

Vzor *dělá – dělal, dělat*

V páté třídě dochází k pravidelné alternaci pouze u kmenotvorné přípony *a(á)/e* (*dělat* > *dělej*). v základu k alternacím nedochází.

Stejnou tendenci v tvorbě imperativu mají i izolovaná slovesa *dát* (*dej*) a *mít* (*měj*).

4.5 Shrnutí k tématu alternací

Hláskové alternace jsou klíčovým faktorem při třídění (nejen) sloves do vzorů, ať už v mluvnických vzorech či z pohledu našeho, tedy při tvorbě vzorů na míru tvarům imperativu. Jedná se o pravidelné jevy, které jsou pro rodilého mluvčího intuitivní. Problémy nastávají až na rovině kodifikace, což způsobuje problém i při vymezení vzorů, je proto třeba definovat výjimky. Jedná se o slovesa, která se vychylují například kvůli dnes již zastřenému historickému vývoji (typickým příkladem je sloveso *být* > *bud*), a která tvoří okrajové jevy. I přesto vzory pokrývají naprostou většinu sloves, protože je možné díky alternacím definovat různá chování a změny jejich základů a zakončení při tvorbě odvozených tvarů.

U mluvnického systému slovesných vzorů jsme si všimli, že tvorba imperativů není v rámci jednoho vzoru vždy jednoznačná. Je tedy dle našeho názoru přinejmenším pro počítačové zpracování velice výhodné utvořit izolovanou soustavu takových vzorů, které se zaměřují pouze na imperativní tvary a jejich tvorbu tak přehledně systematizují.

5 Homonymie imperativních tvarů

Vladimír Petkevič (2014) se ve své knize *Morfologická homonymie v současné češtině* velice podrobně mapuje problematiku homonymie, která byla do éry počítačové lingvistiky spíše na okraji zájmu lingvistů (Petkevič, 2014, s. 27), předmětem zkoumání byl především její původ. Při počítačovém zpracování jazyka způsobuje homonymie značné komplikace. v tomto případě není podstatné, proč k homonymii došlo, ale kde k ní dochází a jak homonymní tvary automaticky rozpoznávat.

Během projektu se s homonymií potýkáme pouze okrajově. Rozhraní, ve kterém anotujeme slova dle vzorů, nabízí nápovědu, která označuje nabízené tvary, jsou-li obsažena v korpusu. To může být problém, pokud slovo v korpusu doloženo je, jako imperativ ale neexistuje nebo existuje, ale ne od daného základu. Rozlišujeme (dle našich potřeb) homonymii čtyř typů:

1) Shoda správného tvaru imperativu:

Imperativ je k danému infinitivu správně utvořen, homonymie při anotacích nečiní problém. (Je ale všeobecně hlavním problémem při dalších krocích jazykové analýzy.)

a) s imperativem jiného základu

Př.: *koupit* / *kupit* > *kup*
Př.: *přeplácat* / *přeplácet* > *přeplácej*
Př.: *vědět* / *vězet* > *věz*
Př.: *hnát* / *ženit* > *žeň*

b) s jiným slovesným tvarem

Př.: *rozeslat* > *rozešli* (vzor *poslat*) – 3. os. pl. l-ového příčestí slovesa *rozejít*
Př.: *zvát* > *zveme*, *zvete* (vzor *zvát*) – tvary indikativu téhož slovesa

c) s tvarem jiného slovního druhu

Př.: *doprodat* > *doprodej* (vzor *zkoumat*) – substantivum
Př.: *být* > *buď* – příslovce

2) Shoda chybného tvaru imperativu s tvary imperativu jiného základu:

Daný tvar imperativu existuje, ale pro jiný slovní základ, což při anotacích může znamenat komplikace.

Př.: *chcát* > *chci*, *chceme*, *chcete* (vzor *kašlat*) – od slovesa *chtít*
Př.: *zapít* > *zapni* (vzor *začít*) – od slovesa *zapnout*
Př.: zapít > *zapěj* (vzor *chtít*) – od slovesa *zapět*

3) **Shoda neexistujícího tvaru imperativu:**

Tvar je zvýrazněn jako „doložený v korpusu“, ale jako imperativ neexistuje. I na to je potřeba dávat při anotacích pozor.

a) **s existujícím slovesným tvarem:**

Př.: *smažit* > *smaži*, *smažeme*, *smažete* (vzor *určit*) – indikativ od *smazat*
Př.: *sálat* > *sáli* (vzor *kašlat*) – 3. os. pl. l-ového příčestí od *sát*

b) **s tvarem jiného slovního druhu:**

Př.: *hrabat* > *hraběte* (vzor *řvát*) – 2. pád sg. substantiva *hrabě*
Př.: *sochat* > *soch* (vzor *ukázat*) – 2. pád pl. substantiva *socha*
Př.: *přísť* > *přes* (vzor *nést*)

4) **Homonymie způsobená jazykovými chybami:**

Př.: *přivést* > **přiveťte*

Homonymie je pro počítačové zpracování jazyka velmi zajímavá a důležitá oblast. Pro naše účely však výpis výše stačí, dále se jí tedy věnovat nebudeme. Více viz Petkevič (2014).

6 Analýza tvarů imperativu

Imperativní tvary mají v celé české flexi speciální podobu. Upustíme od popisu formální morfologie a pokusíme se analyzovat tvary imperativu strojově jako řetězce znaků. Vycházíme z dokladů v korpusu SYN v5[12], který má velikost přes 4,6 miliard pozic a je automaticky anotován pomocí nejnovějšího (a neúspěšnějšího) českého morfologického analyzátoru. Díky tomu předpokládáme dostatečný vzorek imperativů k tomu, abychom získali objektivní výsledek (i přes uvědomění si jisté chybovosti v desambiguaci imperativních tvarů).

Problém rozložíme na dvě roviny. V první, stěžejní, se zaměříme na základní tvar imperativu ve 2. osobě singuláru. Ve druhé části shrneme tvary plurálové.

6.1 Postup a výsledky prohledávání korpusových dat

Prohledávání korpusových dat je závislé na správném zadání dotazů a případných filtrů. v této podkapitole popíšeme postup, kterým jsme daná data získali, a výsledná fakta, která nám pomohou utřídit imperativy do skupin se stejnou nebo podobnou charakteristikou ukončení.

6.1.1 Tvary imperativu ve 2. osobě singuláru

Zadáme-li v korpusovém manažeru *Kontext* dotaz na všechny imperativy v 2. os. sg., získáme 944 577 výskytů (**[tag="Vi.S...2.....[1-]."]**).

V tabulce 11 charakterizujeme imperativní zakončení, která se objevují s určitou (v korpusových datech viditelnou) pravidelností. Vedle popisu ukončení imperativu v tabulce popisujeme i případné znaky ve tvaru infinitivu. Jejich četnost výskytů v korpusu odráží předposlední sloupec. Nakonec uvádíme, kolik lemmat tvoří imperativ utvořený od popsaného infinitivu daným zakončením. Součty lemmat podskupin se nemusí shodovat s celkovým počtem lemmat dané skupiny (př. sloves na *-j*). Jedno lemma se totiž může objevit vícekrát (*hrát* > *-aj*, *-ej*), případně nejsou vyčteny všechny podskupiny (př. u sloves na *-i*).

[12] Spojení korpusů řady SYN, verze 5 z 24. 4. 2017

Tabulka 11: Analýza imperativních tvarů

zakončení		omezení tvaru infinitivu	počet výskytů	počet lemmat
libovolné		bez omezení	944 577	9 191
-i			138 451	1 735
	-ni		65 946	1 064
		bez *-n-* v základu	17 273	10
	-mi		7 880	27
		bez *-m-* v základu	6 011	2
	-di		22 074	88
		bez *-d-* v základu	19 740	21
-j			323 687	4 850
		bez *-j-* v základu	321 402	4 773
	-ej	bez *-j-* v základu	205 000	2 309
	-uj	bez *-j-* v základu	51 239	2 130
	-ěj	bez *-j-* v základu	25 023	183
	-ůj	bez *-j-* v základu	14 702	4
	-ij	bez *-j-* v základu	8 782	82
	-oj	bez *-j-* v základu	7 898	2
	-aj	bez *-j-* v základu	6 490	35
	-yj	bez *-j-* v základu	1 189	28
-ˇ	**-[ďťňžšč]**		222 913	1 492
		bez -ˇ- v základu ([dtnsz])	112 681	797
	-[ďťň]		137 738	774
		bez -ˇ- v základu ([dtn])	91 361	672
		Alternace *s/d/ď*	3 119	31
		Alternace *s/t/ť*	895	38
	-[šž]		59 667	382
		bez -ˇ- v základu (*[sz]*)	21 281	121
	-č		25 508	337
		bez -ˇ- v základu (*c*)	2 209	49
ostatní	**-z**	*-d-* v základu	4 997	9
		Alternace Ø/*e*	15 201	75

Je třeba počítat také s mírným zkreslením výsledných čísel, jelikož jsme neprocházeli všechny výskyty a nekontrolovali správnost jejich desambiguace.

Celý výzkum je postaven na tvarech, které jsme definovali jako imperativy v 2. os. sg. Na pozici č. 15 *Varianta, stylový příznak apod.*[13] jsme vedle možnosti *nevyplněno* zadali i možnost *varianta, víceméně rovnocenná ("méně častá")*:

[tag="Vi.S...2.....[1-]."][14]

Tím získáváme vzorek jádra jazyka, vyjma nespisovné či jinak příznakové okrajové jevy. Jedná se o základ našeho výzkumu, ze kterého vždy postupně vybíráme slovesa daného ukončení pomocí filtrů. v dílčích prohledáváních korpusových dat je pokaždé znovu základním dotazem.

1) **Tvary na *-i*:**

Pozitivní filtr (word): **.*i**

Pomocí pozitivního filtru vyhledáme tvary na *-i*. Poté dělíme tvary podle ukončení na 2 znaky a dále řešíme, zda je předposlední znak součástí infinitivu nebo nikoliv. Pokud není, předpokládáme specifický jev, který bude třeba popsat složitějším pravidlem. Jedná se o následující tři situace:

a) *-ni*:

Pozitivní filtr (word): **.*ni**

Negativní filtr (lemma): **.*n(([aáeěiíýsz]|ou)t|ci)**

U sloves s tímto imperativním zakončením se většinou jedná o slova s *n v* infinitivním základu (z lingvistického hlediska jako kmenotvorná přípona). Naopak je tomu u sloves . **říci, načít, počít, začít, ukrást, nalézt, sežít(?)*.

b) *-mi*:

Pozitivní filtr (word): **.*mi**

Negativní filtr (lemma): **.*m(([aáeěiíýsz]|ou)t|ci)**

Výjimku mezi imperativy na *-mi* jsou pouze slovesa . **vzít.*

c) *-di*:

Pozitivní filtr (word): **.*di**

Negativní filtr (lemma): **.*d(([aáeěiíýsz]|ou)t|ci)**

Slovesa bez *d v* základu jsou pouze odvozeniny od . **jít*.

---

[13]Pražský tagset

[14]Pozici 15 jsme cíleně omezili, abychom odstranili tvary typu *dejž, staniž*, kterým se v této práci nevěnujeme. Ostatní tvary jsou: *vem, di, poď, pocem, (ne)zab, davaj, (ne)řikej, přiď, (s)mej, přikrej, zakrej, vobleč.*

2) **Tvary na *-j*:**

Pozitivní filtr (word): **.*j**

Takové zakončení je velmi časté, podle předcházející samohlásky jsme je rozdělili do osmi skupin.

Existují však imperativní tvary, jejichž řetězec končí na *-j*, koncovku však mají nulovou. Jedná se o ta slovesa, která mají v infinitivu *-j-* před ukončením (př. *utajit* > *utaj-*∅, *bujet* > *buj-*∅).

Vyhledaná data (pomocí původního dotazu viz výše) jsme omezili filtrem:

Negativní filtr (lemma): **.*j([aáeěiíýsz]|ou)t**[15]

Pomocí filtru jsme odstranili nechtěná slovesa s jiným typem infinitivu. Spolu s nimi jsme však odfiltrovali i slovesa typu *krájet* (*kráj-ej*), jejichž infinitivní základ končí na *-j-*, ale imperativ pomocí *-j* tvořen je. Abychom je mohli do výzkumu zahrnout, vyhledáme si je zvlášť a k ostatním vyhledaným je přičteme ručně[16].

Nové hledání po základním dotazu vyžaduje pozitivní filtr, který vyhledá doplňující slovesa typu *krájet* (*krájej*):

Pozitivní filtr (CQL): **[lemma=".*j([aáeěiíýsz]|ou)t" & word=".*j.j"]**[17]

Jednotlivé skupiny jsme opět hledali pomocí pozitivního filtru (př. **.*ej**).

U skupin *-oj* a *-ůj* sledujeme velmi nízký počet lemmat. Předpokládáme, že se jedná o okrajový jev.

3) **Tvary na *-[ďťňžščř]*:**

Pozitivní filtr (word): **.*[ďťňžščř]**

Tuto skupinu tvoří slovesa, u kterých dochází k alternaci (hláskové či jen grafické) měkčení (př. *hla**d**it* > *hla**ď**, vr**t**ět* > *vr**ť***, mi**n**out > *mi**ň***, *ukázat* > *uka**ž***, *psát* > *pi**š***, *pé**c**t* > *pe**č***, *o**r**at* > *o**ř***[18]). Abychom získali výpis sloves, která mají v imperativu jednu z definovaných měkkých souhlásek způsobenou měkčením, musíme zadat následující filtr, kterým získáme lemmata s neměkčenými hláskami *c*, *d*, *n*, *r*, *s*, *t*, *z* v infinitivu a měkčené *č*, *ď*, *ň*, *ř*, *š*, *ť*, *ž*, na konci imperativního tvaru singuláru (měkčení pomocí ˇ)

Pozitivní filtr (CQL): **[lemma=".*[dtnzscr](([aáeěiíýsz]|ou)t|ci)" & word=".*[ďťňžščř]"]**

Je třeba zmínit, že u souhlásek *ď*, *ť*, *ň* se v některých případech jedná o měkčení pouze grafické (*krá**t**it* > *kra**ť**, se**d**ět* > *se**ď***), měkkost v infinitivu je způsobena následující měkčící samohláskou.

---

[15] Vyhledáno 320 323 výskytů, 4 739 lemmat.
[16] Mohli bychom je vyhledat společně s ostatními pomocí složitého CQL dotazu. Ten je však velmi složitý a vzhledem k velkému množství dat, které zpracovává, je z hlediska využití operační paměti velmi náročný.
[17] Vyhledáno 1 079 výskytů, 37 lemmat.
[18] Tento tvar není uveden v tabulce, doklad v korpusu SYN v5 nebyl nalezen.

Tuto skupinu dále dělíme do několika podskupin, postupně je vyhledáváme pomocí pozitivního filtru. (V tabulce 11 jsou uvedeny jejich výskyty i bez předešlého omezení na ty, u nichž dochází k měkčení.):

a) *-[ďťň]*

Pozitivní filtr (lemma): **.*st**

Všimli jsme si také, že se měkké zakončení v imperativu objevuje i v případě, kdy jsou slovesa v infinitivu na *-st*. U těchto sloves dochází k alternaci *s/d/ď* (*vést* > *ve**de*** > *ve**ď***) a *s/t/ť* (*má**st*** > *ma**te*** > *ma**ť***).

b) *-[šž]*

V imperativem s tímto zakončením často dochází v běžném jazyce k chybám (*pomoci* > **pomož/pomoz*, *přihlásit* > **přihlaš/přihlas*). Je třeba tuto oblast prozkoumat a pečlivěji v pravidlech popsat.

c) *-č*

Alternace *c/č* se objevuje bez výjimky u sloves s infinitivním zakončením *-ci* (*péci* > peč, tlouci > *tluč*). Jedná se o uzavřenou skupinu, pro kterou budeme moci snadno vytvořit vlastní vzor.

d) *-ř*

Připomínáme existenci takového imperativu u sloves *orat* > *oř*, v korpusu však doloženo není.

4) **Ostatní**

Další velkou skupinu zakončení už jsme blíže nespecifikovali. Zaměřili jsme se na zbývající imperativy a běžné hláskové alternace, ke kterým dochází jak na konci, tak uvnitř tvaru, srovnáváme-li jej s tvarem infinitivu.

a) Vkladné *e*, alternace Ø/*e*:

Pozitivní filtr (CQL): **[lemma=".*[^e].([aáeěiíý]|ou)t" & word=".*e."]**

Filtr najde ta slovesa, která v infinitivním kmeni *-e-* nemají, v imperativu se však objevuje před posledním znakem v řetězci (*p-**Ø**-rát* > *p**e**r*). Abychom získali nezkreslená data, musíme si odfiltrovat imperativy na *-ej* (*hrát* > *hrej*).

Z mluvnice víme, že tato alternace je vlastní slovesům od mluvnického vzoru *b-Ø-rát* (> *b-e-r*), korpusová data nás v této skutečnosti pouze utvrzují.

b) Alternace *d/z*

Pozitivní filtr (CQL): **[lemma=".*d([aáeěiíý]|ou)t" & word=".*z"]**

Došli jsme k velmi specifické skupině sloves od základů .**vidět* a .**vědět*, která bude v pravidlech zastoupena vlastním vzorem.

6.1.2 Plurálové tvary imperativu

Tabulka 12: Plurálové tvary imperativu

Zakončení ve 2. os. sg.		**Tvar plurálu**
-i		-[eě][mt]e
	-ni	-ě[mt]e
	-mi	-ě[mt]e
	-di	-ě[mt]e
-j		-j[mt]e
-ˇ		-[mt]e
-Ø		-[mt]e

Plurálové tvary jsou obecně charakterizovány jako ukončení *-me/te*. Podle toho, co jim předchází, se dále dělí na čtyři skupiny.

1) *-ěte*, *-ete*

Tyto plurálové tvary se vyskytují u takových sloves, jejichž imperativ v singuláru končí na *-i*. Mezi těmito imperativy tvoří většinu typ *-ěte*. Viz tab. 13, která znázorňuje zastoupení tvarů a lemmat v korpusu SYN v5.

Tabulka 13: Analýza plurálových tvarů imperativu vedle singuláru na *-i*

ukončení	**počet výskytů**			**počet lemmat**		
	1. osoba pl.	**2. osoba pl.**	**vše**	**1. osoba pl.**	**2. osoba pl.**	**vše**
-ě[mt]e	63 645	443 457	507 102	597	1498	1523
-e[mt]e	5 173	122 531	127 704	134	290	311

V korpusu se hledají velice snadno, a to zadáním př. **".*e[mt]e"** jako pozitivní filtr.

2) *-jte*

U této skupiny je hledání v korpusu složitější. Zadáme filtr **".*j[mt]e",** ovšem musíme vytřídit ty případy, kdy *-j* je znakem základu slovesa a kdy už patří k ukončení *-ej, -aj, -uj* apod.

3) *-te*

Největší problém s vyhledáváním tvarů je u skupiny, která v imperativu plurálu připojuje pouze *-me/te*. Vyhledáme ji tím, že odstraníme ukončení na **".*[eě][mt]e"** a část sloves na **".*j[mt]e".** Nesmíme však

odstranit ta, která mají *-j* už v základu slovesa a není tedy součástí ukončení. Problém vyvstává u těch sloves, která mají obě varianty imperativu.

Samostatné vyhledávání imperativu v plurálu není tak snadné ani datově přínosné. Při výzkumu je získáváme na základě tvarů 2. os. singuláru. Rozložení typů imperativů v plurálu mezi veškeré výskyty předpokládáme velmi podobné, jako je tomu u singuláru.

7 Tvorba vzorů pro imperativy

V první řadě je třeba zdůraznit, co je u definic vzorů naším nejvyšším cílem. Je to čtveřice vlastností: korektnost, exaktnost, přesnost a jednoduchost.

Korektností myslíme generování správně utvořených tvarů, které jsou v souladu se zákonitostmi českého tvarosloví (správná podoba slovesného základu, správně utvořené ukončení, správně zvolené alternace apod.).

Exaktnost znamená co nejpodrobnější popis vzoru tak, aby zahrnoval co nejvíce sloves se stejným principem tvoření imperativu.

Přesnost je velmi důležitá při anotacích slovníkových sloves. Čím méně potenciálních vzorů systém ke slovesu nabídne, tím větší je pravděpodobnost, že bude vybrán ten správný. Ideální situace nastává ve chvíli, kdy je pro jedno sloveso nabídnut právě jeden korektní vzor. (Existuje-li k jednomu základu více variant imperativu, nabídne se ideálně právě tolik vzorů.)

Jednoduchost je cíl, kterého se chceme držet na všech rovinách popisu. Samotný systém vzorů by měl být průhledný, tvořen logicky a pravidelně. Stejně tak se snažíme o jednoduchost, co se formy zápisu týče. XML kód prokládáme komentáři, které pomáhají se v zápisu vzorů vyznat.

Při tvorbě samotných vzorů jsme vycházeli z informací o zakončení imperativů i infinitivů (jak je popsáno výše). Hlavním zdrojem výzkumu se stal korpus SYN, verze 4, výsledky jsme ověřovali ve výkladových slovnících a IJP.

Vzory se formálně liší primárně ukončením v infinitivu, rozhodující je znak před koncovým *-t* a zakončením *-ci*. v některých případech jsme do jednoho vzoru spojili ukončení s rozdílem kvantity (např. *-at* + *-át*), které se chovají stejně a změnu kvantity způsobuje přidání prefixu základnovému slovesu (*brát* > *ber*, *zabrat* > *zaber*). Vzory jsou tedy v dokumentu *.xml*[19] poskládány podle základní dvouznakové charakteristiky infinitivních ukončení (viz příloha).

Některé vzory jsou kromě infinitivní koncovky totožné (např. *pros-it* > *pros, běž-et* > *běž, trp-ět* > *trp*). Ty jsou spojeny do nadskupiny v rámci skryté hierarchie, která je zaznamenána uvnitř popisu vzorů pomocí atributů *category* a *parrent*. Nejvyšší ve skryté hierarchii je tedy členění vzorů podle imperativního ukončení. Každý vzor má v *category* zaznamenán druh imperativního ukončení (*-j, -i,* -Ø), *parrent* případně označuje vždy jen nejbližší nadřazený vzor (nejbližšího „rodiče“) v kategorii. Tento způsob jsme zvolili především proto, aby byl zápis vzorů přehlednější, jelikož infinitivních zakončení je více než imperativních.

[19] viz příloha A

Mezi vzory se vyskytují také vzory s přiřazenou prioritou (př. *priority="100"*). Splňují-li slova definice takového vzoru, je vysoká pravděpodobnost, že tvoří imperativní tvary právě podle příslušného vzoru. Nemusí to však znamenat, že nemůže generovat další varianty imperativu sloves i jiný vzor. Čím větší priorita, tím větší pravděpodobnost přiřazení slovesa ke vzoru. Více o přiřazování priority viz kap. 8.3.

7.1 Postup při popisu vzorů

Výzkum směřuje od infinitivu k imperativu – pracujeme s množinou sloves se stejným infinitivním ukončením a třídíme je do vzorů podle ukončení v imperativu. Zpětnou cestou (imperativ → infinitiv) potom ověřujeme správnost našeho výsledku.

Kupříkladu jsme si všimli, že se imperativní tvary na *-j,* (*-jme, -jte*) často odvozují od infinitivních tvarů s koncovým řetězcem v podobě *vokál* + *konsonant* + *vokál* + *t* (-VKV*t*).

- CQL: **[tag="Vi.*" & word=".*j([mt]e)?"]**[20]

Ověříme si to tím, že vyhledáme všechny imperativy na základě infinitivního ukončení (lemmatu) v následující podobě:

- CQL: **[lemma=".*[aeiouyáéěíóůúý].[aeiouyáéěíóůúý]t"]** → ~2,5 mil. výskytů (100 %)

Hledané imperativy *".*j([mt]e)?"* tvoří asi polovinu všech nalezených výskytů (1,3 mil.). Podíváme-li se na zbývající polovinu, tvoří ji z velké části imperativy s nulovou koncovkou od sloves na *-it* (*nič, porad*). Negativním filtrem těchto sloves dostaneme pouhých cca 230 tisíc výskytů, tj. 9,2 %. Mezi těmito zbývajícími imperativy jsou tvary archaické (typ *dávejmež*) a tvary s nulovou koncovkou, které jsou často ke hledaným imperativům variantou (*plav* i *plavej*).

Vrátíme se zpět ke slovesům na **".*[aeiouyáéěíóůúý].it"** v infinitivu a zjistíme, že téměř polovina z nich je v imperativu taktéž končící na *-j*. Toto pravidlo však není tolik spolehlivé, jako u ukončení s ostatními vokály před koncovým *-t*.

Na základě této krátké sondy můžeme obecně předpokládat, že je u sloves na -VKV*t* vysoká pravděpodobnost, že imperativ z nich tvořený bude zakončen na *-j, -jme, -jte*.

Podobným způsobem postupujeme při tvorbě celého systému vzorů. Nebudeme však uvádět celý postup výzkumu. Popíšeme tvorbu vzorů od infinitivů na *-at*, ostatní vzory jsou tvořeny na stejném principu.

[20]Tentokrát hledáme všechny imperativy bez rozdílu příznaku.

7.1.1 Imperativy od infinitivů na *-at*

V korpusu vyhledáme všechny imperativy s lemmatem na *-at*. Postupně pak prozkoumáme každou skupinu ukončení (pozitivním filtrem). Až bude zpracována, z celého vzorku imperativů ji odfiltrujeme (negativním filtrem) a hledáme skupinu další.

- CQL: **[tag="Vi.*" & lemma=".*at"]** → 1 608 005 výskytů

Imperativ na *-j*

V první řadě se zaměříme na nejvíce zastoupenou skupinu imperativů, a to se zakončením *-j, -jme, -jte*. Získáme ji pomocí pozitivního filtru z předešlého dotazu:

Pozitivní filtr: **[word=".*j([mt]e)?"]** → 1 248 729 výskytů (100 %)

Postupným prohledáváním získáváme následující skupiny:

a) **[word=".*ej([mt]e)?"]** → 893 546 výskytů (71,6 %)
b) **[word=".*uj([mt]e)?"]** → 362 497 výskytů (29,0 %)
c) **[word=".*ij([mt]e)?"]** → (~0 %)
d) **[word=".*aj([mt]e)?"]** → (~0 %)
e) **[word=".*ěj([mt]e)?"]** → (~0 %)

Každou skupinu se pokusíme definovat v co nejužší možné míře tak, aby nedocházelo k nadbytečnému přegenerování, ani nežádoucímu vytěsnění slov, která do této skupiny patří. Toho docílíme tím, že vymezíme konkrétní definici infinitivu. Tu pak zadáme do korpusu a ověřujeme, zda pro odpovídající slovesa existují tvary i s jiným ukončením. Pokud je zkoumané ukončení pro daný typ imperativu ideálně většinové, zavádíme jej do popisu vzoru.

a) Vzor ZKOUMAT: *-ej (-ejme, -ejte)*

S téměř třemi čtvrtinami tvarů ze všech imperativů na *-j* tvoří *-ej* největší zastoupení.

Za předpokladu, že je v jazyce tendence vyhýbat se alternacím, hledáme imperativy od sloves na *-kat, -hat, -chat*. U sloves na *-kat* nacházíme vždy imperativní ukončení na *-ej* (*koukej, smrkej, praskej*). Jediné doklady v korpusu s jiným imperativem jsou slovesa .**plakat*, která mají vedle *plakej* také variantu *plač*. Slovesa na *-hat* krom hledaných imperativů tvoří i imperativy od sloves .**lhat* (*lži*). Slovesa na *-chat* mají navíc variantní tvary s nulovou koncovkou u sloves .**nechat* (*nech*). Můžeme tedy definovat téměř plně spolehlivou definici:

(Def. 1) **.*[kh]at**[21]

[21] Hláska *ch* je graficky chápána jako 2 znaky, proto stačí do definice zapsat *h*.

Stejně tak vyhledáme slovesa, která mají před *-at* souhlásky *d, t* nebo *n.* U sloves na *-dat* jsme nenašli jediný doklad s jiným imperativem než na *-ej.* U sloves na *-tat* nalézáme výjimky u sloves *.*stat* (*přestaň*) a na *-nat* u sloves *.*hnat* (*vyžeň*). Od ostatních lemmat se tvoří vždy pomocí *-ej* (*podej*, *motej*, *vychutnej*).

(Def. 2) **.*[dtn]at**

Zkoumali jsme také zakončení s konsonanty *l* a *r* před *-at*. Zjistili jsme, že vzoru odpovídají slovesa na *-rat* nebo *-lat,* předchází-li tomuto ukončení vokál. (Slovesa s konsonantem před *-lat* mají ukončení na *-i: kašli, pošli*). Navíc je třeba zavést sloveso *kašlat*, které má i variantu s *-ej.* U *-rat* přidáváme ještě předcházející *(c)h* (*čechrat* > *čechrej*) a *t* (*větrat* > *větrej*), které mají vždy imperativ na *-ej*.

(Def. 3) **.*[aeiouyáéěíóůúý][rl]at**

(Def. 4) **.*[ht]rat**

Retnice před *-at* mají téměř u všech sloves také imperativ na *-ej*. Je však třeba omezit ukončení na *-pat.* Má požadovanou podobu imperativu jen v případě, že mu předchází vokál či slabikotvorný konsonant. U sloves na *-vat* je častý imperativ na *-uj* kvůli slovesům typu *kupovat,* neplatí to však vždy (*chovat, kovat*, *klovat*, *plovat* apod.). Přesto zůstává imperativní *-ej* silně zastoupené, takže *-vat* zavádíme do definice stejně jako *-pat*.

(Def. 5) **.*[bmf]at**

(Def. 6) **.*[aeiouyáéěíóůúýrl][vp]at**

Definovali jsme ještě obecné ukončení -K[*rl*]K*at* (*šilhat*, *mlsat, drbat*), které má vždy alespoň variantní imperativ na *-ej* (např. *drb* i *drbej*).

(Def. 7) **.*[^aeiouyáéěíóůúý][rl][^aeiouyáéěíóůúý]at**

Mohli bychom pokračovat dále a projít každou hlásku před *-at*, zda a jak pravidelně tvoří požadované tvary imperativu. Jedná se však o největší imperativní skupinu, proto si můžeme dovolit ukončit výčet definicí obecnou, která nabídne toto ukončení u všech sloves na *-at.* Má-li být přiřazování na základě vzorů automatické, kdy na základě těchto vzorů bude vzory přiřazovat guesser, navrhujeme počkat na jeho výsledky a případné omezení pro infinitivy v tomto vzoru doplnit.

(Def. 8) **.*at**

b) vzor KUPOVAT: *-uj (-ujme, -ujte)*

Jedná se o vzor s jedinou definicí. Imperativy tohoto typu se pravidelně tvoří vždy od sloves na *-ovat.* Až na pár výjimek (*chovat, kovat, plovat, klovat*) generuje imperativy na *-uj.* Mohou však být součástí jiného slovesa spadajícího pod tento vzor (*opa-kovat*, *dete-kovat*, *stra-chovat*), imperativ se pak liší (*strach-uj* x *vychov-ej*).

(Def. 1) **.*ovat**

c) Vzor CHCÁT: *-ij (-ijme, -ijte)*

Pro sloveso *chcát*[22] a jeho prefigované deriváty jsme vytvořili samostatný vzor, v němž tvoříme jako jediné ze všech sloves na *-at* (*-át*) imperativ pomocí *-ij.* Jedná se o vzor s nejvyšší prioritou.

(Def. 1) **.*chc[aá]t**

d) Imperativy na *-aj (-ajme, -ajte)*

Jedná se o nespisovný tvar imperativu od slovesa *dávat* > *davaj.* Tento tvar (alespoň v tuto chvíli) do systému nezavádíme.

e) Vzor HUHŇAT: *-ěj (-ějme, -ějte)*

V korpusu jsme narazili na imperativy s *-ěj*. Jedná se o 4 slovesa, z nichž dvě jsou *huhňat* a *hihňat* a tři jsou příznaková (nářečí) či chybně odvozená (*sháňat*, *vyzváňat* a *naháňat*). Vytváříme pro ně nový vzor *huhňat*, který tento způsob tvoření imperativu pokrývá.

(Def. 1) **.*ňat**

Imperativ na -Ø

Vzory, které tvoří imperativy pomocí nulové koncovky, dělíme podle druhu hláskové alternace nebo změn v koncovém řetězci. Tvoří po slovesech na *-j* druhou největší skupinu imperativů od sloves na *-at* (viz tab. 15, 16 a 17).

V korpusu je vyhledáme tak, že postupně odstraníme všechny následující skupiny. Musíme počítat s tím, že kolem 250 zbývajících výskytů (69 lemmat) jsou archaické tvary typu *dávejmež*.

[22] Sloveso je z lingvistického hlediska velmi zajímavé, protože může mít v češtině největší počet variantních infinitivů, tím pádem i od nich odvozených tvarů (*chčít/ščít/chcát/scát* + varianty se zkráceným vokálem *chčit/ščit/chcat/scat*). V systému (alespoň prozatím) uvádíme pouze základní variantu, jedná se o krajně substandardní jevy.

a) **Vzor BRÁT**

Tuto skupinu jsme v systému vzorů spojili do jednoho vzoru, protože se zde vyskytují infinitivy s krátkým *-at* jako odvozeniny od sloves na *-át*. Slovesa typu *brát*, *-brat* jsou tvořena pravidelně pomocí vkladného *-e-*. Jedná se o uzavřenou skupinu sloves, kterou jsme definovali následovně:

(Def. 1) **.*[bdpsž]r[áa]t**

Se stejnými formálními změnami tvoříme imperativy i od slovesa *stlát* a *-stlat.* Proto je zařadíme též ke vzoru *brát.*

(Def. 2) **.*stl[aá]t**

b) **Vzor HNÁT**

Slovesa *hnát*, *-hnat* se chovají stejně jako slovesa vzoru *brát.* Mimo vkladné *-e-* zde dochází navíc k alternacím *h* > *ž* a *n* > *ň*. Tvoříme pro ně tedy samostatný vzor s nejvyšší prioritou.

(Def. 1) **.*hn[aá]t**

c) **Vzor PSÁT**

Pro slovesa *psát, -psat* definujeme vlastní vzor. Chovají se stejně jako slovesa od vzoru *brát,* rozdíl je v tom, že je charakteristický vkladným *-i-* (*napsat* > *napiš*) namísto *-e-*.

(Def. 1) **.*ps[aá]t**

d) **Vzor UKÁZAT**

Nejpočetnější skupinou sloves tvořících imperativy na -Ø jsou slovesa typu *ukázat.* Počítáme s případným krácením kořenového vokálu a měkčením posledního znaku v imperativu před koncovkami (*-Ø, -me, -te*).

Tento vzor popisujeme velmi konkrétně, podařilo se nám pokrýt všechna slovesa, která podle něj tvoří imperativy s nulovou koncovkou (ale počítáme s přegenerováním). Nevylučují však, že se ze sloves na stejná infinitivní ukončení tvoří i jiné formy imperativu.

Prohledáme případy, kdy je v infinitivní ukončení *-sat* nebo *-zat.* Definujeme ukončení na základě předcházejícího vokálu:

(Def. 1) **.*[aeiouyě]sat**

S tímto zakončením jsme v korpusu i během anotací narazili např. na slovesa *česat* > *češ, tesat* > *teš.*

(Def. 2) **.*[aeiouyě]zat**

Této definici odpovídají slovesa typu *mazat* > *maž, řezat* > *řež.*

Ukončení *-zat* doplníme ještě o předcházející dlouhé vokály, u kterých dochází ke krácení, jako je tomu u sloves *lízat, tázat*:

(Def. 3) **.*[áí]zat**

Po dalším průzkumu definujeme skupinu písmen před *-at*, a to *b, ch, m, p* a *v,* kterým předchází buď libovolný krátký (def. 4), nebo dlouhý vokál *á* (def. 5), např. *hrabat* > *hrab* (i *hrabej*), *nechat* > *nech* (i *nechej*), *klamat* > *klam* (i *klamej*), *šlapat* > *šlap* (i *šlapej*), *plavat* > *plav* (i *plavej*).

(Def. 4) **.*[aeiouyě]([bmpv]|ch)at**
(Def. 5) **.*á([bmpv]|ch)at**

Definicí č. 5 jsme nalezli pouze odpovídající slovesa od .**lámat* (> *lam*). Ponecháváme však všechna ukončení.

Na základě předchozího výzkumu přidáváme ke vzoru definici pro sloveso *plakat.* Dochází zde k alternaci koncového k na *č* (*plač*).

(Def. 6) **.*plakat**

Stejně tak je tomu u slovesa *skákat* (*skač*). Dále přiřadíme k tomuto vzoru sloveso *orat* s alternací *r* > *ř* (*oř*) a *drbat* bez alternací.

(Def. 7) **.*skákat**
(Def. 8) **.*orat**
(Def. 9) **.* drbat**

Imperativ na *-ň*

Tento typ imperativů spadá do skupiny nulové koncovky. Dochází zde však k přidání znaku *ň,* které nemá v infinitivním tvaru s čím alternovat.

a) **Vzor VSTÁT**

Mezi slovesy na *-tat* jsme narazili na slovesa typu *zůstat*, která jsou prefigovaná k základu *stát.* Jako jediné netvoří imperativ pomocí *-j* ale přidáním *-ň* (*zůstaň*). Jedná se o výjimku v celé skupině sloves na *-at,* ke které přidáváme sloveso základové (*stát*) a zakládáme vzor *vstát* s nejvyšší prioritou.

Imperativ na *-i*

Imperativy s *-i v* 2. os. sg. se dále dělí také kvůli plurálovým tvarům, rozlišujeme zakončení *-ete* a *-ěte.* Jedná se o malou skupinu sloves.

Pozitivní filtr: **[word=".*i|[eě][mt]e"]** → 8 146 výskytů

Všeobecně předpokládáme, že na *-i* končí v imperativu taková slovesa, která mají před *-at* dva konsonanty.

a) **Vzor ŘVÁT: *-i (-ěme, -ěte)***

Slovesa, která mají v plurálových tvarech imperativu *-ěte*, musí mít na pozici před tímto ukončením takové hlásky, za kterými může *ě* následovat. Jsou to hlásky *d, t, n, b, m, p* a *v*. Vyhledáme taková slova a zkusíme vypsat takové definice, které by je pokryly. Opět u tohoto vzoru platí, že spojujeme skupiny na *-at* a *-át,* jelikož se často jedná o stejný základ a tvary od nich se tedy tvoří stejně.

(Def. 1) **.*[pv][aá]t**

Definujeme slovesa s retnou souhláskou *p* nebo v před *-at/-át* (*spát* > *spi*, *naštvat* > *naštvi*).

Záplata *zvát_Z*

Dle vzoru *ŘVÁT* tvoříme imperativ i u slovesa *zvát* a odvozenin na .**zvat*, které má ve slovnících (jako jediné v této skupině) dubletní tvary v plurálu na *-eme, -ete*. *zvi, zvěte* i *zvete*)[23]. Právě toto je ukázkový případ pro vytvoření tzv. záplaty.

(Def. 1.) ***.*zv[aá]t***

b) **KAŠLAT: *-i (-eme, -ete)***

Oproti předchozímu typu se slovesa dle *kašlat* liší v plurálu vokálem *-e-* namísto *-ě-* před *-me/te.* Po zkoumání těchto sloves v korpusu jsme zjistili, že se jedná o ukončení, kde se před *-at* nachází dva konsonanty, z nichž druhý je jiný než tvrdý či obojetný.

(Def.: 1) **.*[^aieouyěáéíóůúý][^aieouyěáéíóůúýbfmpvdghknt][áa]t**

Během anotací jsme však nenarazili na jiná slova, než která jsou v imperativních tvarech doložena v korpusu. Nabízí se často u sloves, které příslušné imperativy určitě netvoří (*osedlat, odhodlat, zviklat, plát*), nebo sloves, ze kterých v rámci nespisovného jazyka/nářečí se v úzu objevit případně mohou, doklad jsme však nenašli (*dudlat, brblat*). Navrhujeme tedy omezit tento vzor na 2 definice konkrétních sloves:

(Def.: 2) **.*kašlat**

(Def.: 3) **.*rž[aá]t**

Vzhledem ke slovesu *ržát* a jeho prefigovaným odvozeninám jsme opět tento vzor rozšířili o ukončení *-át*.

[23]Obě varianty uvádí výkladové slovníky SSČ, SSJČ i PSJČ. IJP a Morfologický analyzátor AJKA uvádí pouze variantu *zvěte, zvěme*.

c) **POSLAT: *-i (-eme, -ete)***

Během zkoumání předešlých vzorů jsme narazili na slovesa typu *poslat*, ve kterém při tvorbě imperativu dochází k hláskové změně *sl* > *šl* (*pošli*). Vzniká tak nový vzor s nejvyšší prioritou *poslat.*

(Def. 1) **.*slat**

d) **Vzor LHÁT: *-i (-eme, -ete)***

Při zkoumání sloves na *-hat* u imperativního ukončení *-j* a definování vzoru *kašlat* jsme narazili na slovesa .**lh[áa]t*. Vznikl tak vzor *lhát* s nejvyšší prioritou, který zahrnuje sloveso *lhát* a jeho prefigované podoby na *-lhat* (*zalhat*). Mají imperativ tvořený zakončením *-i, -eme, -ete* a při jeho tvorbě dochází k alternaci *h* > *ž.*

(Def. 1) **.*lh[aá]t**

V této kapitole jsme představili, jakým způsobem jsme tvořili vzory pro slovesa na *-at.* Na stejném principu probíhal výzkum i u ostatních sloves. Některé vzory tak vznikly na základě korpusových dat, jiné jsme doplnili až během anotací, když jsme narazili na sloveso, které systém vzoru zatím nebyl schopen správně zpracovat.

8 Analýza vzorů

Celý systém vzorů pro tvorbu imperativů z infinitivu čítá 68 vzorů. v následující kapitole se budeme věnovat jejich analýze. Zhodnotíme je na základě vztahu imperativních versus infinitivních ukončení, na základě hláskových a grafických alternací, které je charakterizují a kterými se od sebe liší. Budeme se věnovat anotacím, jak jsou během nich vzory nabízeny, kdy jsou anotace jednoznačné a kdy je třeba vybírat z většího množství vzorů – kdy je to problém a kdy nikoliv.

Nejprve uvedeme tabulku (14), která reprezentuje všechny vzory. Představuje sloveso, které je zastupuje, jednoduchou charakteristiku obsahující způsob imperativního ukončení a změn, definice určující vzor a názorné příklady.

Tabulka 14: Přehled vzorů

VZOR	POPIS	DEFINICE	PŘÍKLADY
-[sz]t			
nést	imp.: -Ø (*-te*) é/á/í > e	.*[nřl][ééí][sz]t .*v[éí]zt .*ýst	nést > nes třást > třes lézt/lízt > lez nýst > nes
****pást***	imp.: -Ø (*-te*) á > a	.*pást	pást > pas
vést	imp.: -Ø (*-te*) s > ď é/í > e ou > u	NE .*ukrást .*[bflmpsvzř][éí]st .*[rl]ást .*oust	 dovést > doveď příst > předʼ krást > kraď klást > klaď houst > huď
mást	imp.: -Ø (*-te*) s > ť s > sť é > e á > a í > ě ů > o	.*ést .*íst .*ást .*ůst	zamést > zameť hníst > hněť zmást > zmať růst > rosť
****číst***	imp.: -i (*-ěte*) s > t í > Ø	.*číst	číst > čti
****nalézt***	imp: -ni (-něte) Ø > n é/í > e	.*[éí]zt	nalézt > nalezni nalízt > nalezni
****ukrást***	imp.: -ni (-něte) Ø > n s > d á > a	ukrást	ukrást > ukradni

VZOR	POPIS	DEFINICE	PŘÍKLADY
***hrýzt**	imp.: -Ø (-*te*) ý > y	.*hrýzt	hrýzt > hryz
***jíst**	imp.: -Ø (-*te*) s > z í > e/ě	.*jíst sníst	dojíst > dojez sníst > sněz
***zábst**	imp.: -Ø (-*te*) sb > b á > e	.*zábst	zábst > zeb
-ct/ci			
péci	imp.: -Ø (-*te*) c > č é/í/ý > e ou > u	NE .*říc[ti] .*[éíý]c[ti] .*ouc[ti]	 obléci > obleč vyslíct > vysleč týct > teč tlouci > tluč
***říci**	imp.: -ni (-něte) Ø > n c > k í > e	.*říci	říci > řekni
***pomoci**	imp: -Ø (-*te*) c > z (ž)	.*.moc[ti]	pomoci > pomoz/pomož
-at/át			
zkoumat	imp.: -ej (-ejte)	.*[kh]at .*[dtn]at .*[aeiouyáéěíóůúý][rl]at .*[ht]rat .*[bmf]at .*[aeiouyáéěíóůúýrl][vp]at .*[^aeiouyáéěíóůúý][rl][^aeiouyáéěíóůúý]at .*at	cákat > cákej máchat > máchej[1] pomáhat > pomáhej bádat > bádej létat > létej hárat > hárej cválat > cválej čechrat > čechrej skuhrat > skuhrej chátrat > chátrej zanedbat > zanedbej lámat > lámej doufat > doufej obávat > obávej čerpat > čerpej trsat > trsej hltat > hltej jásat > jásej
***huhňat**	imp.: -ěj (-ějte)	.*ňat	ňuhňat > ňuhněj
kupovat	imp.: -uj (-ujte)	.*ovat	tetovat > tetuj

VZOR	POPIS	DEFINICE	PŘÍKLADY
ukázat	imp.: -Ø (-te) k > č s > š r > ř z > ž á > a í > i	.*[aeiouyě]([bmpv]|ch)at .*á([bmpv]|ch)at .*[aeiouyě]zat .*[aeiouyě]sat .*[áí]zat .*plakat .*skákat .*orat .*drbat	nechat > nech hrabat > hrab klamat > klam klapat > klap plavat > plav škrábat > škrab lámat > lam mazat > maž česat > češ kázat > kaž lízat > liž plakat > plač skákat > skač orat > oř drbat > drb
****poslat***	imp.: -i (-ete) sl > šl	.*slat	zaslat > zašli
sát	imp.: -aj (-ajte) á > a	.*[hz]rát .*[klst]át	hrát > hraj zrát > zraj kát > jak lát > laj nasát > nasaj tát > taj
přát	imp.: -ej (-ejte) á > e	.*hrát NE .*zrát NE .*[mv]át NE .*[bdpsž]át .*át	zahrát > zahrej znát > znej
smát	imp.: -ěj (-ějte) á > ě	.*[mv]át .*nadát	zasmát > zasměj vát > věj nadát > naděj
bát	imp.: -oj (-ojte) á > o	.*bát	bát > boj
stát	imp.: -ůj (-ůjte) á > ů	.*stát	stát > stůj
brát	imp.: -Ø (-te) Ø > e	.*[bdpsž]r[aá]t .*stl[aá]t	brát > ber vyprat > vyper stlát > stel
****psát***	imp.: -Ø (-te) Ø > i s > š	.*ps[aá]t	psát > piš napsat > napiš
vstát	imp.: -ň (-ňte) Ø > ň á > a	.*st[aá]t .*plát	vstát > vstaň zůstat > zůstaň vzplát > vzplaň

VZOR	POPIS	DEFINICE	PŘÍKLADY
***hnát**	imp.: -ň (-ňte) ∅ > ň h > ž ∅ > e	.*hn[aá]t	hnát > žeň dohnat > dožeň
***řvát**	imp.: -i (-ěte)	.*[pv][aá]t	rvát > rvi nacpat > nacpi
***zvát_Z**	imp.: -i (-ete)	.*zv[aá]t	zvát > zvete nazvat > nazvete
***kašlat**	imp.: -i (-ete)	.*[^aeiouyáéěíóůúý][^aeiouyáéěíóůúý]at	kašlat > kašli
***lhát**	imp.: -i (-ete) h > ž	.*lh[aá]t	zalhat >zalži
***ustát**	imp.: -oj (-ojte) á > o	ustát	ustát > ustoj
***chcát**	imp.: -ij (-ijte) c > č á > i	.*chc[aá]t	chcát > chčij vychcat > vychčij
-et			
zkoušet	imp.: -ej (-ejte)	.*[aeěiouyáéíóúůý]jet .*[aeěiouyáéíóúůý][^aeěiouyáéíóúůý]et .*mýšlet .*mejšlet .*et	krájet > krájej kráčet > kráčej vymýšlet > vymýšlej vymýšlet > vymejšlej zet > zej
jet	imp.: -∅ (-te) ∅ > ď	.*jet	zajet > zajeď
běžet	imp.: -∅ (-te) í > i	NE .*ou[aeěiouyáéíóůý]et NE .*mýšlet .*ílet .*rl[^aeěiouyáéíóúůý]et .*[aeěiouy][^aeěiouyáéíóúůý]et	 šílet > šil mlčet > mlč křičet > křič
řinčet	imp.: -i (-ete)	NE .*jet NE .*mýšlet NE .*mejšlet .*[^aeiouyáéěíóůúý][^aeiouyáéěíóůúýrl][^aeiouyáéěíóůúý]et .*[aeiouyáéěíóůúý][^aeiouyáéěíóůúý][^aeiouyáéěíóůúý]et	 drnčet > drnči myslet > mysli
-ět			
spět	imp.: -ěj (-ějte)	.*ět	spět > spěj
trpět	imp.: -∅ (-te) d > ď n > ň t > ť	NE .*[áéíóůúý][^aieouyěáéíóůúý]ět .*[aieouyěrl][^aieouyěáéíóůúý]ět .*ět	 civět > civ hledět > hleď klanět > klaň letět > leť
pištět	imp.: -i (-ěte)	.*[^aieouyěáéíóůúý][^aieouyěáéíóůúý]ět	pištět > pišti

VZOR	POPIS	DEFINICE	PŘÍKLADY
vědět	imp.: -Ø (-te) d > z	.*v[iě]dět	vědět > věz vidět > viz
****opomnět***	imp.: -Ø (-te) n > ň Ø > e	.*mnět	opomnět > opomeň
-it			
prosit	imp.: -Ø (-te) d > ď n > ň t > ť á > a é > e í > i ý > y ou > u	.*ou.it .*[(aeěiouyáéíóůúýrl][^aeiouyáéěíóůúý]it .*[sš]tit křivdit .*chtit .*jit .*it	rdousit > rdus bránit > braň praštit > prašť křivdit > křivď dychtit > dychť tajit > taj cítit > cit' toužit > tuž
kazit_Z	imp.: -Ø (-te) s > š z > ž á > a í > i ou > u	.*sit .*zit	hlásit > hlaš kazit > kaž
zanítit_Z	imp.: -Ø (-te) t > ť í > ě	.*nítit .*štípit	zanítit > zaněť vštípit > vštěp
učit_Z	imp.: -Ø (-te)	[dp]oučit	doučit > douč poučit > pouč
zjistit	imp.: -i (-ete)	.*[dntbmpv]it	
určit	imp.: -i (-ěte)	NE .*[dntbmpv]it .*[^aieouyěáéíóůúý][^aieouyěáéíóůúý]it .*[aéíóůý][^aieouyěáéíóůúý]it .*ou[^aieouyěáéíóůúý]it	
****pospíšit***	imp.: -Ø (-te) í > ě	.*spíšit	pospíšit > pospěš
-ít			
chtít	imp.: -ěj (-ějte) í > ě	.*[bdmnptv]ít	odít > oděj mít > měj chtít > chtěj skvít > skvěj
zabít	imp.: -Ø (-te)	zabít	zabít > zab
sít	imp.: -ej (-ejte) í > e	NE .*[bdmnptv]ít .*ít	lít > lej
žít	imp.: -ij (-ijte) í > i	.*ít	žít > žij

VZOR	POPIS	DEFINICE	PŘÍKLADY
mlít	imp.: -Ø (-te) Ø > e	ml[eí]t	mlít > mel mlet > mel
jít	imp.: po-, -Ø (-te) Ø > ď	NE .+jít jít	jít > pojď
****přijít***	imp.: -Ø (-te) Ø > ď	přijít	přijít > přijď
sejít	imp.: -i (-ěte) Ø > d	.*jít	sejít > sejdi
třít	imp.: -i -(ete)	.*ít	třít > tři
ctít	imp.: -i (-ěte)	.*[dntbmpv]it	ctít > cti
začít	imp.: -ni (-něte) Ø > n	.*ít	začít > začni
převzít	imp: -mi (-měte) Ø > m Ø > e	.*vzít	převzít > převezmi
****vzít***	imp.: -Ø (-te) Ø > m Ø > e	vzít	vzít > vem
-out			
tisknout	imp.: -i (-ěte)	.*[aeěiouyáéíóůúý][nm]out	tisknout > tiskni najmout > najmi
zapomenout	imp.: -Ø (-te) n > ň	.*[aeěiouyáéíóůúýrl]nout	zapomenout > zapomeň
pomátnout	imp.: -Ø (-te) t > ť á > a	.*mátnout	pomátnout > pomať
plout	imp.: -uj (-ujte) ou > u	NE .*nout .*out	zout > zuj
-ýt			
krýt	imp.: -yj (-yjte) ý > y ej > y	.*ýt .*ejt	krýt > kryj krejt > kryj
****být***	imp.: -Ø (-te) ý > u ej > u Ø > ď	.*být .*bejt	být > buď pobejt > pobuď

8.1 Imperativ ve vztahu k infinitivu

V tabulce 15 vzory dělíme na 13 skupin dle imperativního ukončení roztříděných na základě typu ukončení infinitivu.

Sledujeme, že zdaleka největší počet vzorů (28) vzhledem k infinitivnímu zakončení mají imperativy na -Ø (*-me, -te*) a *-ň/-ď* (*-me, -te*). Tento fakt přisuzujeme častým a různým hláskovým/grafickým změnám (viz tabulka 16), které způsobuje nulová koncovka (a tedy obnažení základu slovesa). Tento typ ukončení se vyskytuje u všech naznačených druhů infinitivu.

Ukončení na *-j* (*-jme, -jte*) je v tabulce rozděleno podle předcházejícího vokálu. Celkově je zastoupeno 17 vzory, nejčastěji se objevují u ukončení *-ej*. Nikdy takto netvoříme imperativ u infinitivů na -K*t* a *-it*. Nejvíce různých imperativů na *-j* tvoříme od sloves na *-át* (7 vzorů).

Třetí velkou skupinou imperativů je zakončení na *-i*. To dále dělíme dle tvarů plurálu (*-ete/-ěte*) a vyčleňujeme skupinu, která je společně charakteristická hláskami *d, n, m* před ukončením. Vzorů této skupiny je v systému 19. Více jich je u plurálového *-ěte* (12) než *-ete* (7). Všímáme si, že není tento imperativ nikdy odvozován od sloves na *-ýt*. Imperativy na *-[nmd]i* jsou tvořeny pouze od sloves na -K*t* a *-ít*.

Z pohledu infinitivního ukončení můžeme na základě informací v tabulce 15 tvrdit, že nejrozmanitější tvary imperativů pozorujeme u sloves na *-át* a *-at* (po 14 vzorech). Slovesa na -K*t* tvoří uzavřenou skupinu sloves, imperativy tvoří pouze pomocí nulové koncovky nebo *-i* (*-ěte*). Nejméně vzorů nalézáme u sloves na *-ýt*. Imperativy tvoří pravidelně pomocí *-yj* vyjma sloves od *být*.

Některé vzory jsou v tabulce zavedeny dvakrát. To proto, že zastupují vedle sloves na *-át* také slovesa na *-at* (zde jsou označeny znakem +).

Tabulka 15: Imperativní zakončení ve vztahu k zakončení infinitivu

	-ýt	**-out**	**-ít**	**-it**	**-ět**	**-et**	**-át**	**-at**	**-c[ti]**	**-[sz]t**	
-Ø	být	zapomenout pomátnout	zabít mlít vzít_Z	prosit kazit_Z zanítit_Z učit_Z pospíšit	trpět vědět opomnět	běžet	brát psát hnát	ukázat +brát +psát +hnát	péci pomoci	nést pást vést mást jíst zábst hrýzt	**28**
-ň/ď			jít přijít			jet	vstát	+vstát			**4**
-i, -ete			třít	určit		řinčet	zvát_Z lhát	poslat kašlat +zvát_Z +lhát			**7**
-i, -ěte		tisknout	ctít	zjistit	pištět		řvát	+řvát		číst	**6**
-[nmd]i, -[nmd]ěte			začít převzít sejít						říci	ukrást nalézt	**6**
-ej			sít			zkoušet	přát	zkoumat			**4**
-ěj			chtít		spět		smát	huhňat			**4**
-aj							sát				**1**
-ij			žít				chcát	+chcát			**2**
-oj							bát ustát				**2**
-uj		plout						kupovat			**2**
-ůj							stát				**1**
-yj	krýt										**1**
celkem	**2**	**4**	**13**	**7**	**5**	**4**	**14**	**14**	**3**	**10**	**68**

8.2 Imperativ vzhledem k hláskovým změnám

Většina vzorů je charakteristická tím, že ve tvarech k nim řazených dochází k různým alternacím. Tabulka 16 ukazuje i alternace, které nejsou z formy vzorového slovesa zřejmé. Jeden vzor však nereprezentuje jednu konkrétní alternaci, ale celý typ. Např. vzor *nést* zastupuje slovesa na *-st/-zt* a je charakteristický změnou kvantity vokálů *é* > *e* (*n**é**st* > *n**e**s*) a kvality vokálů *á* > *e* (*tř**á**st* > *tř**e**s*), *í* > *e* (*l**í**zt* > *l**e**z*).

V následujících dvou tabulkách představíme dělení vzorů dle alternací vzhledem k imperativnímu/infinitivnímu ukončení.

Bez alternací (hláskových i grafických) tvoří imperativy následujících 18 vzorů: *zkoumat, huhňat, kupovat, řvát, zvát_Z, kašlat, zkoušet, řinčet, spět, pištět, zjistit, učit_Z, určit, zabít, třít, ctít, tisknout, pomátnout.*

Tabulka 16: Tabulka alternací u vzorů vzhledem k infinitivnímu ukončení

	krácení V	**kvalita V**	**alternace Ø/V**	**kvalita (skup.) K**	**(de)palatalizace K**	**grafické změny**
-Ø	NÉST: é > e PÁST: á > a VÉST: é > e, ou > u MÁST: é > e, á > a UKRÁST: á > a HRÝZT: ý > y PÉCI: é > e, ou > u UKÁZAT: á > a, í > i BĚŽET: í > i PROSIT (všechny V) KAZIT: á > a, í > i, ou > u POMÁTNOUT: á > a	NÉST: á/í > e VÉST: í > e MÁST: í > ě, ů > o JÍST: í > e/ě ZÁBST: á > e PÉCI: í/ý > e ZANÍTIT: í > ě POSPÍŠIT: í > ě BÝT: ý/ej > u	BRÁT: Ø > e HNÁT: Ø > e PSÁT: Ø (> í) > i OPOMNĚT: Ø > e MLÍT: Ø > e VZÍT: Ø > e	VÉST: s > ď UKRÁST: s > d JÍST: s (> d) > z MÁST: s > ť, s > sť ZÁBST: bs > b POMOCI: c > z (ž) HNÁT: h > ž	PÉCI: c > č UKÁZAT: k > č, s > š, r > ř, z > ž PSÁT: s > š TRPĚT: d > ď, n > ň, t > ť VĚDĚT: d > z OPOMNĚT: n > ň PROSIT: d > d, n > ň, t > ť KAZIT: s > š, z > ž ZANÍTIT: t > ť ZAPOMENOUT: n > ň POMÁTNOUT: t > ť	HNÁT: Ø > ň VZÍT: Ø > m
-ň/ď	VSTÁT: á > a		JÍT: í > Ø PŘIJÍT: í > Ø			JET: Ø > ď JÍT: Ø > ď PŘIJÍT: Ø > ď BÝT: Ø > ď VSTÁT: Ø > ň
-i (-ete)				POSLAT: sl > šl		
-i (-ěte)			ČÍST: í > Ø	ČÍST: s > t	LHÁT: h > ž	
-[nmd]i, (-ěte)	NALÉZT: é > e	NALÉZT: í > e ŘÍCI: í > e	PŘEVZÍT: Ø > e SEJÍT: í > Ø		ŘÍCI: c > k	SEJÍT: Ø > d ZAČÍT: Ø > n UKRÁST: Ø > n ŘÍCI: Ø > n NALÉZT: Ø > n PŘEVZÍT: Ø > m VZÍT: Ø > m
-ej		PŘÁT: á > e SÍT: í > e				
-ěj		SMÁT: á > ě CHTÍT: í > ě				
-aj	SÁT: á > a					
-ij	ŽÍT: í > i	CHCÁT: á > i			CHCÁT: c > č	
-oj		BÁT: á > o USTÁT: á > o				
-uj	PLOUT: ou > u					
-ůj		STÁT: á > ů				
-yj	KRÝT: ý > y	KRÝT: ej > y				

Vedle hláskových alternací dochází k dalším změnám, které jsou ze synchronního hlediska nejasné. Těmto změnám říkáme alternace grafické, které mají v tabulkách 15 a 16 vlastní kolonku. K těmto změnám přistupujeme jako ke vznikové alternaci znaku na jedné pozici v řetězci.

Jak již bylo několikrát zmíněno, vzory se od sebe (mimo jiné) liší především hláskovými/grafickými změnami. Všímáme si, že před ukončením *-j* může dojít ke změně v předcházejícím vokálu, základ sloves, který mu předchází, se však nemění. Opačně je tomu u sloves s imperativem na -Ø *(-ň/ď)*, která podléhají změnám nejčastěji. Nejvíce se napříč tabulkou objevuje alternace krácení vokálu, můžeme tedy tvrdit, že se jedná o nejběžnější alternaci při tvorbě imperativů.

Dále můžeme vypozorovat, že alternace *-e-* s nulou je typická pro slovesa, která tvoří imperativy pomocí nulové koncovky (a u slovesa *vzít* > *vezmi/vem*). Grafické alternace s *n*, *ň*, *m* a *ď* se vyskytují u sloves s imperativem na -Ø*, -ň/ď, -[nmd]i.*

Následuje tabulka 17 znázorňující alternace, ke kterým dochází ve vzorech vzhledem k jejich infinitivnímu ukončení. Nejméně dochází k alternacím u sloves na *-et*, jedinou alternací, ke které zde může dojít je krácení vokálu (*šílet* > *šil*), vzor *jít* je specifický pro malou skupinu od něj odvozených sloves. Stejně tak je tomu u sloves na *-ýt,* kde dochází k jediné alteraci *ý* > *y* a nespisovné *ej* > *ý,* vzor *být* je zastoupen týmž slovesem a jeho odvozeninami.

Tabulka 17: Tabulka alternací u vzorů vzhledem k imperativnímu ukončení

	krácení V	**kvalita V**	**alternace Ø/V**	**kvalita (skup.) K**	**(de)palatalizace K**	**grafické změny**
-[sz]t	NÉST: é > e PÁST: á > a VÉST: é > e, ou > u MÁST: é > e, á > a NALÉZT: é > e UKRÁST: á > a HRÝZT: ý > y	NÉST: á/í > e VÉST: í > e MÁST: í > ě, ů > o NALÉZT: í > e JÍST: í > e/ě ZÁBST: á > e	ČÍST: í > Ø	VÉST: s > ď MÁST: s > ť, s > sť ČÍST: s > t UKRÁST: s > d JÍST: s (> d) > z ZÁBST: bs > b		UKRÁST: Ø > n NALÉZT: Ø > n
-c[ti]	PÉCI: é > e, ou > u	PÉCI: í/ý > e ŘÍCI: í > e		POMOCI: c > z (ž)	PÉCI: c > č ŘÍCI: c > k	ŘÍCI: Ø > n
-at	UKÁZAT: á > a, í > i			POSLAT: sl > šl	UKÁZAT: k > č, s > š, r > ř, z > ž	
-át	SÁT: á > a VSTÁT: á > a	PŘÁT: á > e SMÁT: á > ě CHCÁT: á > i BÁT: á > o STÁT: á > ů USTÁT: á > o	BRÁT: Ø > e HNÁT: Ø > e PSÁT: Ø (> í) > i		PSÁT: s > š CHCÁT: c > č HNÁT: h > ž LHÁT: h > ž	VSTÁT: Ø > ň HNÁT: Ø > ň
-et	BĚŽET: í > i					*JET: Ø > ď
-ět			OPOMNĚT: Ø > e		TRPĚT: d > ď, n > ň, t > ť OPOMNĚT: n > ň VĚDĚT: d > z	
-it	PROSIT (všechny V) KAZIT: á > a, í > i, ou > u	ZANÍTIT: í > ě POSPÍŠIT: í > ě CHTÍT: í > ě SÍT: í > e			PROSIT: d > d, n > ň, t > ť KAZIT: s > š, z > ž ZANÍTIT: t > ť	
-ít	ŽÍT: í > i		SE JÍT: í > Ø JÍT: í > Ø PŘIJÍT: í > Ø MLÍT: Ø > e PŘEVZÍT: Ø > e VZÍT: Ø > e			*ZAČÍT: Ø > n *SEJÍT: Ø > d JÍT: Ø > ď PŘIJÍT: Ø > ď PŘEVZÍT: Ø > m VZÍT: Ø > m
-out	PLOUT: ou > u POMÁTNOUT: á > a				ZAPOMENOUT: n > ň POMÁTNOUT: t > ť	
-ýt	KRÝT: ý > y	KRÝT: ej > y BÝT: ý/ej > u				*BÝT: Ø > ď

8.3 Přiřazení priority

Pro strojové učení, které bude z našeho výzkumu vycházet, bude důležitá informace o přiřazení priority. Vypíšeme tedy všechny vzory, které mají určitou prioritu, a vysvětlíme, co která priorita znamená.

a) priority="100"

Jedná se o slovesa s nejvyšší prioritou. Znamená to, že pokud sloveso odpovídá charakteristice takovému vzoru, má mu být přiřazen právě a pouze tento vzor. Většinou jsou to okrajové jazykové jevy způsobené vývojovými tendencemi, které jsou ze synchronního hlediska nejasné (*být*). Dále jsou jimi jednotlivá slovesa a jejich prefigované odvozeniny (*lhát*) nebo příznakové varianty imperativu (*vzít* > *vem*).

Nejvyšší prioritu má těchto 22 vzorů: *pást*, *číst*, *nalézt*, *ukrást*, *hrýzt*, *jíst*, *zábst*, *říci*, *pomoci*, *huhňat, poslat*, *psát*, *hnát*, *zvát_Z*, *lhát*, *ustát*, *chcát*, *opomnět*, *pospíšit*, *přijít*, *vzít*, *být*.

b) priority="90"

Priorita 90 se liší od nejvyšší pouze tím, že může být slovesu krom daného vzoru přiřazen ještě tvar dubletní. Opět se jedná o okrajové jevy (*převzít* > *převezmi*) a o jednotlivá slovesa a jejich prefigované odvozeniny (*sejít*).

Prioritu "90" přisuzujeme 7 vzorům: *vstát*, *kašlat*, *zabít*, *jít*, *sejít*, *převzít*, *pomátnout*.

c) priority="75"

Poslední skupinou, které přiřazujeme prioritu, jsou takové vzory, které jsou přiřazeny všem odpovídajícím slovesům až na jednotlivé výjimky. Nemusí se jednat o okrajové jevy (u vzoru *kupovat*: *strachovat* > *strachuj* x *chovat* > *chovej*).

Tuto skupinu tvoří 4 vzory: *kupovat*, *vědět*, *zapomenout*, *krýt*.

8.4 Anotace v číslech

V této podkapitole se budeme zabývat anotacemi sloves, ke kterým jsme přiřazovali imperativní vzory. Jednak budeme zkoumat, v jaké míře a u kterých typů sloves se vzory nabízí, jednak shrneme výsledky anotací, tedy slovesa již správně přiřazená ke vzorům.

8.4.1 Vzory nabízené během anotací

Zaměříme se na data, se kterými pracujeme při anotacích jednotlivých sloves. Ideální stav je takový, že má-li sloveso pouze jeden možný způsob tvoření imperativu, k anotacím se nabízí právě tento jeden vzor. Má-li dvě imperativní varianty, pak se nabízí právě dva příslušné vzory. Čím lépe je systém vzorů popsaný, tím blíže k tomuto stavu jsme. Na jedné straně se snažíme přiblížit k tomuto stavu co nejvíce (nejblíže je tomu slovník), na druhé straně chceme automatickou tvorbu (statistika aj.). S novým

systémem vzorů se pohybujeme někde mezi těmito póly. Snažíme se vzory maximálně popsat, zavádíme osamělá slovesa, která se chovají jinak než ostatní stejného ukončení, avšak spoléháme na strojové učení (guesser), které v další fázi projektu umožní alespoň částečně automatický výběr z nabízených vzorů a jejich přiřazení ke slovům. Můžeme si proto dovolit některé složitější či obsáhlejší skupiny popsat obecněji a počkat na výsledky guesseru.

Ve fázi projektu, ve které vzniká popis našich vzorů, se ve většině případů nabízí vzorů více, než kolik je jich přiřazeno. Tabulka níže (18) ukazuje míru přegenerování systému vzorů.

Tabulka 18: Počet nabízených vzorů během anotací sloves

počet vzorů	**počet sloves**	**vzory**
1	4 499	20: *vést*, *mást*, *zábst**, *hrýzt**, *péci*, *říci**, *pomoci**, *zkoumat*, *přát*, *smát*, *brát*, *sát*, *zkoušet*, *spět*, *prosit*, *plout*, *tisknout*, *zapomenout**, *krýt*, *být**
2	4 569	
3	4 008	
4	182	*-át*, *-ět*, *-ít*
5	139	*ustát*, *-ít*
6	5	*-ít*

*nejvyšší priorita, odpovídající slovesa mají vždy a pouze daný vzor

Je-li přiřazen právě jeden vzor, je vždy správný. Nemusí tomu tak být u všech sloves daného vzoru, ale pouze u jedné z definic, která odpovídá jen tomuto vzoru a žádnému jinému. V jiném případě může být tentýž vzor nabízen s dalšími vzory.

8.4.2 Přiřazené vzory

Celkem jsme anotovali 13 402 sloves[24]. v následující tabulce uvádíme rozložení těchto sloves mezi různá infinitivní ukončení.

[24] Z původní sady sloves, která jsme získali z korpusu czTenTen na základě charakteristiky ukončení infinitivu, jsme vytřídili ta, která nejsou slovesa (*kabát*), obsahovala gramatické chyby či překlepy a taková, od kterých se imperativ netvoří (*muset*).

Tabulka 19: Rozložení anotovaných sloves mezi infinitivní ukončení

infinitivní ukončení	počet sloves	%
-[sz]t	205	1,53 %
-c[tI]	87	0,65 %
-at	7 090	52,90 %
-ovat	3 684	27,49 %
-át	127	0,95 %
-et	567	4,23 %
-ět	406	3,03 %
-it	3 183	23,75 %
-ít	300	2,24 %
-out	1 392	10,39 %
-ýt	45	0,33 %
celkem	**13 402**	**100 %**

Zajímají nás především skupiny s vysokými čísly, které mohou předznamenat produktivní, neuzavřené vzory (*-ovat*), nebo naopak složitý a roztříštěný podsystém vzorů (ostatní na *-at*[25]).

Další tabulka obsahuje informace o tom, kolikrát byl který vzor nabídnut k anotaci slovesa a kolikrát byl opravdu přiřazen.

Tabulka 20: Úspěšnost při nabízení vzorů během anotací

vzor	nabídnut	přiřazen	úspěšnost	vzor	nabídnut	přiřazen	úspěšnost
-[sz]t	***414***	***205***	**49,52 %**	***-ět***	***922***	***469***	**50,87 %**
nést	98	62	63,27 %	*spět*	406	291	71,68 %
**pást*	6	6	100,00 %	*trpět*	327	121	37,00 %
vést	68	39	57,35 %	*pištět*	162	43	26,54 %
mást	166	56	33,73 %	*vědět*	19	11	57,89 %
**číst*	24	24	100,00 %	**opomnět*	8	3	37,50 %
**nalézt*	36	2	5,56 %	***-it***	***5 608***	***3 513***	**62,64 %**
**ukrást*	1	1	100,00 %	*prosit*	3 183	2 611	82,03 %
**hrýzt*	1	1	100,00 %	*kazit_Z*	200	60	30,00 %
**jíst*	13	13	100,00 %	*zanítit_Z*	5	4	80,00 %
**zábst*	1	1	100,00 %	*učit_Z*	2	2	100,00 %

[25] Z lingvistického hlediska je typ *-at* produktivní, pouze skupina sloves daných výčtem patří k uzavřeným vzorům *brát* a *mazat* (podle klasické mluvnice, např. Mluvnice češtiny). Proto mají imperativ nejednoznačný.

vzor	nabídnut	přiřazen	úspěšnost	vzor	nabídnut	přiřazen	úspěšnost
-ct/ci	***87***	***87***	**100,00 %**	*zjistit*	1 628	507	31,14 %
péci	68	68	100,00 %	*určit*	588	327	55,61 %
**říci*	6	6	100,00 %	**pospíšit*	2	2	100,00 %
**pomoci*	13	13	100,00 %	***-ít***	***1 349***	***348***	**25,80 %**
-at/át	***16 354***	***7 443***	**45,51 %**	*chtít*	106	15	14,15 %
zkoumat	7 081	3 206	45,28 %	*zabít*	2	1	50,00 %
huhňat	4	4	100,00 %	*sít*	194	36	18,56 %
kupovat	3 684	3 661	99,38 %	*žít*	300	99	33,00 %
ukázat	4888	236	4,83 %	*mlít*	12	12	100,00 %
**poslat*	11	11	100,00 %	*jít*	1	1	100,00 %
**kašlat*	182	8	4,40 %	**přijít*	1	1	100,00 %
**lhát*	20	6	30,00 %	*sejít*	22	21	95,45 %
stát	13	11	84,62 %	*třít*	300	102	34,00 %
brát	75	67	89,33 %	*ctít*	106	49	46,23 %
**psát*	21	20	95,24 %	*začít*	300	9	3,00 %
vstát	60	17	28,33 %	*převzít*	4	1	25,00 %
**hnát*	21	17	80,95 %	**vzít*	1	1	100,00 %
řvát	98	35	35,71 %	***-out***	***1 447***	***1 405***	**97,10 %**
zvát_Z	8	8	100,00 %	*tisknout*	1 273	1 264	99,29 %
**chcát*	5	5	100,00 %	*zapomenout*	117	111	94,87 %
sát	63	61	96,83 %	*pomátnout*	9	2	22,22 %
přát	95	50	52,63 %	*plout*	48	28	58,33 %
smát	21	17	80,95 %	***-ýt***	***45***	***45***	**100,00 %**
bát	3	2	66,67 %	*krýt*	31	31	100,00 %
**ustát*	1	1	100,00 %	**být*	14	14	100,00 %
-et	***888***	***583***	**65,65 %**				
zkoušet	567	363	64,02 %				
jet	76	22	28,95 %				
běžet	219	172	78,54 %				
řinčet	26	26	100,00 %				
CELKEM					**27 118**	**14 098**	**51,99 %**

U jednotlivých vzorů můžeme sledovat jejich úspěšnost v tom, zda jsou nabídnuty zrovna u slovesa, k němuž patří. Celková úspěšnost nabízení vzorů je 51,99 %. I to však považujeme za úspěch, neboť se jedná pouze o data připravená pro guesser. Bude-li třeba systém vzorů vylepšit, zjistíme až na základě

výsledků guesseru. Nejhorší výsledek sledujeme u vzoru *začít* (3 %). Jedná se však o vzor pro uzavřenou skupinu sloves. Nepovažujeme za potřebné daná slovesa do vzoru přímo vypisovat. Předpokládáme, že takové případy guesser dokáže vyřešit. Naopak vzory produktivní mají úspěšnost téměř stoprocentní (*tisknout*, *kupovat*).

Při anotacích musíme počítat s tím, že ne vždy jsou tvary od jednoho slovesa tvořeny pomocí jediného vzoru. Sledujeme, v jakých situacích dochází k tzv. variantám, tedy u kterých sloves a vzorů jsme přiřadili vzory dva. Žádnému slovesu nejsou přiřazeny více než dva vzory.

Tabulka 21: Přehled vzorů nabízejících variantní imperativy

vzor1	**vzor2**	**počet sloves**	**příklad**
ukázat	zkoumat	191	škrab/škrabej
kašlat	zkoumat	6	kašli/kašlej
sát	přát	16	hraj/hrej
vstát	stát	4	staň/stůj
běžet	zkoušet	15	šil/šílej
řinčet	zkoušet	2	řinči/řinčej
trpět	spět	41	šum/šuměj
pištět	spět	22	lpi/lpěj
zjistit	prosit	150	vábi/vab
určit	prosit	111	léči/leč
zabít	žít	1	zab/zabij
chtít	žít	2	naděj/nadij
sít	žít	20	lej/lij
chtít	ctít	4	přezděj/přezdi
třít	sít	2	tli/tlej
převzít	vzít_Z	1	vezmi/vem
prosit	kazit_Z	60	has/haš
tisknout	pomátnout	2	pomátni/pomať
celkem		**650 (4,85 %)**	

Ve srovnání s tabulkou č. 18 lze tvrdit, že každé sedmé anotované sloveso, kterému jsou nabídnuty 2 vzory, tvoří variantní imperativy podle obou vzorů.

Závěr

Na to, aby byly mluvnické slovesné vzory použity pro automatickou morfologickou analýzu, až příliš spoléhají na jazykovou intuici. Jeden vzor má často více možností tvorby některých z odvozených tvarů (např. imperativu), a to v systému založeném jak na kmeni prézentním, tak kmeni minulém/infinitivním (kap. 3.3.3). Taková nesystematičnost není v počítačovém zpracování přípustná, vyžaduje následně zbytečné popisování téměř totožných pravidel pro celé paradigma s rozdíly třeba jen u jednoho z tvarů. Proto jsme se rozhodli tvořit přímo na sobě nezávislé vzory pro subparadigmata, tzv. parciální vzory (kap. 1.2.1). Přistupujeme k jednotlivým dílčím vzorům samostatně a předpokládáme mezi nimi silné vazby, které odhalí strojové učení, jež na tento výzkum navazuje. Vyvstanou tak častá spojení např. mezi imperativním a prézentním vzorem, které jsou často přiřazovány ke slovesům společně.

Za možnost přístupu k velkému množství dat, na nichž je třeba ověřovat různé jazykové jevy, vděčíme počítačovému zpracování jazyka a korpusové lingvistice, které jsou dnes v prostředí i tak malého jazyka, jako je čeština, velmi úspěšné a dá se na jejich výsledky spoléhat. Využili jsme největší označkovaný korpus Českého národního korpusu SYN (verze 5), který je textově vyvážený a odráží běžnou psanou podobu českého jazyka. Díky kvalitnímu značkování jsme odhalili spoustu slovesných skupin, které jsou charakteristické nejen společným infinitivním či imperativním ukončením, ale také alternacemi uvnitř sloves při změně řetězce infinitivního tvaru na řetězec znaků odpovídající imperativu. Právě tyto charakteristiky jsou pro tvorbu nových vzorů stěžejní.

Vytvořili jsme systém o 68 vzorech pro imperativy, které pokrývají většinu běžných i okrajových sloves české slovní zásoby. Všechny jsou zapsány i se základním popisem v tabulce 14 (kap. 8). Počet vzorů je poměrně vysoký, důležitějším číslem, na základě kterého je možné hodnotit úspěšnost výzkumu, je počet vzorů nabízených při anotování sloves. Při výběru vzorů se totiž nabízí maximálně 6 vzorů, a to jen v pěti případech u sloves na *-ít* (*zabít*, *jít*, *vzít*), která jsou těžko formálně popsatelná nebo mají stanovenou vysokou prioritu. Slovesa, která spadají pod produktivní vzory (a shodují se s produktivními mluvnickými vzory), jsou pečlivě zapracována a při anotacích jsou téměř jednoznačná (př. slovesa na *-ýt,* vzor *tisknout*). Z tabulky 19 je zřejmé, že téměř u všech sloves, která jsme anotovali, systém nabídl jeden až tři vzory. Necelé třetině byl vzor určen jednoznačně. Další asi třetinu anotací tvoří slovesa, u nichž vybíráme ze dvou vzorů, každé sedmé přísluší právě oběma vzorům. Velkou roli zde hraje i předpoklad, že guesser bude vzhledem k charakteristikám vzorů a přiřazeným prioritám umět procento jednoznačně anotovaných sloves ještě zvýšit (např. u sloves na *-ovat* se nabízí vždy 4 vzory, přičemž dva z nich nikdy nejsou přiřazeny, téměř vždy se jedná o vzor *kupovat* > *-uj,* zřídka *zkoumat* > *-ej*). Přísné hodnocení vzorů, které jsme vytvořili, je úspěšné z 51,99 %. Nejhůře zpracovaná jsou slovesa na *-it*, u kterých jsme stěží hledali společné jevy, dále slovesa na *-at*, která jsou velmi členitá a těžko popsatelná. Popisy vzorů však z praktických důvodů nejsou přehlceny negativními

charakteristikami, které by vylučovaly skupiny slov vždy spadajících pod jiný vzor. Předpokládáme, že strojové učení snadno samo tyto třecí plochy odhalí.

Podrobně jsme se věnovali hláskovým a grafickým změnám, které je nutné ve vzorech explicitně popsat. Tabulky 9 a 10 v kapitole 4 obsahují přehled hláskových alternací, se kterými se běžně setkáváme v českém tvarosloví. Ty a další změny grafické pak analyzujeme vzhledem k vytvořeným vzorům v kapitole 8.2 (tabulky 16 a 17). Výsledky výzkumu potvrzují teorii o tom, že je v jazyce tendence se hláskovým změnám vyhnout, je-li to možné. V mluvnici to dokazuje přechod mezi vzory a třídami, př. 1. třída (vzor *mazat*: *maž*, vzor *brát*: *syp*) přechází ke třídě 5. (*mazej*, *sypej*). I tato tendence je charakteristická pro vzory produktivní (př. v mluvnici i novém systému vzor *tisknout* > *tiskni*, vzor *kupovat* > *kupuj*). Přesto k alternacím dochází velmi často. Typické je měkčení a další změny na koncovém konsonantu, dochází-li k obnažení slovního základu (*vést* > *veď*, *péci* > *peč*, *ukázat* > *ukaž*, *svítit* > *sviť*, *sedět* > *seď*, *minout* > *miň*) přidáním nulové koncovky (v 2. os. sg.). Právě takové imperativy vyžadují při jejich tvorbě nejvíce hláskových/grafických změn. Naopak imperativy na *-j* mění oproti imperativu (krom koncovky) pouze znak odpovídající kmenotvorné příponě, žádné další alternace se u nich nevyskytují. Vokály střídající se s nulou jsou typické pouze pro uzavřenou skupinu sloves, pro vzory, které mají omezené množství přiřazených sloves. Celkově nejrozsáhlejší a tedy typickou změnou při tvorbě imperativů z infinitivu je krácení vokálu napříč všemi druhy sloves, ať už dle infinitivního nebo imperativního zakončení. Velmi časté je také různorodé střídání kvality vokálů.

Stěžejní jsou v systému vzory, které pokrývají nejvíce sloves. Jsou to vzory produktivní a předpokládáme u nich další řazení dosud nezapracovaných (neznámých) sloves. Vedle již zmíněného vyhýbání se hláskovým alternacím je předpoklad takových vzorů odvoditelný také z tabulek 8 a 19, které znázorňují rozložení sloves v úzu i v našem anotovaném vzorku. Nejvíce zastoupena jsou slovesa na *-at* (*-ovat*), *-it* a *-out*. Právě takovou podobu přebírají slovesa převzatá z cizích jazyků (př. *streamovat*, *googlit*, *trumfnout*), která tvoří jednu z největších skupin nových slov v češtině. Úspěšnost vzorů, které generují imperativy od těchto sloves, fungují úspěšně. Vzor *určit* nemá úspěšnost tak vysokou, 55,61 % a nabízí se vždy společně se vzorem *prosit*. Charakteristika těchto dvou vzorů je složitá, některé typy sloves mají dokonce obě varianty imperativu. Vzor *kupovat* má úspěšnost 99,38 %, při anotacích jej vybíráme až ze 4 vzorů. Vzor *tisknout* je úspěšný v 99,29 %, většinou je nabízen sám, případně s jedním dalším vzorem.

Jsme si vědomi jistých nedostatků, které ztěžují ruční anotace. Je však třeba uvědomit si, že se jedná o vzory, na základě kterých budou probíhat anotace automatické pomocí guesseru, který se na ručních anotacích bude učit. Problémy při určování vzorů tak budou buď odstraněny, nebo řešeny až na základě výsledků strojového učení. Nový systém vzorů jako východisko pro navazující práci proto považujeme za úspěšný.

Seznam použité literatury

Tištěné zdroje

ČECHOVÁ, Marie (2011). Čeština – řeč a jazyk. 3., rozšířené a upravené vydání. Praha: SPN – pedagogické nakladatelství, 2011. ISBN 978-80-7235-413-9.

FROULÍKOVÁ, Lenka (2008). *Adam a Eva v Českém ráji: učebnice češtiny pro cizince-začátečníky*. Praha: Academia, 2008. ISBN 978-80-200-1673-7.

HAVRÁNEK, Bohuslav a Alois JEDLIČKA (1992). *Stručná mluvnice česká*. 24., upr. vyd. Praha: Státní pedagogické nakladatelství. Pomocné knihy pro žáky základních škol. ISBN 80-04-26228-7.

HRONOVÁ, Karla a Milada TURZÍKOVÁ (1993). Čeština pro cizince: základní kurs. Praha: Státní pedagogické nakladatelství. ISBN 80-04-23734-7.

KARLÍK, Petr, NEKULA, Marek a Jana PLESKALOVÁ (2016). Nový encyklopedický slovník češtiny. Praha: NLN, Nakladatelství Lidové noviny. ISBN 978-80-7422-480-5.

KOMÁREK, Miroslav (1986). *Mluvnice češtiny 2, Tvarosloví*. Praha: Academia.

OSOLSOBĚ, Klára (1996). *Algoritmický popis české formální morfologie a strojový slovník češtiny*. 1996. Disertační práce. Masarykova univerzita.

OSOLSOBĚ, Klára (2011). *Morfologie českého slovesa a tvoření deverbativ jako problém strojové analýzy češtiny*. Brno: Masarykova univerzita. Spisy Masarykovy univerzity v Brně, Filozofická fakulta. ISBN 978-80-210-5565-0.

OSOLSOBĚ, Klára (2014). *Česká morfologie a korpusy*. 1 vyd. Praha: Karolinum, 235 s. ISBN 978-80-246-2562-1.

PETKEVIČ, Vladimír (2014). *Morfologická homonymie v současné češtině*. Praha: NLN, Nakladatelství lidové noviny. Studie z korpusové lingvistiky. ISBN 978-80-7422-367-9.

RUSÍNOVÁ, Zdenka (1982). *Současná česká morfologie*. Praha: Státní pedagogické nakladatelství. ISBN 80-210-0087-2.

SEDLÁČEK, Radek (1999). *Morfologický analyzátor češtiny*. Brno. Diplomová práce. Masarykova univerzita.

ŠLOSAR, Dušan a Zdenka RUSÍNOVÁ (1994). *Základní kurs češtiny pro cizince*. Brno: Filosofická fakulta Masarykovy university. ISBN 80-210-0007-4.

ŠMERK, Pavel (2010). *K počítačové morfologické analýze češtiny*. Brno. Disertační práce. Masarykova univerzita.

ŠTÍCHA, František (2011). Kapitoly z české gramatiky. Praha: Academia. Lingvistika. ISBN 978-80-200-1845-8.

ŠTÍCHA, František (2013). Akademická gramatika spisovné češtiny. Praha: Academia. Lingvistika. ISBN 978-80-200-2205-9.

ZIKOVÁ, Markéta (2015). Morfonologické alternace v současné češtině. In: Oldřich ULIČNÝ a kol., eds. *Preliminária k moderní mluvnici češtiny*. Olomouc: Univerzita Palackého, s. 187–193. ISBN 978-80-244-4667-7.

Elektronické zdroje

Český národní korpus – SYN [online] (2014). Verze 3. Praha: Ústav Českého národního korpusu FF UK, 27. 1. 2014 [cit. 24. 10. 2016]. Dostupné z: http://www.korpus.cz

Český národní korpus – SYN [online] (2016). Verze 4. Praha: Ústav Českého národního korpusu FF UK, 16. 9. 2016 [cit. 3. 5. 2017]. Dostupné z: http://www.korpus.cz

Český národní korpus – SYN [online] (2017). Verze 3. Praha: Ústav Českého národního korpusu FF UK, 24. 4. 2017 [cit. 11. 3. 2017]. Dostupné z: http://www.korpus.cz

Český národní korpus – SYN2015 [online] (2015). Praha: Ústav Českého národního korpusu FF UK, [cit. 14. 5. 2017]. Dostupné z: http://www.korpus.cz

GRÁC, Marek (2006). *Tvorba morfologické databáze z neoznačkovaného korpusu* [online]. Brno, [cit. 5. 4. 2017]. Dostupné z: http://is.muni.cz/th/50728/fi_m/. Diplomová práce. Masarykova univerzita, Fakulta informatiky. Vedoucí práce Pavel Šmerk.

HUSÁROVÁ, Dagmar (2015). *Určování tvarotvorných vzorů neznámých slov* [online]. Brno, [cit. 5. 4. 2017]. Dostupné z: http://is.muni.cz/th/333279/ff_m/. Diplomová práce. Masarykova univerzita, Filozofická fakulta. Vedoucí práce Marek Grác.

Internetová jazyková příručka (IJP) [online] (2017). Praha: Ústav pro jazyk český AV ČR [cit. 8. 6. 2017]. Dostupné z: http://prirucka.ujc.cas.cz/

DEBDict – obecný prohlížeč slovníků [online]. Brno: Centrum zpracování přirozeného jazyka, FI MU.[cit. 20. 6. 2017]. Dostupné z: https://deb.fi.muni.cz:8005/debdict/

HAJIČ, Jan (2001). *Popis morfologických značek – poziční systém* [online]. Praha: Ústav Českého národního korpusu FF UK [cit. 29. 3. 2015]. Dostupné z: http://ucnk.ff.cuni.cz/doc/popis_znacek.pdf

Kontext [online] (2017). Brno: FI MU, NLP centrum. [cit. 18. 4. 2017]. Dostupné z: https://kontext.korpus.cz/

SEDLÁČEK, Radek (2006a). *Morfologický analyzátor češtiny. Morphological analyser of Czech – ajka* [online]. [cit. 22. 5. 2017]. Dostupné z: http://nlp.fi.muni.cz/projekty/ajka/ajkacz.htm

SEDLÁČEK, Radek (2006b). *ajka tagset* [online]. [cit. 22. 5. 2017]. Dostupné z: http://nlp.fi.muni.cz/projekty/ajka/tags.pdf

Přílohy

Přílohy jsou velmi rozsáhlé, proto je nepřikládáme k tištěné verzi diplomové práce. Jsou uloženy pouze v archivu této závěrečné práce v Informačním systému Masarykovy univerzity, a to v podobě samostatných souborů.

A: Soubor s popisem vzorů: *slovesa_imp.xml*

Jedná se o *.xml* soubor obsahující všechny definované vzory pro tvorbu imperativů z infinitivu.

B: Soubor s popisem hlásek: *vowels.xml*

Soubor obsahuje všechny hlásky a k nim přiřazené charakteristiky.

C: Použitý program: *langusta3.jar*

Příloha C je program, který je potřeba ke spuštění anotací.

D: Anotovaná slovesa + všechny nabízené vzory: *anotac.out*

Textový soubor obsahuje seznam anotovaných slov a k nim všechny vzory, které systém nabízí.

E: Anotovaná slova + přiřazené vzory: *anotace.imp*

Textový soubor obsahuje seznam anotovaných slov s příslušnými přiřazenými vzory.