

Kombinace konsonantu s vokálem v českých slovech cizího původu

ALEŠ BIČAN

Consonant-vowel combinations in Czech words of foreign origin

ABSTRACT: Analyzing a corpus of over 36,000 words of foreign origin and 33,000 native words, the paper presents a detailed account and comparison of the distribution of consonant-vowel combinations in Czech. These combinations are discussed according to their vowel class. Both combinations typical for foreign-origin words and combinations typical for native words were found. In some cases, the typicality is predictable from the phonemes in the combinations (i.e. from the phonemes typical for either type of vocabulary), but there are many combination types for which this is not possible. The paper concludes with a discussion of the factors that may influence the distribution of consonant-vowel combinations, including sound changes, phoneme frequency, morphological processes, and universal preferences in the combinability of consonants with vowels. Special attention is paid to loanwords from English.

Key words: phonotactics, loanword, phonological corpus, Anglicism, expressive word

Klíčová slova: fonotaktika, výpůjčka, fonologický korpus, anglicismus, expresivum

1. Úvod

Slova cizího původu neboli výpůjčky či přejímky z cizích jazyků představují důležitý materiál nejen pro lingvistické teorie, ale i pro jazykové popisy (Paradis – LaCharité, 2005). Vyčleníme-li je jako samostatnou kategorii, nabízí se otázka, jak se liší od slov ostatních, především od slov původu domácího. Rozdíly existují na různých jazykových rovinách a jsou si jich vědomi jak lingvisté, tak laici. Nejmarkantnější jsou rozdíly formální, tj. grafické nebo fonologické. Ty se týkají jak konkrétních jednotek, nejčastěji fonémů (ale i suprasegmentálních prostředků), tak způsobu, jakým se fonologické prvky navzájem kombinují (Haugen, 1950).

Z hlediska ontogeneze řeči, výskytu v jazycích světa a frekvence (Gordon, 2016) je nejzákladnějším typem spojení fonémů kombinace konsonantu s vokálem (dále jen CV kombinace). Pro češtinu popisy tohoto typu spojení sice existují (Ludvíková, 1968; Bičan, 2013), a to včetně údajů o frekvenci jednotlivých kombinací v lexiku a v textech (Ludvíková – Kraus, 1966; Fonologický korpus češtiny), ale již dávno před nimi upozornil Mathesius (1931) na to, že takový popis je neúplný a vlastně i nepřesný, nevezmeme-li v úvahu kontext, v němž se kombinace vyskytuje. Odlišnosti mohou existovat mezi různými druhy lexika, především mezi slovy neutrálními a slovy expresivními nebo mezi slovy domácího a cizího původu, ale také mezi výskytem na morfologickém švu a jinde. Mathesiův krátký text zůstal po dlouhou dobu jediný svého druhu, byť na něj navázaly studie fonologických způsobů vyjádření expresivity v češtině (Zima, 1961; Rechziegel, 1998).

Ačkoliv má na distribuci CV kombinací vliv i původ slov, systematickému popisu rozdílů mezi domácími a přejatými slovy se doposud nikdo nevěnoval. Přesto existují

kombinace, které jsou možné pouze v přejatých slovech, nebo se v nich většinou vyskytují. Existují ovšem i kombinace typické pro slova domácí, které ve výpůjčkách lze nalézt jen okrajově, nebo vůbec. V mnoha případech jejich okrajovost vyplývá z fonémů, z nichž jsou kombinace tvořeny. Některé fonémy se v češtině totiž vyskytují pouze nebo převážně ve výpůjčkách. Stejně tak existují i fonémy s většinovým zastoupením v domácích slovech. Přesto neplatí stoprocentně, že by takové fonémy tvořily kombinace typické buď pro přejatá, nebo domácí slova. A stejně tak existují i kombinace, jejichž typičnost pro určitý typ lexika není možné odvodit z povahy jejich složek. Odhalení takových fonologických zvláštností a jejich charakteristika je hlavní náplní tohoto textu.

Článek má následující strukturu. Druhý oddíl shrnuje předchozí popisy pravidel, kterými se distribuce CV kombinací v češtině řídí, především pravidlo postulované Mathesiem (1931). Třetí oddíl představuje popisovaný materiál čítající více než 36 tisíc přejatých slov a 33 tisíc domácích slov, oddíl 4 pak způsob, jakým lze pomocí procentuálního zastoupení rozlišovat mezi prvky typickými pro určitý typ lexika. V následujících oddílech postupně probírám kombinace podle vokalické třídy, rozděluje je na vokály zadní (odd. 5), střední (odd. 6), přední středové (odd. 7) a přední vysoké (odd. 8). Samostatný oddíl je věnován spojení s /j/ (odd. 9), poněvadž většina takových kombinací náleží cizím slovům. Po podrobném popisu spojení konsonantu s vokálem nastiňuji, jaké faktory mohly vést k rozdílu v míře zastoupení a ve frekvenci jednotlivých kombinací (odd. 10). Oddíl 11 shrnuje hlavní závěry analýzy. Zvláštní pozornost je věnována anglicismům, které v současnosti představují nejvíce se rozšiřující soubor slov cizího původu.

2. Mathesiovo pravidlo

Mathesius (1931) vytyčil dva druhy distribučně zvláštních spojení konsonantu a vokálu. Do první skupiny patří kombinace konsonantů /t/, /d/, /š/, /ž/, /ň/, /ř/, /j/, /č/, /c/, které nazývá palatální, a vokálů /o/, /ó/, /u/, /ú/, /ou/, pro něž jednotný termín nepoužívá, ale které jsou tradičně označovány jako vokály zadní. Ačkoliv se ne všechny uvedené konsonanty pod palatály vždy řadí, budou zde pro zjednodušení takto označovány. Druhým typem spojení jsou kombinace velár /k/, /x/, /h/ a vokálu /e/. Mezi veláry by mělo být zařazeno i /g/, byť ho Mathesius nezmiňuje. Uvedené kombinace jsou podle Mathesia téměř výlučně omezeny na tři kontexty (viz (1)). Určitá spojení konsonantu a vokálu jsou tedy v češtině nedoložená nebo jsou vzácná uvnitř morfému, především uvnitř kmene domácích neexpresivních slov, která ve shodě s Mathesiem budu nazývat *domácí intelektuální slova*.

(1) Distribučně specifické kontexty:

- (a) Morfologická hranice
- (b) Citově zabarvená a zvukomalebná slova (expresivní slova)
- (c) Slova cizího původu (přejatá slova)

Mathesiovo pravidlo bylo často citováno (např. Trnka, 1937; Vachek, 1940; Ludvíková, 1968; Zima, 1961) hlavně v souvislosti s rozdíly mezi intelektuální (neutrální) a expresivní slovní zásobou. Fonologickým vyjádřením expresivity se věnovala Rechzieglová

(Grygarová-Rechzieglová, 1993; Rechziegel, 1998; Rechzieglová, 1998). Ta navrhla rozšíření pravidla týkajícího se palatál o střední vokál /a/ (stejně tak už Zima, 1961, s. 12). Argumentem jsou jí nejen fonetická fakta, ale především to, že podle ní je výskyt spojení palatálních konsonantů (s výjimkou /c/, viz odd. 6) s vokálem /a/ omezen na kontexty (1). Jelikož se věnovala expresivitě, své tvrzení dokládá na expresivních slovech, nikoliv na slovech přejatých. Z výkladu Rechzieglové není bohužel jasné, zda za distribučně specifické pokládá i kombinace s ostatními a-ovými vokály, tedy s /á/ a diftongem /au/.

Další rozšíření Mathesiova pravidla, tentokrát pravidla o vokálu /e/, navrhuje *Mluvnice češtiny I* (Dokulil et al., 1986). V oddíle o anomálních konfiguracích fonémů (s. 158–161) upozorňuje na to, že podobné distribuční omezení jako veláry má s e-ovými vokály i vibranta /r/, tj. takové kombinace se vyskytují především v kontextech (1), byť existují mnohé výjimky.

3. Materiál

Jelikož Mathesiovo pravidlo ve všech jeho rozšířeních nemusí vyjadřovat jedinou pravidelnost nebo tendenci v distribuci konsonantů ve spojení s vokály, rozhodl jsem se takové kombinace podrobit detailnější analýze. Předmětem mého zájmu bude jejich výskyt v cizích slovech a částečně i v expresivech v porovnání s domácími, popř. domácími neexpresivními slovy. Výskytu na morfologické hranici věnuji malou pozornost, jelikož tam lze najít téměř jakoukoliv kombinaci jak v domácích, tak v přejatých slovech. Za morfologickou hranici lze považovat takové místo ve struktuře slova, jež je odhalitelné synchronní morfologickou analýzou. Za expresivní (tj. zvukomalebná a citově zabarvená) budu pokládat slova, jež takový příznak ve výkladových slovnících současně češtiny mají.

Apel na synchronní kritéria se nutně dostává do konfliktu s pojmy „cizí slovo“, „slovo cizího původu“ či „přejímka“ nebo „výpůjčka“. Ty jsou obvykle chápány jako synonyma a takto budou užívány i zde. Označují slova, která jsou sice součástí určitého jazyka, zde češtiny, ale na rozdíl od slov domácích nejsou pokračováním výrazů z předchozího vývojového stádia jazyka, zde praslovanštiny, nýbrž byla v minulosti převzata z jazyků cizích. Jsou to tedy pojmy primárně diachronní. Synchronně obhajitelným kritériem pro označení lexému jako cizího může být jeho zařazení do slovníku cizích slov a skutečně tyto výrazy takto identifikovány bývají (Poplack, 2018). Problematická je ovšem potom skutečnost, že někdy jsou jako cizí vnímána slova domácího původu (např. výraz *robot*, jenž je zařazen do *Akademického slovníku cizích slov*, byť je jeho zdrojem domácí *robota*).¹ Zároveň existuje mnoho lexémů, které jsou sice z etymologického hlediska výpůjčkami,

¹ Ačkoliv lze tento fakt použít k zpochybnění uvedeného kritéria pro vymezení pojmu „cizí slovo“, je vhodnější na něj pohlížet jinak. Slovo *robot* se nejspíše vyznačuje fonologickou strukturou, která je pro domácí česká slova málo obvyklá, zatímco výraz *robota* může mít strukturu obvyklejší (jednou z možných příčin je koncová zavřená slabika v *robot* oproti slabice otevřené v *robota*). Na obecné úrovni jsou to právě takové fonologické zvláštnosti charakteristické pro různé typy lexika, které mě zajímají. Dodejme, že podobně cize působí slovo *kovral* zařazené do slovníku přejatých slov *Výslovnost spisovné češtiny* (VSC), byť je jeho původ opět domácí (*KOberce VRAtisLavice*; Pokorná, 1979).

ale byly již do češtiny natolik integrovány, že je slovníky cizích slov nezmiňují (např. *škola, kalhoty*). Avšak jak analýza ukázala, některé z těchto starých výpůjček se vyznačují fonologickými vlastnostmi, které stále svědčí o jejich cizím původu (např. *réva, junák*).

Poněvadž předmětem zájmu tohoto článku jsou fonologické zvláštnosti cizích slov, zvolený materiál vychází ze slovníků cizích/přejatých slov, zahrnuje tedy výrazy, u nichž lze předpokládat cizí původ a jež jsou jako cizí skutečně vnímány. Dva takové slovníky jsou součástí Fonologického korpusu češtiny (FKČ; viz níže). Jsou jimi seznam přejatých slov z ortoepické příručky *Výslovnost spisovné češtiny* z r. 1978 (VŠČ) a kniha *Akademický slovník cizích slov* z r. 1995 (ASCS).

Volba slovníků VŠČ a ASCS jako zdroje cizích slov s sebou nese několik omezení. V prvé řadě tyto slovníky neobsahují aktuální přejímky. Nedostatek částečně kompenzuje sloučení zvoleného materiálu s Fonologickou databází českých anglicismů (FDČA; Bičan et al., 2020), která obsahuje 4 612 anglicismů. Většinu z nich (4 220) tvoří výrazy, u kterých slovníky *Nový akademický slovník cizích slov* (NASCS), *Nová slova v češtině: Slovník neologizmů* (SN), *Co v slovnících nenajdete* (CSN) a *Slovník spisovného jazyka českého* (SSJČ) uvádějí anglický původ.² Jelikož Fonologický korpus češtiny zahrnuje slovník ASCS, jehož je NASCS druhé a rozšířené vydání, velké množství těchto anglicismů je již obsaženo v primárním materiálu získaném z ASCS a VŠČ.³ Kromě těchto slov FDČA zaznamenává 392 výrazů excerpovaných ze seznamu 28 000 nejfrekventovanějších lemmat podle Českého národního korpusu (báze SYN2015) a z vlastního hesláře autorů FDČA. Tyto výrazy nejsou obsaženy v žádném z hlavních výkladových slovníků (*Příruční slovník jazyka českého* [PSJČ], SSJČ, *Slovník spisovné češtiny pro školu a veřejnost* [SSČ], CSN, SN, ASCS, NASCS a *Slovník současné češtiny* [SSouČ]) a patří mezi nově přejatou slovní zásobu (např. *cheesecake, cloud*). Alespoň částečně tak zaplňují výše zmíněnou absenci aktuálních přejímek ve zkoumaném materiálu, i když pochopitelně angličtina není jediným jazykem, z něhož si čeština v posledních letech slova vypůjčuje. Je však nejčastějším zdrojem.

Zvolený materiál byl dále nevyvážený z opačného hlediska. Slovníky VŠČ a ASCS obsahují souborně 53 432 hesel, ale mnoho z nich tvoří výrazy, které jsou spíše okrajové nebo specifické pro určitou vědní oblast (např. lékařská nebo fyzikální terminologie). Jelikož jsem chtěl zkoumat běžně užívaná slova, bylo nutné materiál omezit. Nabízely se dvě možnosti. Za prvé jsem mohl užít frekvenční kritérium a zvolit výrazy, které mají dostatečnou frekvenci (např. podle seznamů lemmat z Českého národního korpusu). Druhou možností bylo do analýzy zahrnout pouze cizí slova, která jsou zaznamenána v hlavních výkladových slovnících češtiny. Lze totiž předpokládat, že takové výrazy jsou nebo aspoň byly často užívány. Vzhledem k tomu, že jsou slovníky VŠČ a ASCS součástí Fonologického korpusu češtiny, který takové standardní slovníky obsahuje (PSJČ,

² Tyto slovníky byly vybrány ze dvou důvodů. Kromě toho, že uvádějí původ slov, jejich elektronické verze umožňují automaticky filtrovat podle původu slova.

³ FKČ je založen na Databázi heslářů (<<https://lexiko.ujc.cas.cz/heslare/>>), která vychází z prvního vydání ASCS z r. 1995, zatímco FDČA využívá materiál z rozšířeného vydání tohoto slovníku (NASCS) z r. 2005. Obsahuje proto některé anglicismy (např. *stilton*), které jsou zařazené do NASCS, ale nikoliv do ASCS.

SSJČ, SSČ, CSN, SN), rozhodl jsem se vycházet z druhé možnosti. Získaný soubor lexémů jsem rozšířil o slova doložená ve VSC nebo ASCS a zároveň ve *Frekvenčním slovníku češtiny* (FSC), který zpracovává seznam nejfrekventovanějších lemmat z Českého národního korpusu z databáze SYN2000. Konečně jsem do souboru cizích slov zařadil výrazy, jež se opět objevují ve VSC nebo ASCS a jež se nacházejí v seznamu 50 000 nejfrekventovanějších lemmat (s malým počátečním písmenem) vygenerovaném ze souborné databáze všech synchronních korpusů SYN Českého národního korpusu. Takto jsem zkomboval obě výše zmíněné možnosti výběru neokrajových a běžných cizích slov.

Aby bylo možné provést kontrastivní analýzu cizích a domácích slov, bylo nutné získat i databázi domácích lexémů. Zde jsem použil SSČ, který stále patří mezi standardní výkladové slovníky češtiny. Z něj jsem nejdříve automaticky odstranil cizí slova, která jsou součástí hesláře ASCS či VSC, a dále 42 vlastních jmen (zkoumám apelativa). Ručně jsem potom vyhledal pravopisné varianty přejímek a jejich zjevné deriváty a přesunul je do databáze cizích slov. Takto jsem dospěl k souboru potenciálně domácích slov. Jak již bylo řečeno, z faktu, že slovo není uvedeno ve slovníku cizích slov, sice nutně nevyplývá, že se jedná o slovo původu domácího, ale je to velmi pravděpodobné. Navíc takové výrazy nebudou přinejmenším působit cize.

Po získání databáze domácích slov jsem seznam cizích slov dále upravil. Vyřazena byla kompozita zjevně obsahující domácí složku, jelikož je nelze jednoznačně pokládat za cizí ani domácí (např. *stometrovy*). Zachoval jsem však výrazy, které obsahují domácí prefix nebo sufix (např. *přeinštalovat*, *marketingový*). Bylo sice možné odstranit i je, poněvadž taková slova mohou obsahovat určitou CV kombinaci právě v domácím afixu, ale nakonec jsem se přiklonil k tomu je ponechat. Srovnání výskytu v domácích sufixech u cizích slov oproti výskytu v přejatých kmenech a na morfologické hranici mezi přejatým a domácím či přejatým afixem ukázalo, že pouze v sedmi případech lze hovořit o vlivu domácích afixů na četnost CV kombinací (u /ní/, /ví/, /ře/, /kí/, /ňe/, /va/, /ti/).⁴

Výsledkem právě popsaného postupu jsou dva soubory slov (lemmat), jejichž základní charakteristiky shrnuje tabulka 1. Databáze cizích slov (dále CS) má rozsah 36 567 unikátních lemmat (typů) a je v ní doloženo 120 949 CV kombinací (tokenů). Databáze domácích slov (dále DS) obsahuje 33 966 lemmat a 96 526 kombinací.

V obou souborech je každému slovu přiřazena fonologická forma přeepsaná jako posloupnost fonémů s určením slabičné a slovní hranice. Fonologická forma vychází z výslovnosti, která je uvedena ve slovnících nebo kterou slovníky implikují (u slov, jejichž výslovnost v podstatě odráží pravopis, např. *automobil*).

Specifikem cizích slov je častá výslovnostní variantnost, jež je obvykle důsledkem konkurence mezi aproximací původní výslovnosti a výslovností podle pravopisu (Duběda, 2018). Jak dokládá Štěpánová (2013, 2015, 2016), variantnost se z velké části týká vokalickej kvantity (např. *procedura* [-du:ra] × [-dura]), znělosti (*diskuse* [-kuse] × [-kuzε]) nebo výslovnosti pravopisných skupin *Vi* (V = vokalickej písmeno, např. *kofein* [kofeji:n])

⁴ Tyto případy jsou popsány v textu. U všech z nich více než 50 % jejich výskytu v cizích slovech připadá na domácí afixy.

Tabulka 1: Základní charakteristiky analyzovaného materiálu

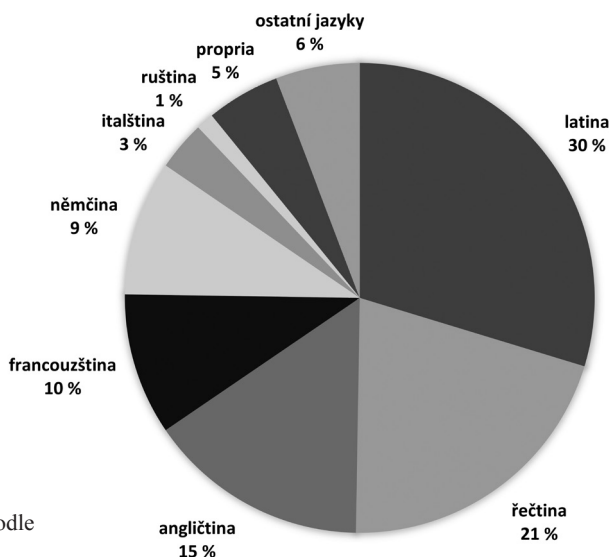
Databáze	Zdroje	Úpravy zdrojů	Počet slov (typů)	Počet CV kombinací (tokenů)
Cizí slova (CS)	ASCS, VSČ	zahrnutý pouze výrazy a) doložené v PSJČ, SSJČ, SSČ, CSN, SN a FSC b) doložené mezi 50 000 nejfrekventovanějšími lemmaty (ČNK: SYN) sloučeno s FDČA odstraněna kompozita s domácí složkou přidány deriváty a pravopisné varianty ze SSČ	36 567	120 949
Domácí slova (DS)	SSČ	odstraněny a) výrazy obsažené ve VSČ a ASCS b) vlastní jména c) deriváty a pravopisné varianty slov z VSČ a ASCS	33 966	96 526

× [kofejn]). I když výslovnostní variantnost nelze při zkoumání fonologie cizích slov zcela ignorovat, v analýze jsem se rozhodl pro výběr jedné reprezentativní výslovnosti, tj. jedné fonologické formy. Zařazení variant by totiž rozšířilo soubor cizích slov o téměř 3 300 lemmat, pro něž databáze cizích slov výslovnostní variantnost zaznamenává, což by vedlo k větší nevyváženosti v rozsahu zkoumaných lexikálních souborů. Navíc by varianty zkreslily kvantitativní údaje o výskytu některých CV kombinací. Např. skupina /ku/ by se vyskytovala dvakrát ve formách /diskuse/ a /diskuze/, které by však odpovídaly jednomu lemmatu *diskuse*.

Databáze cizích slov obsahuje výslovnostní údaje ze tří slovníků, ASCS, VSČ a SSJČ (samozřejmě pouze pro slova, která jsou v nich zaznamenána). Jako reprezentativní byla preferována výslovnost z novějších slovníků, tj. ASCS oproti VSČ a SSJČ, SSJČ oproti VSČ. Pokud slovník uvádí více výslovnostních podob, zvolil jsem první uvedenou. Přihlédl jsem i k výzkumům výslovnostního úzu popsanych Štěpánovou ve výše citovaných statích, dále k výslovnosti, kterou uvádí nejnovější ucelený výkladový slovník SSouČ, a k výslovnostním pravidlům, jimiž se zjevně řídí hesla se záhlavím A v *Akademickém slovníku současné češtiny*.⁵ Vlastní zkoumání uzuální výslovnosti nebylo provedeno.

V databázi cizích slov je zaznamenán i původ slov tak, jak ho uvádějí slovníky ASCS a SSJČ. Informace o původu není dostupná pro každé slovo, protože jednak některé výrazy nejsou součástí zmíněných slovníků, jednak ani tyto slovníky údaje o původu všech lexémů nenabízejí. Obrázek 1 zobrazuje rozložení slov podle původu pro 24 626 výrazů, pro něž je zdroj původu dostupný. Z FDČA jsou započítány pouze anglicismy, které nejsou deriváty jiných anglicismů. Jak již bylo řečeno, ASCS a SSJČ neuvádějí původ u všech lexémů a takovými položkami jsou především (potenciální nebo skutečné) deriváty jiných výrazů. Rozložení jednotlivých jazyků je tedy orientační a nemusí odpovídat skutečným podílům v celé databázi.⁶

⁵ Zbývající hesla nebyla v době dokončení tohoto článku zveřejněna. Pravidla se týkala např. výslovnosti chemických názvů zakončených na *-in* (např. *amin*). Analogicky jsem posuzoval ostatní názvy takto zakončené.



Obrázek 1: Rozložení cizích slov podle původu (pro soubor 24 626 slov)

Z povahy zvoleného materiálu vyplývají některé další důsledky a omezení. Kromě již zmíněných faktorů (omezení se na lemmata doložená ve slovnících, nezohlednění výslovnostních variant) předkládanou analýzu ovlivňují dvě důležité okolnosti. Za prvé, analýza bere v potaz pouze spisovnou zvukovou realizaci slov (např. *býk* místo *bejk*), ačkoliv reálné nepřipravené projevy zřídka mívají ortoepickou podobu. Nicméně ortoepická nebo spisovná výslovnost slouží jako standard nebo ideál, podle něhož jsou nepsisovné realizace hodnoceny, a proto i níže prezentované závěry mohou sloužit jako referenční materiál pro budoucí výzkum mluvené češtiny. Distribuci CV kombinací v nepsisovných podobách slov nechávám zcela stranou.

Druhé a podstatnější omezení se týká absence různých flektivních tvarů slov. Slovníky totiž zaznamenávají pouze reprezentativní podoby lexémů, což jsou u sklonných slov nominativy singuláru a u sloves infinitivy. Některé CV kombinace se však vyskytují právě ve flektivních tvarech, např. /ňů/ v *sršňů*. Jelikož zde zkoumaný materiál vychází ze slovníkových tvarů (lemmat), nemůže být takový výskyt zohledněn. Přesto má toto omezení i své výhody, protože umožňuje konstatovat, že jisté CV kombinace nejsou v základních tvarech slov doloženy. Jak z hlediska laického, tak lingvistického lze nominativy singuláru a infinitivy pokládat za základní podobu lexému. Dále je třeba

⁶ Údaje o původu byly primárně čerpány z ASCS, přičemž jako zdrojový jazyk je chápán ten, ze kterého bylo slovo do češtiny přejato, bez ohledu na to, zda se jedná o výraz původní, nebo vypůjčený z jiných jazyků (např. *budget* je anglicismus, byť byl do angličtiny přejat z francouzštiny). U slov, která nejsou v ASCS, jsem údaje čerpal ze SSJČ. Případné rozpory mezi těmito slovníky jsem podrobněji nezkoumal. Jako původ mnoha slov oba slovníky uvádějí vlastní jména bez ohledu na zdrojový jazyk (např. *camembert*). Dodejme, že v grafu nejsou zohledněna hybridní kompozita obsahující prvky z různých jazyků (např. *kinosál*, kde je *kino* z řečtiny a *sál* z němčiny).

si uvědomit, že začleněním flektivních tvarů bychom získali především informace o výskytu CV kombinací na hranicích kmene a domácího flektivního afixu nebo přímo v domácím flektivním afixu (např. /řo/ a /vi/ v *císařovi*).

Poněvadž jsou zvolené primární zdroje slov omezené, ověřuji výskyt CV kombinací ve Fonologickém korpusu češtiny. Tento korpus je fonologicky transkribovanou databází lemmat a celých textů. Jeho lexikální část zahrnuje lemmata (především apelativa) z hlavních slovníků češtiny (PSJČ, SSJČ, VSC, SSC, CSN, SN, ASCS a dalších) a dále vlastní jména a názvy z různých databází (názvy obcí a jejich částí, zoologické a botanické názvy, křestní jména a příjmení, názvy hradů a zámků). Textová část korpusu obsahuje fonologické transkripty českých textů, převážně beletrie. Počet slov (typů) v celém FKČ přesahuje hodnotu 460 000 položek.⁷

V následující analýze budu pro zjednodušení mluvit o výskytu CV kombinací v cizích a domácích slovech, jako by se tyto závěry a zobecnění týkaly všech cizích a domácích slov v češtině. Primárně se nicméně všechny závěry vztahují na analyzované databáze a jejich obecná platnost je proto pouze aproximativní. To ovšem platí pro všechny výzkumy částečných vzorků z celé populace jevů, především pak pro jazyková data, která jsou díky otevřenosti a dynamičnosti jazyka téměř vždy dílčí a omezená.

4. Zastoupení fonologických prvků

Při porovnávání rozdílů v kombinovatelnosti konsonantů s vokály se nebudu omezovat na skupiny, jež jsou doloženy výhradně v cizích, nebo pouze v domácích slovech. Téměř každá kombinace se totiž objevuje v obou typech lexika. Domácí a cizí slova se však odlišují v míře, v jaké jsou kombinace zastoupeny. Zastoupení fonologického prvku (fonému, kombinace fonémů apod.) označuje, kolik procent z jeho celkového výskytu připadá na jeden ze dvou typů lexika. Hodnota zastoupení umožňuje definovat prvky typické pro cizí, popř. domácí slova. Za prvky typické pro cizí slova pokládám ty, které v nich mají zastoupení větší nebo rovné 90 %, zatímco prvky se zastoupením menším nebo rovným 10 % pokládám za typické pro slova domácí. V tabulkách níže rozlišuji mezi dalšími mírami zastoupení (viz tab. 2).

Zastoupení lze samozřejmě vyčíslit pro jakýkoliv fonologický prvek. V případě CV kombinací je nejrelevantnější informací zastoupení jednotlivých fonémů. Fonémy vyskytující se pouze v CS jsou velára /g/, diftong /eu/ a dále znělé afrikáty /dz/ a /dž/. Kromě nich můžeme za typicky cizí fonémy pokládat labiálu /f/ a vokály /au/ a /ó/, jejichž zastoupení přesahuje 90 %. V DS se /f/ objevuje ve slovech *doufat* a *zoufat* (a derivátech) a dále ve slovech citoslovečného původu a jejich derivátech. Některé si ponechávají svůj expresivní charakter (např. *fičet*, *hafan*), jiné jsou již zcela neutrální (např. *výfuk*). Vokály /au/ a /ó/ nejsou v DS doloženy jinde než v citoslovcích *mňau* a *hó*, *jó*, *vijó*.

⁷ Viz <<http://www.ujc.cas.cz/phword>> a Bičan et al. (2020, s. 27–28). Údaje se vztahují k druhé, dosud nepublikované verzi korpusu, která je oproti publikované verzi rozšířená např. o příjmení a slovní tvary excerpované z Českého národního korpusu.

Tabulka 2: Kategorizace zastoupení a doložení fonologických prvků v cizích a domácích slovech a ve Fonologickém korpusu češtiny

Zastoupení	Symbol	Vysvětlení
100 %	100	doloženo pouze v CS
větší než 90 %	> 90	typické pro cizí slova, doloženo více než 90 % v CS
větší než 66 %	> 66	vyskytující se spíše v cizích slovech, doloženo více než 66 % v CS
v rozpětí 34–66 %	≈	srovnatelné zastoupení v cizích a domácích slovech, v CS a DS
menší než 33 %	< 33	vyskytující se spíše v domácích slovech, doloženo méně než 33 % v CS
menší než 10 %	< 10	typické pro domácí slova, doloženo méně než 10 % v CS
0 %	0	doloženo pouze v DS
bez zastoupení	×	v CS a DS nedoloženo, nedoloženo ani ve FKČ
bez zastoupení	×	nedoloženo ani v CS, ani v DS, ale doloženo ve FKČ

Výčet typicky cizích fonémů není překvapivý a potvrzuje dobře známá fakta. Co ale v popisech nebývá zmíněno, je existence fonémů typických pro domácí lexikum. Kupř. pro angličtinu lze za typicky domácí foném pokládat znělou frikativu /ð/ (např. ve slově *this*), která se až na malé výjimky ve slovech cizího (tj. neanglosaského) původu neobjevuje. V případě češtiny se žádný z fonémů nevyskytuje výhradně v domácích slovech, avšak typickými domácími fonémy jsou konsonanty /ř/, /ʀ/, /dʀ/ a diftong /ou/. Všechny mají v CS menší než 10% zastoupení.

Jestliže se určité fonémy vyskytují převážně nebo výhradně v jednom typu lexika, je nasnadě předpokládat, že to platí i pro jejich kombinace s jinými fonémy. U všech fonémů typických pro cizí slova se tato hypotéza potvrzuje. Jejich CV kombinace jsou doloženy buď zcela v CS (/dz/, /dž/, /g/, /eu/), nebo mají více než 90% zastoupení (/f/, /au/, /ó/). Jedinou výjimkou je skupina /ňau/, která se objevuje pouze v DS, a to v již zmíněném citoslovci *mňau*. V případě fonémů typických pro domácí slova se ovšem předpoklad nepotvrzuje takto silně. Většina takových CV kombinací sice v CS nepřesahuje 33% hranici, ale objevují se i podstatné výjimky. Kromě /tʰá/, která má srovnatelné zastoupení v obou typech lexika, sem patří skupiny /dʰa/, /dʰá/ (obě zastoupení v CS větší než 66 %) a /jou/ (100% zastoupení v CS). Výjimkám se podrobně věnuji níže.

Jelikož srovnávám dva soubory slov, tj. CS a DS, hodnota zastoupení také vyjadřuje poměr mezi výskytem CV kombinace v cizích slovech a výskytem v domácích slovech. Výskyt v expresivech oproti intelektuálním slovům není kvantifikován, protože tento rozdíl je relevantní především pro domácí slova a zde zkoumám rozdíly mezi cizími a domácími slovy. Stejně tak nejsou kvantifikovány rozdíly ve výskytu v kmeni oproti morfologické hranici (popř. oproti afixům), jelikož slovo posuzují jako celek.⁸ Hodnota zastoupení rovněž neříká nic o celkovém počtu konkrétních kombinací nebo o tom, zda je rozdíl mezi výskytem v CS a DS statisticky významný. Primárně analýzu zakládám na hodnotě zastoupení, jež vyjadřuje poměr, jakým se cizí slova jako celek podílejí na

⁸ Výjimkou jsou domácí afixy přítomné v cizích slovech, u nichž jsem posuzoval míru, do jaké ovlivňují výskyt CV kombinací v cizích slovech.

výskytu CV kombinací. Teprve je-li tento poměr větší než 90 % (= kombinace typická pro cizí slova) nebo menší než 10 % (= kombinace typická pro domácí slova), zkouám, v jakých slovech (intelektuální × expresivní) nebo na jakém místě (kmen × morfologická hranice, popř. afix) se kombinace vyskytuje. Převládá-li určitý typ slov nebo místo výskytu, je tato skutečnost zmíněna.

Seznam CV kombinací, jejich frekvence a zastoupení je k dispozici na webových stránkách Fonologického korpusu češtiny (viz Zdroje za textem).

5. Kombinace se zadními vokály

Jednu z distribučně specifických kombinací tvoří palatály se zadními vokály. Jejich zastoupení v CS je shrnuto v tabulce 3.⁹ Mathesiovo pravidlo predikuje, že se tyto kombinace vyskytují především v kmenech expresivních nebo cizích slov a dále na morfologické hranici. Analyzovaný materiál predikci potvrzuje, avšak mnohé z kombinací jsou v CS vzácné, nebo se v nich vůbec neobjevují. Charakterizují tedy spíše domácí slova. Výjimky představují kombinace s fonémy typickými pro cizí slova a kombinace s vokálem /ú/ a konsonantem /j/ (k němu viz odd. 9).

Tabulka 3: Procentuální zastoupení kombinací palatál se zadními vokály v CS (pro symboly viz tab. 2)

	/t'/	/d'/	/ň/	/š/	/ž/	/ř/	/j/	/č/	/c/	/dž/	/dz/
/o/	< 10	0	< 10	≈	≈	< 33	> 90	< 33	< 33	100	100
/ó/	×	×	×	100	100	×	> 90	100	100	100	×
/u/	< 33	×	< 33	≈	> 66	0	> 90	≈	< 33	100	×
/ú/	×	×	100	100	100	×	> 90	0	0	100	×
/ou/	0	0	0	≈	0	0	100	< 33	≈	×	×

V prvé řadě je zjevné, že v CS jsou vzácné kombinace s palatálními okluzivami /t'/ a /d'/ . Pro /t'/ jsou doloženy pouze skupiny /t'o/ (1× v *potomkinský*, jinak na morfologické hranici, např. *perleťový*) a /t'u/ (2×, *baťuška*, *ťubetějka*) a pro /d'/ se v CS žádné kombinace neobjevují. Málo doložené jsou i kombinace s /ň/ : /ňu/ (ve *šňupat* a derivátech), /ňo/ (např. *kaňon*) a /ňu:/ (k /u:/ viz níže). Pro foném /ř/ CS dokládají jenom skupinu /řo/, která stojí výhradně na morfologické hranici (např. *císařovna*).

Co do počtu kombinací existuje v CS nejvíce spojení s palatálními frikativami, chybí pouze /žou/ (srov. /šo/ *šok*, /šó/ pouze v anglicismu *shortlist*, /šu/ *šutr*, /šú/ *kešú*,¹⁰ /šou/ *show*, /žo/ *žold*, /žó/ pouze v *demižón*, /žu/ *žurnál*, /žú/ *abažúr*). S výjimkou /šú/, /žú/, /šó/ a /žó/ žádná z nich nemá v CS větší než 90% zastoupení. Kombinace /šó/ a /žó/ mají každá po jednom dokladu.

⁹ Z tabulky 3 a dalších tabulek lze také vyčíst, které CV kombinace jsou v češtině doložené. Jsou to všechny, které nejsou označeny křížkem v bílém pozadí. Křížek v šedém pozadí označuje kombinace doložené ve Fonologickém korpusu češtiny, avšak nikoliv ve zde zkoumaných databázích cizích a domácích slov.

¹⁰ Tvar *kešú* s dlouhým /ú/ je kodifikovaná podoba slova. Podle Internetové jazykové příručky (heslo *kešú*) se v úzu častěji vyskytuje *kešu* s krátkým /u/ (viz <<https://prirucka.ujc.cas.cz/>>).

Velký počet kombinací také tvoří afrikáty /c/ a /č/, avšak s výjimkou skupin s /ó/ nejsou opět žádné z nich specifické pro cizí slova. Navíc kombinace s dlouhým /ú/, tj. /čů/ a /ců/, se v CS neobjevují (viz též níže). Ostatní kombinace doklady mají (např. /čo/ *čokl*, /čó/ *čór*, /ču/ *čufy*, /čou/ pouze v anglicismu *choke*, /co/ *cop*, /có/ *tencóna*, /cu/ *tsunami*, /cou/ *coul*). Kombinace se znělými afrikátami se objevují výhradně v CS (/dzo/ *podzol*, /džo/ *bendžo*, /džó/ *stagiona*, /džu/ *džus*, /džú/ *jukebox*).

Zvláštní pozornost si vyžadují kombinace s vokálem /ú/. Ty jediné lze totiž pokládat za charakteristické pro cizí slova, ačkoliv je lze nalézt i na morfologické hranici. Druhý typ výskytu není v materiálu téměř doložen, protože morfologickou hranicí je v tomto případě hlavně šev s pádovou koncovkou -ů (srov. *chořů*, *sledů*, *tchořů*, *koňů*, *hrabošů*, *nožů*, *rýčů*, *hřebců*, *zlodějů*). Další výjimkou je /jú/, která se v DS objevuje jen v citoslovci *jú* (viz odd. 9). Kromě těchto dokladů jsou v materiálu kombinace s /ú/ zastoupeny výhradně v CS.

Se zadními vokály se spojují i jiné konsonanty než palatální. Preferencí pro cizí slova se vyznačují pouze dvě skupiny spojení konsonantu s vokálem (viz tab. 4). Za prvé je to kombinace /mu/, která má v CS 93% zastoupení. Tento fakt ovšem musíme přičítat velkému množství přejímek zakončených na *-ismus* (např. *romantismus*), a proto výrazné zastoupení /mu/ svědčí spíše o preferenci určitých typů slov než konkrétních kombinací fonémů. Druhou skupinu spojení konsonantu s vokálem typických pro cizí slova tvoří kombinace s fonémy /g/, /f/, /ó/. Všechny tři fonémy jsou však celkově typické pro cizí slova. V DS jsou doloženy pouze dvě kombinace s /f/, a to /fu/ a /fou/. Objevují se ve slovech s kořenem *fuk/fuč/fouk* (*náfuka*, *fučet*, *foukat*), jež byla původně expresivní povahy. Kombinace /fou/ se dále vyskytuje v expresivech *fouňa* a *fous*. Ačkoliv tato kombinace obsahuje jak foném typický pro domácí slova (/ou/), tak foném typický pro cizí slova (/f/), její výskyt převažuje v DS (32 dokladů). V CS /fou/ najdeme v osmi výrazech, z nichž jsou všechny kromě *profous* anglicismy (např. *smartphone*).

Tabulka 4: Procentuální zastoupení kombinací labiál, velár, /r/ a alveolár se zadními vokály v CS (pro symboly viz tab. 2)

	/p/	/b/	/m/	/f/	/v/
/o/	< 33	≈	> 66	100	< 33
/ó/	100	100	100	100	100
/u/	> 66	> 66	> 90	> 90	> 66
/ú/	< 10	≈	≈	100	< 33
/ou/	< 33	< 33	< 10	< 33	0

	/k/	/g/	/x/	/h/	/r/
≈	100	≈	< 33	≈	
100	100	100	> 66	100	
> 66	100	< 33	≈	≈	
≈	100	< 10	< 33	< 33	
< 33	100	0	< 33	< 33	

	/t/	/d/	/n/	/s/	/z/	/l/
/o/	> 66	≈	≈	≈	> 66	≈
/ó/	100	100	100	100	100	100
/u/	> 66	≈	≈	> 66	> 66	≈
/ú/	> 66	< 33	≈	> 66	< 33	> 66
/ou/	< 33	< 33	< 10	< 10	< 10	< 33

Za zmínku stojí i skupina /tú/, která sice má v CS pouze většinové zastoupení (82 %), ale vyznačuje se několika zvláštnostmi. Kombinace je kromě slov jako *túra* nebo *blue-tooth* možná i ve výrazech zakončených na *-ura* (např. *kultura*, *inventura*), pokud se v nich toto zakončení vyslovuje s dlouhým [u:] (viz odd. 3). Kdybychom takovou výslovnost zohlednili v analyzovaném materiálu, podstatně by se zvýšilo zastoupení skupiny /tú/ v CS (na 95 %). Patřila by tedy mezi prvky typické pro cizí slova. Nebyl by to jediný důvod. Svou roli hraje i fakt, že ve vzorku je pouze devět domácích slov s kombinací /tú/. Kromě expresivních výrazů (*tú*, *tutú*) a zdobnělin (*potůček*) jsou to slova *stůl*, *tůně* a jejich deriváty. Lze tedy konstatovat, že domácí slova se této kombinaci spíše vyhýbají.

Na závěr ještě uvedme, že veláry, labiály a alveoláry též tvoří některé kombinace, které jsou typické pro slova domácí, jelikož mají v CS menší než 10% zastoupení. Patří sem čtyři spojení s diftongem /ou/ (/sou/, /zou/, /mou/ a /nou/), což vzhledem k celkovému postavení tohoto vokálu není překvapivé. Pozoruhodné je to, že se jedná o kombinace s alveolárními frikativami a nazálami (včetně /ňou/ s nulovým zastoupením v CS, viz tab. 3). Cizí slova se tedy takovým typům kombinací vyhýbají. Dalšími skupinami typickými pro domácí slova jsou kombinace /pú/ a /xú/. Na rozdíl od předchozích čtyř kombinací nejsou příliš frekventované a jejich zvýšený výskyt v DS je důsledek velmi malého počtu cizích slov, která je obsahují (6 pro /pú/, např. anglicismus *whirlpool*, 1 pro /xú/, *ischurie*).

6. Kombinace se středními (nízkými) vokály

Jak už bylo zmíněno v oddíle 2, na kontexty (1) jsou podle Rechzieglové omezeny i kombinace palatál s vokálem /a/. Analyzovaný materiál tento závěr nepotvrzuje, alespoň ne v takové míře jako u zadních vokálů (viz tab. 5). Některé kombinace se skutečně objevují v expresivních slovech (/ʔa/ *tapat*, /ňa/ *chňapat*, /řa/ *řach*, /čal/ *čadit* a /ca/ *cavky*) a všechny najdeme i v kmeni cizích slov (/ʔa/ *keʔas*, /dʔa/ *madʔal*, /ša/ *bešamel*, /ža/ *seržant*, /ňa/ *koňak*, /řa/ pouze v rusismu *odřad*, /čal/ *čaj*, /ca/ *car*). Jsou ovšem doloženy i v kmenech řady domácích intelektuálních slov (např. /ʔa/ *šťastný*, /dʔa/ *dʔas*, /ša/ *však*, /ža/ *žalud*, /ňa/ *ňadra*, /řa/ *řada*, /čal/ *čas*, /ca/ *zrcadlo*).

Tabulka 5: Procentuální zastoupení kombinací palatál se středními vokály v CS (pro symboly viz tab. 2)

	/ʔ/	/dʔ/	/ň/	/š/	/ž/	/ř/	/j/	/č/	/c/	/dž/	/dz/
/a/	< 33	> 66	< 33	> 66	< 33	< 10	> 66	< 33	≈	100	×
/á/	≈	> 66	< 33	≈	< 33	0	> 90	< 10	≈	100	×
/au/	×	100	0	×	×	×	100	100	×	100	×

Bližší pozornost si zaslouží /dʔ/ a /š/. Skupiny /dʔa/, /dʔá/ a /dʔau/ jsou omezeny na morfologickou hranici a cizí slova (např. /dʔa/ *madʔal*, /dʔá/ *dʔábel*, /dʔau/ pouze v *dʔaur*), tj. nevyskytují se v kmeni domácích slov. Jedinou výjimkou je výraz *dʔas*. Ten sice není zařazen do slovníků cizích slov (a tudíž naší databáze CS), ale původně je slovo tabuovou

zkráceninou *d'ábel*, jež je cizího původu.¹¹ Zastoupení /d'a/ je v CS 71 % a v případě /d'á/ 80 %, což znamená, že ačkoliv je foném /d'/ jedním z prvků typických pro domácí slovní zásobu, ve spojení se středními vokály ho najdeme více v cizích než domácích slovech.

Distribučně omezené jsou i kombinace středních vokálů s frikativou /š/. Necháme-li stranou jejich výskyt na morfologické hranici, který tvoří většinu dokladů kombinace /ša/, pak se v DS tato skupina nenalézá jinde než ve slovech *však*, *všade*, *ošatka*, která jsou etymologicky skutečně domácího původu, a dále ve slovech *šaty*, *lišaj*, *rozšafný* a *šašek* (a jejich derivátech), jež jsou ovšem z etymologického hlediska starými výpůjčkami. Kombinace /šá/ je v DS přítomná pouze v *šálit*, které je bez jasné etymologie, a v *šátrat* (a derivátech), u něhož se uvažuje o expresivním původu. Zbývající výskyty skupin /ša/ a /šá/ připadají na CS (např. *šach*, *rošáda*). Spojení /š/ a středních vokálů, především /á/, lze tedy hodnotit jako specifikum kmenů cizích slov a morfologických hranic.

Jsou-li kombinace /d'/ a /š/ s vokálem /á/ typické pro cizí lexikum, nabízí se otázka, zda totéž neplatí i pro ostatní palatály. Potvrzuje se to pouze pro kombinaci /já/ s 92% zastoupením (viz dále odd. 9). Kombinace /řá/ se v CS neobjevuje (srov. domácí *řád*). Kombinace /žá/ tam sice doložena je (např. *žánr*), ale větší zastoupení má v DS (např. *žár*). Zbývající skupiny proberu podrobněji.

Skupina /řá/ je v kmeni cizích slov doložena jenom v *křišťál*. Kromě morfologické hranice odpovídají ostatní výskyty hlavně expresivním slovům (např. *tápota*). Skupinu lze nalézt i v kmeni domácích slov, a to ve *šťáva*, u něhož však není vyloučen expresivní původ (Bičan, 2018, s. 17–18). Dále se objevuje v domácích *dřišťál* a *košťál*, jež jsou zakončena stejně jako cizí *křišťál*. Podle Machka (1968) měl ve slově *košťál* sufix *-ál* spolu s doprovodným změkčením předchozího konsonantu expresivní funkci. V *dřišťál* je koncové *-ál* pravděpodobně jiného původu. Přesto lze kombinaci /řá/ s ohledem na její výskyt pokládat za distribučně omezenou hlavně na kontexty (1).

I spojení /ňá/ se objevuje v cizích slovech, jako např. *tuňák*. Vyskytuje se dále na morfologické hranici, přičemž mnoho takových slov má navíc expresivní charakter (*kachňátko*, *mličňák*). V kmeni domácích slov není doloženo jinde než ve výrazech *luňák* a *maňásek*, které nemají jasnou etymologii. Kombinace /řá/ a /ňá/ proto nelze jednoznačně hodnotit jako specifikum určité slovní zásoby, byť jsou časté v expresivech.

Konečně pozornost věnujme afrikátám. Ačkoliv Mathesius (1931) zařazoval /c/ pod palatály a zároveň s tím i pod své fonotaktické pravidlo, Rechzieglová (Grygarová-Rechzieglová, 1993; Rechziegel, 1998) takovou klasifikaci zpochybňuje.¹² Následně ani tuto afrikátu nezačleňuje do svého pravidla týkajícího se vokálu /a/. Distribučně specifická je podle ní pouze skupina /ča/. Analýza materiálu dokládá, že se tato kombinace skutečně objevuje na morfologické hranici, v expresivních slovech (např. *čadit*, *čahoun*) a ve slovech cizího původu (např. *čaj*). Je doložena i v kmeni domácích intelektuálních slov,

¹¹ Neexistují-li zásadní rozpory ve výkladech, etymologická vysvětlení čerpám z Rejzka (2015).

¹² Bičan (2018) nicméně dokazuje, že je zařazení /c/ pod Mathesiovo pravidlo oprávněné.

jako např. *čas, čapka, častovat, čarovat, kolčava*. Není ovšem bez zajímavosti, že zmíněná intelektuální slova nemají jasné etymologické vysvětlení a nelze u nich vyloučit expresivní nebo cizí původ.

Podobné postavení má i kombinace /čá/. Vyskytuje se ve slovech, která lze hodnotit jako expresivní (*brčál, burčák*), a dále ve slovech cizího původu (např. *čáka*). Kombinace je doložena i v několika domácích intelektuálních slovech, jako *část, čáry, čára, čáp* nebo *pečárka*, ale ani ona nemají jasný původ; některá z nich se navíc odvozují od expresiv (*čáp*). S ohledem na velké množství výjimek nemůžeme jednoznačně říct, že by byly kombinace /č/ se středními vokály specifické pro expresivní nebo cizí slova, popř. pro morfologické hranice (to pochopitelně neplatí pro diftong /au/). Naopak, kombinaci /čá/ s jejím méně než 10% zastoupením v CS je možné pokládat za charakteristický rys domácích slov. Je to ovšem důsledek jejího celkově malého výskytu (v CS pouze 8×, z nichž polovina jsou anglicismy, např. *charleston*).

Podíváme-li se blíže na kombinace s afrikátou /c/, zjistíme, že i ony mají omezený výskyt. V DS lze skupinu /ca/ nalézt hlavně v expresivních slovech (např. *capy, cavyky*) a v číslovkách jako *dvacaterý, třicaterý*. Další doklady představují slovo *zrcadlo*, kde je ovšem historicky morfologická hranice, a slovo *ocas* bez jasné etymologie. V CS se skupina /ca/ také objevuje (*car, candát*) a počet jejích dokladů je podobný jako u DS. Obdobné distribuční omezení se týká i skupiny /cá/. Kromě morfologické hranice, číslovek jako *dvacátý, třicátý* a zdobnělin *ocásek* a *zrcátko* (viz výše) je doložena v expresivech (např. *cákat*) a v cizích slovech (*lipicán, palcát*). Dále se vyskytuje ve dvou domácích výrazech s nejasnou etymologií, *cár* a *kocábka*. První by mohla být původně výpůjčka (Rejzek, 2015) nebo slovo expresivního původu (Holub – Kopečný, 1952). Expresivní původ se uvažuje i u druhého slova, *kocábka*. Stručně řečeno, kombinace /c/ a středního vokálu se vyskytují na morfologické hranici, v expresivních a cizích slovech a dále ve tvarech suffixu *-cat-* (*-caterý, -cátý*), v kmenech domácích slov *ocas* a *zrcadlo* a jejich zdobnělinách *ocásek* a *zrcátko*. Doplňme, že kombinace /cau/ není doložena v žádném typu lexika (/čau/ najdeme v CS pouze jednou v pozdravu *čau*).

Zastoupení kombinací středních vokálů s ostatními konsonanty zobrazuje tabulka 6. Necháme-li stranou fonémy /f/, /g/, /au/, nemá žádná kombinace ani v jednom typu lexika převažující zastoupení. Je naopak zřejmé, že se kombinace s labiálními, velárními, alveolárními konsonanty, /r/ a /l/ vyznačují pozoruhodně vyrovnaným rozložením mezi domácí a cizí slova (v tabulce označeno ≈).¹³ Projevuje se to více u krátkého /a/ než u dlouhého /á/. U /a/ jsou výjimkou skupiny /ma/ a /ta/, jež se častěji objevují v CS, ale z fonologického hlediska tento fakt nelze pokládat za nějakou pravidelnost. Je zajímavé, že v případě /á/ se všechny kombinace, které nemají vyrovnané zastoupení, vyskytují větší měrou v DS, tj. /pá/, /vá/, /xá/, /há/, /zá/ (zastoupení v CS < 33 %). Vezmeme-li v úvahu, že v CS má menší než 33% zastoupení i kombinace /žá/ (viz tab. 5), můžeme zde vidět dvě tendence, které se částečně překrývají. Za prvé, v cizích slovech je menší

¹³ U /va/ je však potřeba upozornit, že mnoho cizích slov s touto kombinací obsahuje domácí verbální sufix *-ovat* (např. *skenovat*). V kmenech cizích má tato kombinace také mnoho dokladů (např. *kvantum*).

výskyt spojení znělé frikativy (/v/, /z/, /ž/, /h/) s vokálem /á/.¹⁴ Za druhé, stejně tak je s tímto vokálem omezenější výskyt kombinací s velárními frikativami (/x/ a /h/).

Tabulka 6: Procentuální zastoupení kombinací labiál, velár, /r/ a alveolár se středními vokály v CS (pro symboly viz tab. 2)

	/p/	/b/	/m/	/f/	/v/
/a/	≈	≈	> 66	> 90	≈
/á/	< 33	≈	≈	100	≈
/au/	100	100	100	100	100

	/k/	/g/	/x/	/h/	/r/
≈	100	≈	≈	≈	
≈	100	< 33	< 33	≈	
100	100	×	100	100	

	/t/	/d/	/n/	/s/	/z/	/l/
/a/	> 66	≈	≈	≈	≈	≈
/á/	≈	≈	≈	≈	< 33	≈
/au/	100	100	100	100	100	100

7. Kombinace s předními středovými vokály

Dalším typem distribučně specifických CV kombinací je podle Mathesia (1931) spojení veláry a předního středového vokálu /e/. Zatímco první ze svých pravidel Mathesius zjevně vztahoval jak na krátké, tak dlouhé vokály a diftongy, zde není zřejmé, zda pod druhou skupinu kombinací řadil i spojení s dlouhým /é/, popř. i s diftongem /eu/. Zastoupení kombinací labiál, velár a /r/ se všemi e-ovými vokály ve zkoumaném materiálu zobrazuje tabulka 7; tabulka 8 pak udává hodnoty pro zbývající konsonanty. Diftongu /eu/ nevěnuji bližší pozornost, poněvadž jsou všechny jeho kombinace, jichž navíc není mnoho, doloženy výhradně v CS.

Tabulka 7: Procentuální zastoupení kombinací labiál, velár a /r/ s předními středovými vokály a /eu/ v CS (pro symboly viz tab. 2)

	/p/	/b/	/m/	/f/	/v/
/e/	> 66	≈	> 66	100	≈
/é/	≈	> 66	> 66	100	≈
/eu/	×	×	×	100	×

	/k/	/g/	/x/	/h/	/r/
> 66	100	> 66	> 66	> 66	
> 66	100	100	> 66	> 90	
×	100	×	100	100	

Spojení veláry s /e/ se skutečně objevuje na morfologické hranici a dále v expresivních a cizích slovech, srov. expresivní *kecat*, *chechtat se*, *hepčí* a cizí *raketa*, *chemie*, *herbář*. Jsou doložena i v kmenech domácích intelektuálních slov, jako např. *keř*, *věchet*, *hebký*, což poznamenává i *Mluvnice češtiny I* (s. 159). Opět tedy existují výjimky, ale na rozdíl od výjimek u prvního Mathesiova pravidla zde nelze tvrdit, že by taková slova měla nejasnou etymologii nebo by byla starými výpůjčkami či původními expresivy. Spojuje je ovšem jeden fakt, a to, že je v nich vokál /e/ výsledkem vokalizace jerů, kterážto

¹⁴ V tabulce 6 je sice u kombinace /vá/ uvedeno vyrovnané zastoupení (označeno symbolem ≈), avšak přesné zastoupení je 33,41 %.

změna proběhla až po palatalizaci velár (právě palatalizace velár je hlavní důvod, proč jsou spojení veláry a /e/ v domácích slovech tak málo častá – původní veláry se před předními vokály změnily v (post)alveoláry). I přesto materiál potvrzuje, že kombinace velár s /e/ jsou doloženy hlavně v CS, byť nepřesahují 90% hranici. Jejich zastoupení je 86 % pro /xe/, 85 % pro /ke/ a 82 % pro /he/.¹⁵

Již citovaná *Mluvnice češtiny I* (s. 159–160) rovněž upozorňuje na to, že uvedené kombinace nejsou jediné distribučně specifická spojení konsonantu s vokálem /e/. Patří sem i kombinace /re/, která se kromě morfologické hranice vyskytuje v expresivních slovech (srov. *breptat*, *šeredný*) a ve slovech cizího původu (např. *regál*), popř. i ve vlastních jménech (např. *Renata*). Objevuje se také v několika domácích slovech intelektuálních (např. *ret*, *rež*), v nichž je /e/ opět jerového původu. Analýza materiálů toto pozorování potvrzuje. Nejenže má /re/ v CS 89% zastoupení (např. *adresa*, *kompresor*), ale celkově jsou spojení vibranty /r/ a e-ových vokálů typická pro cizí slova (viz níže).

Zastoupení ostatních kombinací konsonantu s e-ovými vokály je shrnuto v tabulce 8 (alveoláry a palatály) a tabulce 7 (labiály). V případě /e/ nemá žádná z nich v CS převažující zastoupení. Naopak, kombinace /t/, /d/ a /ř/ jsou, ovšem podle očekávání, typické pro domácí slova. Vůbec se spojení palatál a /e/ vyskytují více v DS, byť výjimkou je skupina /je/, ke které se vrátím v oddíle 9, a /ce/, která je doložena spíše v cizích slovech.¹⁶ U kombinací s labiálami a alveolárami nelze spatřovat nějaké obecnější preference pro určitý typ lexika. Naopak za pozornost stojí, že u alveolár je zastoupení kombinací s vokálem /e/ téměř vyrovnané.

Tabulka 8: Procentuální zastoupení kombinací alveolár a palatál s předními středovými vokály a /eu/ v CS (pro symboly viz tab. 2)

	/t/	/d/	/n/	/s/	/z/	/l/
/e/	≈	> 66	≈	≈	≈	≈
/é/	> 66	> 66	> 66	> 66	100	< 33
/eu/	×	100	100	×	100	100

	/t'/	/d'/	/n'/	/s'/	/z'/	/ř/	/j/	/č/	/c/	/dž/	/dz/
/e/	< 10	< 10	< 33	< 33	< 33	< 10	≈	< 33	> 66	100	×
/é/	×	×	100	100	100	×	> 90	0	> 90	×	×
/eu/	×	×	×	×	×	×	×	×	×	×	×

Zcela jiná a daleko jednoznačnější je distribuce kombinací s dlouhým /é/, což ovšem nevyplývá ani tak z konkrétních kombinací jako z marginálního postavení tohoto vokálu.

¹⁵ To, že jsou v CS nejvíce doloženy kombinace s frikativou /x/, potvrzuje Mathesiovu (1931) poznámku, že právě u /x/ se nejvíce projevuje jeho velární (tj. výjimečná) povaha.

¹⁶ U /ře/ navíc většina výskytů v CS odpovídá domácím prefixům *pře-* a *před-* (např. *přeinstalovat*, *předalpský*). V kmeni cizích slov tuto kombinaci najdeme např. v *řehole*. Stejně tak domácím afixům náleží většina výskytů kombinace /ře/, a to především adverbialnímu sufixu *-ně* (např. *totálně*). V přejatém kmeni je /ře/ doložena např. v *kastaněta*.

Podle Vachka (1968, s. 32–33) se v češtině výskyt /é/ omezuje na (i) sufixy (např. *dobré, dobrému*), (ii) kmeny knižních domácích slov, které mají obecněčeské či hovorové tvary s /í/ (např. *polévka/polívka, mléko/mlíko*), (iii) kmeny knižních či dokonce archaických domácích slov, v nichž takové tvary v hovorovém jazyce doloženy nejsou (bez příkladů, ale např. *bérec, česka* nebo *céva*), (iv) kmeny knižních/archaických domácích slov, v nichž se /é/ střídá s /e/ (např. *péro/pero*, ale také *plémě/plemeno, témě/temeno*),¹⁷ (v) expresivní slova (*jémine, ně*), (vi) cizí slova (např. *téma, pléd*) a konečně (vii) slovo *lék* a jeho deriváty jako *lékař, lékárna*, kde náhrada /í/, popř. /e/ není možná (v nářečí se ovšem objevuje i tvar *lík* místo *lék*).

Analýzovaný materiál Vachkovo pozorování potvrzuje. Mimo uvedené případy /é/ dále najdeme v jednom typu slov, která lze hodnotit jako expresivní, a to v názvech písmen jako *bé, cé, dé* a ve slovech od nich odvozených jako *béčko* nebo *déčko*. Sem patří, byť ve vzorku nedoložené, i kombinace /t'é/, /d'é/ – srov. názvy písmen *t'* a *d'*. To je pozoruhodné, jelikož se spojení palatálních okluzív a vokálu /é/ jinak v češtině neobjevují (Bičan, 2013).

V CS jsou všechny kombinace kromě /p'é/, /v'é/, /l'é/ a /č'é/ doloženy nejméně 66 %. Skupiny /p'é/ a /v'é/ mají srovnatelné zastoupení jak v domácích, tak cizích slovech. U obou je jejich výskyt v DS značně ovlivněn velkým počtem derivátů. U /p'é/ jsou to deriváty slov *péct, péro* a u /v'é/ deriváty slov *vést, vézt, kvést, vévoda*. Kombinace /v'é/ spolu s /h'é/ a /s'é/ zároveň představují jediné kombinace /é/ s frikativou, které jsou v DS doloženy. Výskyt /s'é/ je omezen na kořen *-sév-* (např. *rozsévat*) a etymologicky příbuzné slovo *sémě*, zatímco /h'é/ se neobjevuje jinde než na morfologické hranici (*mnohé*). Další výskyty kombinací /v'é/, /s'é/ a /h'é/ a všechny výskyty s ostatními frikativami odpovídají cizím slovům (např. /v'é/ *kvér*, /s'é/ *sérum*, /h'é/ *hélium*, /f'é/ *aféra*, /z'é/ *arzen*, /š'é/ *šéf*, /ž'é/ *ležérní*, /x'é/ *schéma*).

Pouze v CS lze nalézt kombinaci /g'é/ a dále /ň'é/. Druhá ze zmíněných se ovšem objevuje ve třech slovech, z nichž všechna obsahují element *monseigneur*. Za zvláštnost cizích slov lze pokládat i kombinace /j'é/, /c'é/ a /r'é/, které mají v CS zastoupení větší než 90 %. Situace u /c'é/ je ovlivněna malým počtem domácích slov s touto kombinací – 4 (např. *céva*) oproti 48 v CS (např. *scéna*). Jediným dokladem kombinace s druhou afrikátou, tj. /č'é/, je domácí *česka*. K /j'é/ viz oddíl 9.

Jak již bylo naznačeno výše, v DS se objevuje kombinace /l'é/, která je navíc v celém zkoumaném materiálu nejvíce doloženým spojením konsonantu s /é/. V CS má tato kombinace zastoupení menší než 33 % a jako jedinou ji lze pokládat za fonologický prvek, kterým se vyznačují spíše domácí než cizí slova. Její výjimečnost vynikne především v konfrontaci s druhou likvidou, /r/, jelikož /r'é/ má v CS 98% zastoupení a je tedy pro tento typ lexika typická. V DS se objevuje pětkrát. Dva výskyty spadají na rozhraní kmene a sufixu (*modré, dobré*). Uvnitř kmene ji najdeme ve slově *réva* a jeho dvou derivátech. Přítomnost /r'é/ v kmenech je tedy něco, co má *réva* společného s cizími slovy

¹⁷ Mimo Vachkův seznam stojí i *jméno*; u něj však, aspoň v nespisovné češtině, také dochází ke střídání s /e/, srov. *meno*.

(srov. *rébus*, *réma*), a skutečně se ukazuje, že je to původně slovo cizího původu, které nebylo zařazeno do zkoumaných slovníků cizích slov (Machek, 1968, s. 513; Rejzek, 2015, s. 593). Spojíme-li tento fakt se specifickým výskytem skupiny /re/ zmíněným výše, můžeme konstatovat, že kombinace vibranty /r/ a e-ových vokálů jsou charakteristickým rysem cizího lexika.

8. Kombinace s předními vysokými vokály

Poslední skupinou českých vokálů jsou přední vysoké vokály /i/ a /i/. Zastoupení jejich kombinací je shrnuto v tabulce 9. Žádná z nich není přímo postižena Mathesiovým pravidlem či jeho navrhovanými rozšířeními. Z toho lze usuzovat, že nejsou distribučně omezeny na kontexty (1). Naopak, jak poznamenává Rechziegllová (Rechziegel, 1998, s. 334), tyto kombinace nevykazují žádná omezení. Materiál přesto ukazuje, že zde existují preference některých kombinací jak pro domácí, tak cizí lexikum.

Tabulka 9: Procentuální zastoupení CV kombinací s předními vysokými vokály v CS (pro symboly viz tab. 2)

	/t'/	/d'/	/ň/	/š/	/ž/	/ř/	/j/	/č/	/c/	/dž/	/dz/
/i/	< 33	< 10	< 10	< 33	< 10	< 10	≈	< 33	> 66	100	×
/i/	< 10	< 10	≈	< 33	< 33	< 10	< 33	< 33	< 33	100	100

	/p/	/b/	/m/	/f/	/v/
/i/	> 66	> 66	> 66	> 90	≈
/i/	< 33	< 33	< 33	100	≈

	/k/	/g/	/x/	/h/	/r/
≈	100	> 66	> 66	> 90	
> 66	100	≈	< 10	≈	

	/t/	/d/	/n/	/s/	/z/	/l/
/i/	> 90	> 90	> 90	> 66	> 66	> 66
/i/	≈	≈	< 33	< 33	≈	< 33

Z tabulky 9 je zjevný kontrast mezi kombinacemi /i/ s alveolárami /t/, /d/, /n/ včetně /r/ a kombinacemi /i/ s palatálami /t'/, /d'/, /ň/, /ř/ (k tomu viz též odd. 10). Zatímco první skupina má v CS větší než 90% zastoupení, a lze ji proto pokládat za typickou pro cizí slova, druhá skupina má naopak menší než 10% zastoupení s výjimkou /t'i/, která má zastoupení 15 %. Kombinace palatál s /i/ tedy naopak charakterizují slova domácí. Vyšší zastoupení skupiny /t'i/ je způsobeno velkým počtem názvů jazyků zakončených na domácí sufix *-tina* (např. *vlámština*). V kmenech cizích slov je kombinace /t'i/ vzácná (např. *pobratim*).

Z kombinací /ti/, /di/, /ni/ a /ri/ je pro cizí slova nejcharakterističtější /ni/. Necháme-li stranou sufixy *-iny* a *-y* (např. *prázdniny*, *obranyschopnost*), pak se /ni/ v kmeni domácích slov objevuje pouze ve třech slovech a jejich derivátech, *nyní*, *nyvý* a *unylý*. Všechny jsou to výrazy knižní až archaické (etymologicky jsou *nyvý* a *unylý* slova příbuzná). Zbytek výskytů se nachází v CS (např. *nikl*). O něco více jsou v kmeni domácích slov doloženy kombinace /ti/ a /di/ (např. *ostych*, *dychtivý*), avšak jejich zastoupení v DS je ovlivněno velkým počtem slov obsahujících prvky *čtyř-* nebo *kdy* (např. *nikdy*). Jak /ti/

tak /di/ se vyskytují na morfolo­gické hranici, ale také v místě, jež původně morfolo­gickou hranicí bylo (*motyka, vladyka*).

Zajímavé je i postavení kombinace /ri/. Kromě morfolo­gické hranice se v DS objevuje v expresivních slovech (např. *ryčet, kykyryký*) a dále v několika dalších, intelektuálních slovech. Etymologicky tato slova mohou být původu expresivního (např. *hryzat* nebo *prýč*, viz Holub – Kopečný, 1952, s. v. *hryžti* a *prýč*) nebo byla přejata z jiných jazyků (*brynda, tovaryš, tryzna, čtverylka, pokrytec, rytíř, rys* – znak, nikoliv zvíře), popř. nemají jasnou etymologii a nelze u nich vyloučit přejetí (*ryba*, Holub – Kopečný, 1952). Oproti tomu se kombinace s /l/ sice vyskytují větší měrou v CS, ale nepřesahují 90% hranici.

Rozdíl mezi alveolárami a palátalami se projevuje i u frikativ, byť ne tak silně. Kombinace /si/ a /zi/ mají v CS většinové zastoupení (obě větší než 66 %), kdežto /ši/ a /ži/ se většinově vyskytují v DS. Skupina /ži/ je dokonce v CS doložena méně než 10 % a je tedy zvláštností domácí slovní zásoby. Podobný rozdíl je mezi /či/ (< 33 %) a /ci/ (> 66 %).

Kombinace /ši/ si zaslouží pozornost. Být je z více než 66 % zastoupena v DS, má v tomto typu lexika omezený výskyt. Přičítat to můžeme především její přítomnosti na morfolo­gické hranici. Jinak je skupina /ši/ doložena (i) ve výrazech, které jsou sice v analyzovaném materiálu zařazeny mezi slova domácí, ale které jsou původními vý­půjkami (*šidit, šikmo, šilhat*), (ii) ve slovech, u nichž se o výpůjce uvažuje (*široký* – viz *Etymologický slovník jazyka staroslověnského* 16, s. 942), (iii) ve slovech, která mají nejspíš expresivní původ (*šimrat, šišlat*), popř. se o expresivním původu či způsobu tvoření uvažuje (*šibat* a od něho odvozená *šibenice* – viz Machek, 1968; *šinout* – viz Holub – Kopečný, 1952; *šiška* – viz Rejzek, 2015), a (iv) ve výrazu *mšice*, v němž /ši/ stojí na původní morfolo­gické hranici, a ve výrazu *všimnout* bez jasné etymologie. Je zjevné, že tato kombinace není pro původní lexikum češtiny typická.¹⁸

Větší než 66% zastoupení v CS má i většina zbývajících kombinací konsonantů s krátkým /i/, a to kombinace s labiálami a velárami. Výjimkami jsou /vi/ a /ki/ (obě mají srovnatelné zastoupení v obou typech lexika). Z kombinací velár stojí za zmínku spoje­ní s frikativou /h/. Skupina /hi/ má se svým 87% zastoupením v CS nejbliže k tomu být pokládána za typickou pro cizí slova. Její výskyt je celkově malý (spolu s /xi/ je nejméně doloženou kombinací konsonantu s /i/). Mimo morfolo­gickou hranici je kombinace /hi/ v DS omezena na několik kořenů, jako *hyb-* (*pohyb*), *hy-* (*hynout*), *hyzd-* (*hyzdit*), a dále na expresivní slova (*čehy, hihňat, hije*), popř. slova s nejasným původem (*mohyla*).

Konečně pozornost zaměřme na kombinace s dlouhým /í/. Z tabulky 9 vyplývá něko­lik faktů. V první řadě mají všechny kombinace s /í/ buď srovnatelné zastoupení v obou typech lexika, nebo jsou častější v domácích slovech. Neplatí to pro spojení s cizími foné­my /f/, /g/, /dz/ a /dž/ a s konsonantem /k/, kteréžto kombinace mají větší zastoupení v CS.

¹⁸ Podobná situace je i u kombinace /ši/. Kromě cizích slov (*kšír*) a mimo morfolo­gickou hranici se vysky­tuje v domácích slovech *šít, šídlo, šije*, která jsou etymologicky příbuzná, *všimat* (viz výše *všimnout*) a dále ve slovech *šíp* s nejasnou etymologií a *šilet*, jež nejspíše souvisí s *šálit*, které je také bez jasné etymologie.

U /kí/ je zvýšený výskyt zapříčiněn přítomností této kombinace v domácích sufixech *-(c)ký, -ský* nebo *-ovský* (např. *prozodický, abiturientský, dandyovský*). Výskyt v kmeni je omezen na velmi málo cizích slov (např. *pedikýra*). Preference domácích slov se nejvíce projevuje u spojení labiál s /í/. Skupina /ví/ má sice vyrovnané zastoupení v DS a CS, avšak většina jejích výskytů připadá v cizích slovech na domácí sufixy, především na *-ový* a *-ství* (např. *sekundový, regentství*). V kmeni cizích slov se /ví/ objevuje např. ve *vízum*.

Tabulka 8 rovněž naznačuje, že některé z kombinací s /í/ můžeme dokonce pokládat za specifikum domácích slov, protože mají v CS menší než 10% zastoupení. Kromě /ří/, /ďí/ a /ří/, u nichž takový výskyt nepřekvapuje, jsou to skupiny /hí/ a /čí/. Výskyt první je omezen na 8 cizích slov (např. *sipahí, healer*), zatímco druhá se kromě morfologické hranice (např. *tresčí*) a sufixů (např. *hitlerčik*) objevuje ve výrazu *cappuccino*, ve slovech s prvkem *čin-*, jako např. *čínský*, a v anglicismech jako *cheeseburger*.

Pro vokál /í/ dále platí i to, co bylo řečeno o spojení palatál, /ř/ a /c/ s krátkým vokálem /i/: tyto kombinace jsou z historických důvodů buď typické pro domácí slova (/ří/, /ďí/, /ří/), nebo v nich mají větší zastoupení než ve slovech cizích (/ši/, /ži/, /jí/, /čí/). Jedinou výjimkou je kombinace /ňí/, která má srovnatelný počet dokladů v obou typech lexika. To je způsobeno cizími slovy zakončenými na domácí sufixy *-ní, -ník* a *-níkový* (např. *renesanční, agrárník, bavlníkový*). V kmeni cizích slov tuto skupinu nenajdeme.

Konečně tabulka 9 dokládá, že zatímco kombinace krátkého /i/ s jinými než palatálními konsonanty jsou v CS vždy zastoupeny nejméně 66 %, skupiny s dlouhým /í/ mají naopak zjevnou tendenci převažovat v DS. Kromě již zmíněného /kí/ a kombinací /fí/ a /gí/ mají kombinace nepalatály s /í/ buď srovnatelné zastoupení v obou typech lexika (/ví/, /xí/, /rí/, /tí/, /dí/, /zí/), nebo převažují v DS s menším než 33% zastoupením v CS (/pí/, /bí/, /mí/, /hí/, /ní/, /sí/, /lí/).

9. Kombinace s aproximantou /j/

Přestože aproximanta /j/ nepatří mezi fonémy charakterizující určitý typ lexika, vyznačuje se distribucí, která ji spojuje s konsonanty typickými pro cizí slova. Mathesius ji proto zahrnul do svého pravidla pro zadní vokály a stejně tak ji zařadila Rechieglová do svého rozšířeného pravidla pro vokál /a/. Podle těchto pravidel se příslušné kombinace s /j/ vyskytují především nebo výhradně v expresivních slovech, na morfologické hranici či v cizích slovech. Materiál toto potvrzuje pro kombinace se zadními vokály, byť výjimkou je domácí *junák* s variantou *jonák*. Slovo je nicméně původní výpůjčkou. Na druhé straně se ale rozšíření Mathesiova pravidla o /a/ nezdá být zcela opodstatněné. Skupina /ja/ se totiž objevuje v mnoha kmenech domácích intelektuálních slov, jako např. *jaký, jazyk, nejapný*.

Distribučně specifické nejsou jen kombinace se zadními vokály, nýbrž všechna spojení /j/ s vokály (viz tab. 3, 5, 8 a 9). V prvé řadě pouze cizí slova obsahují kombinace s diftongy, a to včetně /ou/. Ovšem ani v nich nemají kombinace s diftongy mnoho dokladů a objevují se jen v anglicismech. Skupina /jou/ se vyskytuje v *yeoman* (tj. /joumen/) a *talk radio* (tj. /rejdijou/) a skupina /jau/ v *layout* and *playout*. Kombinace /jeu/ není v databázi doložena.

Všechny kombinace /j/ s dlouhými vokály mají větší než 90% zastoupení kromě /ji/ (např. /jó/ *jód*, /jú/ *interview*, /já/ *plejáda*, /jé/ *jégrovky*, /ji/ *vějír*). V DS se /jó/ objevuje ve dvou citoslovcích, *jó*, *vijó*, a /jé/ v jejich variantách, *jé*, *vijé*. Skupina /jú/ se také nachází v jednom citoslovcí, *jú*, a dále na morfologické hranici, a to ve slově *bojůvka*. Jedinou kombinací, s výjimkou /ji/, přítomnou v kmeni domácích intelektuálních slov je /já/ (např. *já*, *jádro*).

Kombinace /ji/ má výjimečné postavení a lze ji jako jedinou pokládat za rys domácích slov (srov. např. *jídlo*, *jiška*). V CS má zastoupení 25 %. Její výskyt z velké části připadá na zakončení *-ain*, *-ein*, *-oin*, *-eida*, *-oida* (např. *betain*, *cystein*, *heroin*, *sklereida*, *planetoida*). Jejich výslovnost není ve výslovnostních a výkladových slovnících češtiny ustálená a např. u *heroin* kolísá mezi [ħerojn], [ħerojm] a [ħeroji:n] (Štěpánová, 2013, s. 124). Pro Fonologický korpus češtiny byla zvolena výslovnost uvedená ve VSC a SSJČ, tj. se skupinou [ji:].¹⁹ Pokud bychom slova tohoto typu vynechali nebo zvolili jinou výslovnost, bude se /ji/ vyskytovat u velmi malého počtu cizích slov (např. *vějír*, *yield*).

Právě uvedená zobecnění se nevztahují na zbývající kombinace s /j/, kam patří kombinace s krátkými nezadními vokály /ja/, /je/, /ji/. Byť mají všechny o něco větší procentuální zastoupení v CS, u žádných z nich nepřesahuje 90 % (/ja/ 75 %, /ji/ 58 %, /je/ 52 %). Přesto ale existuje zásadní rozdíl mezi jejich výskytem v domácích a cizích slovech. Rozlišíme-li tři základní kontexty, v nichž se spojení konsonantu a vokálu může objevit (po vokálu, po konsonantu a na začátku slova), pak se poměr mezi výskytem v jednotlivých situacích značně liší. V CS připadá 92 % výskytů kombinací /ja/, /je/, /ji/ na postvokální kontext (např. *bojar*, *kóje*, *hokejista*), zatímco v DS této situaci odpovídá 22 %. Naopak v DS většina výskytů (61 %) připadá postkonsonantickému kontextu, jemuž v CS přísluší pouze 4 % případů (např. *soljanka*, *adjektivum*, pro /ji/ nedoloženo); zbytek odpovídá slovnímu začátku (16 % v DS, 3 % v CS, např. *jachta*, *jeseter*, *jinan*). Vysvětlení této diskrepance lze hledat v historii češtiny. V domácích slovech je v postkonsonantickém kontextu výskyt /j/ ve spojení s vokály /a/, /i/, /e/ zpravidla výsledkem fonemizace původního palatálního příznaku konsonantů (např. stč. *p'ěna* > mod. č. *pěna* [pjena]). V intervokálním postavení mělo /j/ v průběhu vývoje češtiny naopak tendenci zanikat (srov. psl. *dobroje* > mod. č. *dobré*). Nic z toho se netýká cizích slov. U nich navíc velké množství výskytů /j/ v intervokálním postavení odpovídá tomu, co se tradičně označuje jako hiát (srov. /ja/ *diabetik*, /je/ *chemie*, /ji/ *naivní*).

10. Faktory ovlivňující zastoupení a frekvenci CV kombinací

Již několikrát bylo zmíněno, že distribuční specifičnost některých CV kombinací lze odvodit od specifičnosti konkrétních konsonantů a vokálů. Neplatí to pro všechny takové kombinace, a proto je nutné vysvětlení hledat jinde. Svou roli mohou hrát jak interní faktory týkající se samotného jazyka, jeho vývoje a struktury, tak externí faktory vyplývající z kontaktu jazyka s jazyky jinými, popř. z geografických, historických a kulturních

¹⁹ ASCS žádnou explicitní výslovnost pro taková slova neudává.

souvislostí. Zastoupení a frekvenci CV kombinací snad také podmiňují faktory, které bychom mohli nazvat univerzální a které vyplývají z povahy lidského jazyka. Z artikulačních nebo percepčních důvodů totiž mohou být preferována určitá spojení konsonantu s vokálem.

K interním a externím faktorům patří (a) historické hláskové změny ve vývoji domácích slov a (b) hláskové změny při adaptaci slov z cizího jazyka. Interní změny mohly zapříčinit to, že některé kombinace bez náhrady zanikly. Hláskové změny však pravděpodobně častěji vyústily v kombinace zcela nové či již existující, což v druhém případě zvýšilo zastoupení stávajících kombinací. Jelikož jsou hláskové změny vázané na kontext, nově vzniklé kombinace mohly být vzácné nebo se vůbec nemusely vyskytovat v kontextech, v nichž ke změně nedošlo. Podobné důsledky mívají i externí hláskové změny při adaptaci výpůjček.

Frekvenci a zastoupení fonologických prvků ovlivňuje i (c) frekvence fonémů jak v současném lexiku, tak v dřívějších fázích vývoje jazyka, ale též (d) frekvence fonémů a jejich kombinace v jazycích, z nichž si daný jazyk slova vypůjčil. Je zřejmé, že četnost jednotlivých fonémů má vliv na kombinace, které tyto fonémy tvoří. Je-li foném málo častý, budou vzácné i jeho kombinace. Týká se to i fonémů v dřívějších podobách jazyka. Bez ohledu na to, zda zůstaly v dalším vývoji zachovány nebo se transformovaly na fonémy jiné (popř. zanikly), jejich původní frekvence, díky kontinuálnosti jazykového vývoje, ovlivňuje frekvenci fonémů a jejich kombinací v dalších vývojových stádiích. Stejně tak frekvenci CV kombinací ovlivňuje frekvence fonémů a jejich kombinací v jazyce, jenž je zdrojem lexikálních výpůjček. Tento vliv je ovšem nepřímý, poněvadž primární roli hraje výběr slov, která jsou z cizího jazyka přejata. Např. v anglickém lexiku je foném /θ/ zhruba dvakrát častější než /ð/ (Duanmu, 2008, s. 184), ovšem podíváme-li se do Fonologické databáze českých anglicismů, pak je v češtině 6 anglicismů (např. *rhythm and blues*), jejichž zdrojová slova obsahují /ð/, zatímco anglicismů se zdrojovým /θ/ je 34, tj. nikoliv dvakrát, ale téměř šestkrát více (např. *think-tank*). Je to způsobeno nepochybně tím, že foném /ð/ je v angličtině vzácný v substantivech, která patří mezi nejčastěji přejímané výrazy.

Distribuci a četnost fonologických prvků ovlivňují i nefonologické faktory, a to jak (e) morfologické a slovtvorné procesy vedoucí k většímu počtu určitých typů domácích slov nebo morfémů (např. verbální sufix *-ovat* nutně zvyšuje zastoupení kombinace /va/), tak (f) preference výpůjček určitého morfologického nebo slovtvorného typu či obecně výpůjček určité struktury (např. v odd. 5 zmíněné velké zastoupení kombinace /mu/ díky slovům zakončeným na *-ismus*; viz též pozn. 21). S tímto souvisí další fakt, na nějž jsme v průběhu výkladu několikrát upozornili. Jelikož jsou slova z cizích jazyků při přejetí do češtiny často morfologicky adaptována tak, že přijmou domácí afix (srov. angl. adjektivum *instant* → č. *instantní*), nebo mohou samy o sobě být předmětem další slovtvorné derivace (např. *instantní* → *instantně*),²⁰ výskyt určité CV kom-

²⁰ Bez konkrétních dokladů užití je v praxi obtížné tyto případy odlišit, protože např. *artový* může být stejně domácí derivát od cizího slova *art* jako adaptace anglického adjektiva *art* v sousloví *art film*.

binace může být v cizích slovech ovlivněn její přítomností v domácím afixu. V analyzovaném materiálu se tento vliv výrazněji projevuje u sedmi kombinací, a to u /ňí/, /ví/, /ře/, /kí/, /ňe/, /va/ a /tí/. U všech více než 50 % jejich výskytů v cizích slovech odpovídá domácím afixům.

Vraťme se k fonologickým faktorům. Jelikož by objasnění důvodů distribučních rozdílů u všech CV kombinací, u nichž byly takové rozdíly nalezeny, daleko přesáhlo rámec tohoto článku, demonstrují vliv uvedených faktorů na reprezentativním případu kombinací krátkých vokálů s palatálami /tʰ/, /dʰ/, /ň/, /ř/ (dále pro zjednodušení jen palatály) oproti kombinacím s alveolárami /t/, /d/, /n/, /r/ (dále jen alveoláry). Jejich distribuci lze shrnout následovně (viz tabulky výše): (a) kombinace palatál s vokály jsou v CS vzácné, (b) kombinace palatál s nepředními vokály jsou v DS ve srovnání s kombinacemi palatál a předních vokálů vzácné, (c) kombinace alveolár a předního vysokého vokálu mají většinové zastoupení v CS a (d) kombinace alveolár a nepředních vokálů mají vyrovnané zastoupení jak v CS, tak DS.²¹

Palatály /tʰ/, /dʰ/ odpovídají v postavení před předními vokály praslovanským palatalizovaným alveolárám, jež zde vznikly jako varianty alveolár (Trávníček, 1935; Komárek, 1962; Lamprecht – Šlosar – Bauer, 1986). Přední vokály zůstaly zachovány a zároveň se palatalizace fonologizovala díky jiným změnám, které umožnily nepalatalizovaným alveolárám stát před předními vokály.²² Palatály /ň/, /ř/ mají stejný původ jako /tʰ/, /dʰ/, ale zčásti jsou i střídnicemi skupin /nj/ a /rj/. V nich bylo /j/ nejspíše vždy součástí sufixu (Ślowski, 1974). V praslovanštině stál v takovém sufixu za /j/ buď přední vokál, který pozdějším vývojem svůj přední, tj. palatální charakter neztratil, nebo vokál /a/, který se následným vývojem v postavení po /j/ změnil na přední vokál /e/ (*kanja* > *káně*, tj. /káňe/). Uvedené hláskové změny způsobily, že se v současné češtině palatály /tʰ/, /dʰ/, /ň/, /ř/ vyskytují především před předními vokály.

Před jiné vokály se palatály mohly dostat dvojím způsobem. Za prvé velkou roli hrála analogie a paradigmatické vyrovnávání při flexi a derivaci mezi tvary jednoho slova či kmene. Díky tomu se palatály ocitly v pozici před sufixy začínajícími na jiný než přední vokál (např. *kánata* < *káně*). Spojení palatála + nepřední vokál pak ale stálo na morfolo- gické hranici. Druhý způsob nabízela citově zbarvená a zvukomalebná slova. Palatalizace totiž měla v praslovanštině (a nejspíš i jinde) silný expresivní charakter (Machek, 1930), a proto mnoho expresivních slov obsahuje nejen palatály před předními vokály, ale i před vokály jinými (*tápota*, *d'obat*, *ňufat*, *křupat*). Takových slov je, alespoň v češtině, výrazně méně než neexpresivních slov, a tudíž příliš neovlivňují distribuci palatál.

Dalším zdrojem palatál jsou přejímky z cizích jazyků. Palatály nebo artikulačně a percepčně blízké konsonanty, jako palatalizované alveoláry, mohly být přítomny ve

²¹ Zvýšený výskyt kombinací /to/, /tu/ a /ta/ je v CS zapříčiněn jednak preferencí cizích slov určité struktury (např. pro /ta/ slova zakončená na *-sta* a *-ita*, pro /to/ slova zakončená na *-tor* nebo *-tovat*), jednak malým počtem domácích slov s těmito kombinacemi, což platí především pro /tu/. Malý počet domácích slov je rovněž příčinou zvýšeného výskytu /de/ v CS. Konečně ke kombinaci /re/ viz oddíl 7.

²² Jednou ze změn je depalatalizace palatalizovaných alveolár před /e/, což vedlo k tomu, že výskyt kombinací /te/, /de/, /ne/, /re/ není v domácích slovech omezený jako u kombinací s /i/.

zdrojových jazycích a následně převzaty jako součást přejímaného slova (*mad'al* < maďarské *magyal* [-ja-]). Palatály také mohly vzniknout hláskovými změnami při adaptaci jiných segmentů nebo jejich kombinací (*oř* < středohornoněm. *ors*).²³ Nedovedu posoudit, nakolik byl druhý případ častý, nicméně první případ, tj. přímé přejetí palatál, byl významně omezen dvěma faktory. V první řadě jazyky, z nichž pochází nejvíce minulých i současných přejímek (latina, řečtina, němčina, francouzština, angličtina), palatály nemají nebo je v dřívějších fázích vývoje neměly. Pokud je jako francouzština mají, pak bývají málo frekventované (/ɲ/ je druhým nejméně frekventovaným fonémem ve francouzštině, viz Malécot, 1974). Dalším a souvisejícím faktorem je okolnost, že palatály jsou napříč jazyky velmi vzácné (Maddieson, 1984) a v případě /ʃ/ dokonce extrémně vzácné. I kdyby si tedy čeština vypůjčovala slova z libovolných jazyků bez ohledu na geografické, kulturní a historické okolnosti, je pravděpodobnost přejetí palatál malá.

Výše zmíněná palatalizace alveolár před předními vokály rovněž způsobila, že v současné češtině mají kombinace palatála + nepřední vokál v porovnání s kombinací palatála + přední vokál malý výskyt. Zapříčinila současně také to, že malý výskyt mají i kombinace alveolára + /i/. Více než 90 % dokladů /ti/, /di/, /ni/, /ri/ náleží cizím slovům. Zde do hry opět vstupují jak interní, tak externí faktory. V současné češtině představuje /i/ výsledek splynutí dvou původně odlišných vokálů, a to předního /i/ a zadního /y/. Před druhým z nich k palatalizaci alveolár nedošlo, a proto /t/, /d/, /n/, /r/ mohou nyní stát i před vokálem /i/. Jelikož jsou oba původní vokály stále ortograficky rozlišeny, lze v databázi porovnat počet domácích slov obsahujících skupiny *ty, dy, ny, ry* (tj. původní alveoláry + /y/) se slovy obsahujícími skupiny *ti, di, ni, ři* (= /ʃi/, /dʃi/, /ɲi/, /ʁi/, tj. původní alveoláry + /i/). Hodnoty jsou 463 oproti 4 491. Znamená to, že počet kombinací alveolára + /i/ odstraněných palatalizací byl jen částečně kompenzován kombinací alveolára + /y/.

Frekvence fonémů, resp. jejich kombinací ve starších stádiích češtiny nabízí část vysvětlení velkého zastoupení skupin /ti/, /di/, /ni/, /ri/ v cizích slovech. Neméně zásadní je vysoká četnost slov cizího původu obsahujících uvedené skupiny. To má několik příčin. Především na rozdíl od palatál jsou alveoláry, popř. dentály poměrně frekventovanými fonémy v jazycích, z nichž si čeština slova půjčovala nejčastěji. Časté jsou i napříč různými jazyky (Maddieson, 1984; Gordon, 2016). Jelikož obojí platí i pro přední vysoké vokály, zvyšuje to pravděpodobnost přejetí slov se skupinou dentála/alveolára + /i/. Navíc se kombinace tohoto typu v domácím lexiku vyskytují, a tak mohla být povaha jak konsonantu, tak vokálu při adaptaci výpůjček zachována. K tomu je nutné připočítat další důležitý fakt: Mezi výpůjčkami s kombinacemi /ti/, /di/, /ni/, /ri/ se opakují slova určité struktury. Třetinu slov obsahujících uvedené kombinace tvoří slova, v nichž za alveolárou následují formy sufixů *-ický/-icky, -ika, -izace/-isace, -ismus, -ista, -ie, -ita* a *-ivní*. Vliv na zvýšený výskyt /ti/, /di/, /ni/, /ri/ v cizích slovech je tedy i důsledkem preference určitých slov, resp. derivačních prostředků.

²³ Striktně vzato i v prvním případě jde o fonetickou adaptaci, protože ne vždy jsou hlásky cizího jazyka zcela totožné s domácími hláskami.

Vraťme se ještě k jednomu z faktorů ovlivňujících zastoupení a frekvenci fonologických prvků, a to k frekvenci fonémů (či spíše hlásek, protože fonémy jsou vázány na konkrétní jazyk) a jejich kombinací napříč jazyky. Explikační hodnota frekvence je značně omezena nedostatkem znalostí o většině jazyků světa. Nelze proto tvrdit, že by fonologické vlastnosti popsané pro několik jazyků byly univerzálně platné pro všechny jazyky. Přesto lze stav v češtině srovnat alespoň s dílčími mezijazykovými popisy.

Pro CV kombinace takové popisy a srovnání existují (Janson, 1986; Maddieson – Precoda, 1992; Maddieson, 1992; MacNeilage et al., 2000; Valle – Rossato – Rousset, 2009; Whalen et al., 2012). Analýzy srovnávají různé počty jazyků (5 až 30), jež se zčásti překrývají, a jsou založeny na odlišných typech materiálu (lexikon, tj. lemmata oproti textu, tj. slovním tvarům, psaný text oproti mluvenému textu). Ačkoliv výsledky a závěry jednotlivých studií nejsou stejné, většina z nich se shoduje na tom, že zkoumané jazyky preferují následující druhy spojení:

(2) **Preferované CV kombinace napříč jazyky:**

- (a) Bilabiály a střední vokály (především a-ové vokály)
- (b) Dentály/alveoláry (koronály) a přední vokály
- (c) Veláry (popř. dorzály) a zadní vokály, především zadní zaokrouhlené vokály

Preference se obvykle vysvětlují artikulační ekonomikou. U uvedených kombinací nemusí jazyk činit mnoho různých pohybů – u koronál je vpředu stejně jako u předních vokálů, u velár vzadu stejně jako u zadních vokálů a u bilabiál, u nichž se jazyk aktivně nezapojuje, není jazyk ani vpředu, ani vzadu jako u středních vokálů. Vysvětlení se hledá i v akustice. Například právě kombinace bilabiála a /a/ jsou akusticky (auditivně) zvlášť výrazné. Dodejme, že četné studie MacNeilage, Davisové a dalších (viz MacNeilage et al., 2000; Whalen et al., 2012, a odkazy tam) navíc dokazují, že jsou spojení (2) preferována i u dětí při žvatlání a tvorbě nových slov.

Tabulka 10: Procentuální hodnoty odchylky pro CV kombinace v domácích slovech

	Labiály	Alveoláry	Palatály	Veláry
Přední	-7,16	-10,30	41,49	-30,22
Střední (nízké)	5,43	4,33	-20,51	13,88
Zadní	1,73	5,97	-20,98	16,35

Tabulka 11: Procentuální hodnoty odchylky pro CV kombinace v cizích slovech

	Labiály	Alveoláry	Palatály	Veláry
Přední	-2,81	1,15	16,51	-11,15
Střední (nízké)	3,87	-1,30	-11,01	6,04
Zadní	-1,06	0,15	-5,50	5,12

Tabulky 10 a 11 zobrazují pro češtinu procentuální hodnoty odchylky pro jednotlivé kombinace konsonantů podle místa artikulace a vokálů podle horizontálního členění. Stejně jako u Maddiesona a Precodové (1992) hodnoty vyjadřují rozdíl mezi procentuálním podílem dané vokalické třídy v kombinaci s konkrétní konsonantickou třídou

a celkovým podílem této vokalické třídy ve všech CV kombinacích. Např. procentuální podíl všech předních vokálů bez ohledu na to, s čím se kombinují, je 48,05 %, zatímco podíl předních vokálů v kombinaci s labiálou je 40,89 %. Hodnota odchylky je -7,16 %, což znamená, že se přední vokály s labiálami spojují o 7,16 % méně, než bychom očekávali z celkového výskytu předních vokálů. Hodnoty odchylky vyjadřují rozdíl mezi pozorovanou a očekávanou frekvencí. Pozitivní hodnotu odchylky lze chápat jako míru preference kombinace a zápornou hodnotu jako míru dispreference kombinace. Nulová hodnota by znamenala, že je kombinace zastoupena mírou, jež odpovídá frekvenci dané konsonantické třídy a vokalické třídy.

Z tabulek můžeme vyčíst několik faktů. V první řadě vidíme, že jsou hodnoty odchylky větší v DS než v CS. V cizím lexiku jsou největší rozdíly u palatál a velár, avšak pouze u palatál jsou statisticky významné.²⁴ Celkově jsou v tomto typu lexika jednotlivé kombinace distribuovány téměř rovnoměrně podle typu vokálu a konsonantu. U domácích slov dosahují výrazné a statisticky významné odchylky palatály a veláry. Stejně tak jsou statisticky významné i rozdíly pro alveoláry, i když u nich vykazují menší hodnoty odchylky. Nejmenší rozdíly jsou u labiál a zároveň jsou statisticky nevýznamné. Palatály jsou zjevně preferovány ve spojení s předními vokály, kdežto veláry se spojením s těmito vokály vyhýbají. Vysvětlení souvisí s výše popsaným původem palatál a dále s praslovanskými palatalizacemi velár, které před předními vokály změnily veláry v jiné konsonanty.

Pro domácí lexikum je dále patrný rozdíl mezi předními vokály na jedné straně a nepředními vokály, tj. zadními a středními vokály, na straně druhé. Zatímco labiály, alveoláry a veláry vykazují preferenci pro nepřední vokály a dispreferenci pro vokály přední, palatály upřednostňují přední vokály a vyhýbají se nepředním vokálům. V tomto binárním rozdělení českých vokálů lze vidět spojitost s návrhem Rechzieglové (Grygarová-Rechzieglová, 1993; Rechziegel, 1998), podle něhož by české a-ové vokály měly být klasifikovány spíše jako vokály zadní. Jejím argumentem jsou artikulační vlastnosti vokálu /a/ (resp. jeho charakteristika pomocí jakobsonovských distinktivních rysů) a jeho kombinační schopnosti paralelní k zadním vokálům (viz odd. 6). Ačkoliv nepokládám za nutné interpretovat české a-ové vokály jako skutečně zadní, výsledky analýzy a argumenty Rechzieglové se shodují v tom, že pro české vokály je relevantní rozdělení na vokály přední a nepřední. Stejně členění navrhuje i Němec (1971) s odkazem na fakt, že se vokály /a/, /o/, /u/ vyskytují v produktivních slovotvorných prostředcích, zatímco /e/, /i/ spíše v prostředcích neproduktivních.

Konečně srovnáme hodnoty odchylek kombinací v domácích slovech s výše popsanými preferencemi napříč jazyky (viz (2) výše). Ve shodě s nimi české labiály nejvíce preferují střední vokály a veláry preferují zadní vokály. U alveolár ovšem shoda není.

²⁴ Pro testování statistické významnosti v distribuci určitých prvků se obvykle používá test chí-kvadrát. Ve velkých vzorcích prvků je ovšem nepoužitelný, protože i sebemenší rozdíly budou statisticky významné (Mačutek – Wimmer, 2013). Test jsem proto použil na 20 000 náhodně vybraných vzorcích o 500 domácích slovech a na 20 000 náhodných vzorcích o 500 cizích slovech. Za významné jsem pokládal rozdíly, které byly u více než 95 % vzorků statisticky významné.

České alveoláry preferují zadní vokály místo předních vokálů a naopak pro přední vokály vykazují dispreferenci. Přední vokály jsou upřednostňovány palatálami.

Tato neshoda nemusí být náhodná. Především dvě z výše citovaných srovnávacích studií (Maddieson – Precoda, 1992; Maddieson, 1992) žádnou preferenci pro dentály/alveoláry nepotvrdily, což může znamenat, že oproti labiálům a velárám jsou tyto konsonanty napříč jazyky daleko variabilnější. Za druhé, citované studie zkoumaly buď pouze dentály či alveoláry, nebo analyzované jazyky palatály neměly. Výsledky pro češtinu mohou znamenat to, že pokud jazyk mezi dentálami/alveolárami a palatálami rozlišuje, spojení s předními vokály budou preferovat právě palatály. Ty se totiž z artikulačního hlediska blíží předním vokálům více než dentály/alveoláry.

11. Závěry

Ačkoliv se zde analyzovaný materiál vyznačuje některými omezeními, jeho poměrně velký rozsah (70 533 slov jako typů) opravňuje k předpokladu, že závěry platí pro češtinu obecně. Nejpodstatnějším omezením je absence flektivních tvarů slov, avšak výzkum byl záměrně zaměřen na slovníková lemmata, která lze z laického a lingvistického pohledu zpravidla pokládat za základní tvar slova. Přesto bude nutné závěry ověřit na datech, která zahrnují i různé flektivní tvary slov.

V prvé řadě se potvrzuje, že kombinace konsonantu s vokálem nejsou v češtině rovnoměrně distribuované. Analýza podporuje Mathesiovo (1931) pozorování, že konsonanty a vokály vykazují jiné kombinační možnosti uvnitř morfémů a na rozhraní morfémů a dále ve slovech domácích neexpresivních (slovesch intelektuální povahy) a ve slovesch expresivních (slovesch citově zabarvených nebo zvukomalebných) a slovesch cizího původu. Mathesius spojoval rozhraní morfémů, expresivní a cizí slova do jedné kategorie, ale alespoň v případě CV kombinací lze cizí slova vyčlenit jako samostatný soubor odlišný od ostatních zmíněných typů. Některé kombinace, jako např. /tʰ/, /dʰ/, /ŋ/, /ʃ/ + /o/, /ou/, mají totiž v cizím lexiku minimální zastoupení a vyskytují se většinou nebo zcela v domácích slovesch (v expresivech a na morfémové hranici).

Z velké části se ovšem popsané distribuční rozdíly projevují spíše jako tendence než jako absolutní omezení. Znamená to, že se některé typy kombinací vyskytují převážně v určitých kontextech, ale nejsou na ně nutně omezeny. Platí to i pro Mathesiovo pravidlo týkající se výskytu kombinací palatál a zadních vokálů, o nichž se někdy v literatuře tvrdí, že je čeština dovoluje pouze v expresivních a cizích slovesch a na morfologické hranici (Ludvíková, 1968; Novotná – Karlíková, 2016). Ze synchronního hlediska má pravidlo mnoho výjimek, byť etymologický rozbor svědčí o tom, že z historického hlediska své opodstatnění má (Bičan, 2018). Nelze ho ovšem jednoznačně rozšířit o a-ové vokály, jak někteří lingvisté navrhuji, poněvadž synchronních výjimek je v tomto případě podstatně více. Přesto některá fakta naznačují, že české vokály jsou rozděleny mezi přední a nepřední dimenzi. Nepřední, tj. střední a zadní vokály vykazují jinou schopnost kombinovat se s konsonanty než vokály přední a jsou užity ve fonologických formách znaků odlišným způsobem než vokály přední.

Distribuční specifčnost CV kombinací často vyplývá z povahy fonémů, které kombinaci tvoří, nikoliv ze samotné kombinace. Některé fonémy jsou doloženy výhradně v cizích slovech, popř. v cizích a expresivních slovech (/g/, /ó/, /eu/, /au/, /dž/, /dz/), jiné fonémy se zase objevují téměř výhradně ve slovech domácích (/t/, /d/, /ř/, /ou/). V mnoha případech se potvrzuje, že i kombinace s těmito fonémy jsou typické pro daný typ lexika. Existují však kombinace, jejichž výskyt je i přes přítomnost typicky domácího fonému charakteristický pro slova cizího původu (/ňú/, /d'a/, /d'á/, /ňé/, /jou/). Na obecné úrovni tedy pro češtinu neplatí předpoklad (Maddieson – Precoda, 1992), že fonémy (segmenty) jsou primárními fonologickými prvky, od jejichž vlastností jsou zcela odvoditelné vlastnosti fonémových skupin. Naopak, skupiny fonémů se mohou vyznačovat výrazně odlišnou distribucí než jednotlivé fonémy v této skupině.

Kromě uvedených existují ještě další kombinace typické pro slova cizího původu. Jednu skupinu tvoří kombinace, které jsou v analyzovaném vzorku slov doloženy pouze v cizích slovech, byť některé z nich mimo analyzovaný materiál lze nalézt i na morfolo- gické hranici, především na švu mezi kmenem a flektivní koncovkou (/ňú/, /šú/, /žú/, /jou/, /xé/, /z'é/). Existují ale i takové kombinace, které se jinde než v cizích slovech nejspíš neobjevují, což potvrzuje Fonologický korpus češtiny (obsahující více než 460 000 unikátních slovních tvarů). Jsou to kombinace s vokálem /é/ po palatálách /ň/, /š/, /ž/ (tj. /ňé/, /š'é/, /ž'é/, srov. /monseñer/ *monseigneur*, /šér/ *share*, *ležérní*). Vokál /é/ má v domácím lexiku celkově marginální postavení, ale je poměrně častý v sufixech (srov. *dlouhé*, *drahému*). V nich se ale využívá pouze v tzv. tvrdých typech skloňování, což vylučuje jeho výskyt po „měkkých“ /ň/, /š/, /ž/.

Za specifikum cizích slov lze rovněž pokládat kombinace, jež jsou z více než 90 % doloženy v tomto typu lexika. Vyzvednout zde můžeme tři skupiny kombinací. Za první, spojení aproximanty /j/ s dlouhým vokálem a diftongem. Všechna se vyskytují hlavně v přejatém lexiku. Jedinou výjimkou je kombinace /ji/, která je nejen častější ve slovech domácích, ale navíc spolu s /ji/ patří v češtině do první třetiny nejfrekventovanějších kombinací. Toto zjištění je pozoruhodné v konfrontaci s tvrzením, že se jazyky vyhýbají kombinacím /j/ a předních vysokých vokálů (Janson, 1986; Maddieson – Precoda, 1992; Maddieson, 1992).

Za druhé, cizí výrazy charakterizují kombinace /j/ a krátkých vokálů. Pro zadní vokály to platí ve všech kontextech, zatímco kombinace /ja/, /je/, /ji/ jsou pro cizí slova typické pouze tam, kde před nimi stojí vokál, nikoliv konsonant či začátek slova. Tento fakt naznačuje, že distribuční specifčnost CV kombinací bude v budoucnu nutné posoudit v širším kontextu, nejen pouze izolovaně. Třetí skupinu výhradně cizích kombinací tvoří spojení alveolár /t/, /d/, /n/, /r/ a /i/. To vyvrací tezi (Rechziegel, 1998), že se kombinace nepostížené Mathesiovým pravidlem vyskytují zcela bez omezení nebo preferencí.

Uvedené poznatky potvrzují oprávněnost rozdělování lexika na domácí a cizí slova (Mathesius, 1932; Fries – Pike, 1949; Kučera, 1958; Itô – Mester, 1999; a další). Rozdíly ovšem existují i uvnitř cizích slov, což bylo možné posoudit na materiálu výpůjček z angličtiny získaných z Fonologické databáze českých anglicismů, která je součástí analyzovaného souboru cizích slov. Některé CV kombinace se ve zkoumaném vzorku cizího

lexika neobjevují jinde než v anglicismech (/bou/ *bowling*, /jou/ *yeoman*, /jaʊ/ *playout*) nebo v nich mají dvoutřetinové zastoupení (/ʃou/ *show*, /jú/ *interview*, /daʊ/ *download*, /foʊ/ *smartphone*, /hou/ *homeoffice*, /ňú/ *newsroom*, /bó/ *paintball*, /kou/ *dresscode*, /hú/ *hooligan*; podrobněji viz Bičan et al., 2020). Je pozoruhodné, ale jistě ne náhodné, že všechny zahrnují nepřední vokály a zároveň diftongy nebo dlouhé vokály, především /ou/ a /ú/.

Dalším důležitým závěrem vyplývajícím z předložené analýzy je i to, že v češtině existují nejen fonémy, ale i kombinace specifické pro domácí slova. V literatuře se, pokud je mi známo, mluví pouze o signálech či příznacích cizosti (Mathesius, 1932; Trnka, 1942), expresivnosti (Zima, 1961) nebo o hraničních signálech mezi jednotkami vyššího řádu jako morfémů či slov (Trubetzkoy, 1939). Jelikož některé CV kombinace mají 100% nebo více než 90% zastoupení v domácích slovech, je možné je pokládat za signály domáckosti. To se samozřejmě může přílivem nových výpůjček změnit, ale alespoň ve zde zkoumaném materiálu má takový status většina kombinací s diftongem /ou/ a s konsonantem /ř/ a dále kombinace /t/, /d/ + /o/, /e/, /i/, kombinace /d/ /ň/ /ž/ + /i/, kombinace /č/ + /ú/, /á/, /é/, kombinace /h/, /n/ + /i/ a kombinace /pú/. Všechny zahrnují palatální konsonanty a/nebo nekrátké vokály (dlouhé vokály a diftongy).

Poznatky získané předloženou analýzou nejen prohlubují naši znalost češtiny, ale mohou být využity i mimo fonologii. Jestliže rozbor velkého, avšak nutně omezeného vzorku slov ukazuje, že se CV kombinace vyskytují s odlišnou mírou pravděpodobnosti v různých typech lexika nebo v určitých morfologických kontextech, pak lze toto zjištění použít pro automatickou detekci typů slov nebo pro detekci morfologických hranic (viz Bičan et al., 2020, s. 120–127, kde byla tato metoda s pozitivními výsledky testována na vlastnostech typických pro anglicismy). Získané poznatky je také možné využít v historickém zkoumání češtiny, především v etymologii. Není neobvyklé, že se u některých slov etymologové neshodují na jejich původu a nabízejí různé výklady. Plauzibilitu výkladů je možné posoudit i na základě synchronní fonologické struktury nejasných slov (Bičan, 2018). Obě zmíněné možnosti (a jistě existují i jiné) lingvisté samozřejmě používají. Jejich argumentační síla je omezena mírou znalostí, které o synchronní fonologii češtiny máme.

Poděkování

Příspěvek vznikl v rámci projektu Fonologie českých anglicismů (16-06012S, GA ČR). Práce používá nástroje, které poskytuje výzkumná infrastruktura LINDAT/CLARIAH-CZ (<<https://lindat.cz>>) podporovaná Ministerstvem školství, mládeže a tělovýchovy České republiky (projekt č. LM2018101).

ZDROJE

Akademický slovník současné češtiny. Praha: Ústav pro jazyk český AV ČR, v. v. i. Dostupné online na adrese: <<http://www.slovníkcestiny.cz>>.

ASCS = Petráčková, V. – Kraus, J., et al. (1995): *Akademický slovník cizích slov*. Praha: Academia.

CSN = Sochová, Z. – Poštolková, V. (1994): *Co v slovnících nenajdete*. Praha: Portál.

Český národní korpus. Praha: Ústav Českého národního korpusu FF UK. Dostupné online na adrese: <<http://www.korpus.cz>>.

- Databáze heslářů. Praha: Ústav pro jazyk český AV ČR, v. v. i. Dostupné online na adrese: <<https://lexiko.ujc.cas.cz/heslare/>>.
- FDČA = *Fonologická databáze českých anglicismů*. Nepublikovaná databáze.
- FKČ = *Fonologický korpus češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i. Dostupné online na adrese: <<http://www.ujc.cas.cz/phword/>>.
- FSČ = Čermák, F. – Křen, M. (eds.) (2004): *Frekvenční slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- NASCS = Kraus, J., a kol. (2005): *Nový akademický slovník cizích slov*. Praha: Academia.
- PSJČ = *Příruční slovník jazyka českého*. Praha: Státní nakladatelství – Školní nakladatelství – Státní pedagogické nakladatelství, 1935–1957.
- SN = Martincová, O., a kol. (1998, 2004): *Nová slova v češtině: Slovník neologizmů; Nová slova v češtině: Slovník neologizmů, 2*. Praha: Academia.
- SSČ = Filipce, J. – Daneš, F. – Machač, J. – Mejstřík, V. (eds.) (1978, 1994): *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia.
- SSJČ = Havránek, B. (ed.) (1960–1971): *Slovník spisovného jazyka českého*. Praha: Nakladatelství Československé akademie věd – Academia.
- SSouČ = *Slovník současné češtiny*. Brno: Lingea. Dostupné online na adrese: <<https://www.nechybujte.cz/slovník-soucasne-cestiny/>>.
- VŠČ = Romportl, M., a kol. (1978): *Výslovnost spisovné češtiny: Výslovnost slov přejatých: Výslovnostní slovník*. Praha: Academia.

LITERATURA

- BIČAN, A. (2013): *Phonotactics of Czech*. Frankfurt am Main – Berlin – Bern – Bruxelles – New York – Oxford – Wien: Peter Lang.
- BIČAN, A. (2018): Etymologie a fonologie: Případ Mathesiova fonotaktického pravidla. In: P. Malčik (ed.), *Vesper Slavicus: Sborník k nedožitým devadesátinám prof. Radoslava Večerky*. Praha: Nakladatelství Lidové noviny, s. 13–31.
- BIČAN, A. – DUBĚDA, T. – HAVLÍK, M. – ŠTĚPÁNOVÁ, V. (2020): *Fonologie českých anglicismů*. Praha: NLN.
- DOKULIL, M. – HORÁLEK, K. – HŮRKOVÁ, J. – KNAPPOVÁ, M. (eds.) (1986): *Mluvnice češtiny, 1: Fonetika – Fonologie – Morfonologie a morfejmika – Tvoření slov*. Praha: Academia.
- DUANMU, S. (2008): *Syllable Structure: The Limits of Variation*. Oxford: Oxford University Press.
- DUBĚDA, T. (2018): Fonologická adaptace anglicismů v kvantitativním pohledu. *Naše řeč*, 101, s. 245–267. *Etymologický slovník jazyka staroslověnského*. Praha – Brno: Academia – Tribuna, 1989–2016.
- FRIES, C. C. – PIKE, K. (1949): Coexistent phonemic systems. *Language*, 25, s. 29–50.
- GORDON, M. (2016): *Phonological Typology*. Oxford: Oxford University Press.
- GRYGAROVÁ-RECHZIEGLOVÁ, A. (1993): K otázce měkkosti a kombinatoriky českých fonémů. *Slovo a slovesnost*, 54, s. 255–269.
- HAUGEN, E. (1950): The analysis of linguistic borrowing. *Language*, 26, s. 210–231.
- HOLUB, J. – KOPEČNÝ, F. (1952): *Etymologický slovník jazyka českého*. Praha: Státní nakladatelství.
- ITŌ, J. – MESTER, A. (1999): The phonological lexicon. In: N. Tsujimura (ed.), *The Handbook of Japanese Linguistics*. Oxford: Oxford University Press, s. 62–100.
- JANSON, T. (1986): Cross-linguistic trends in the frequency of CV sequences. *Phonology Yearbook*, 3, s. 179–195.
- KOMÁREK, M. (1962): *Historická mluvnice česká, 1: Hláskosloví*. Praha: Státní pedagogické nakladatelství.
- KUČERA, H. (1958): Inquiry into coexistent phonemic systems in Slavic languages. In: *American Contributions to the Fourth International Congress of Slavists, Moscow, September 1958*. 's-Gravenhage: Mouton, s. 169–189.

- LAMPRECHT, A. – ŠLOSAR, D. – BAUER, J. (1986): *Historická mluvnice češtiny*. Praha: Státní pedagogické nakladatelství.
- LUDVÍKOVÁ, M. (1968): Kombinatorika českých fonémů z kvantitativního hlediska. *Slovo a slovesnost*, 29, s. 56–65.
- LUDVÍKOVÁ, M. – KRAUS, J. (1966): Kvantitativní vlastnosti soustavy českých fonémů. *Slovo a slovesnost*, 27, s. 334–344.
- MACNEILAGE, P. F. – DAVIS, B. L. – KINNEY, A. – MATYEAR, C. L. (2000): The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development*, 71, s. 153–163.
- MAČUTEK, J. – WIMMER, G. (2013): Evaluating goodness-of-fit of discreet distribution models in quantitative linguistics. *Journal of Quantitative Linguistics*, 20, s. 227–240.
- MADDIESON, I. (1984): *Patterns of Sounds*. Cambridge: Cambridge University Press.
- MADDIESON, I. (1992): The structure of segment sequences. *UCLA Working Papers in Phonetics*, 83, s. 1–8.
- MADDIESON, I. – PRECODA, K. (1992): Syllable structure and phonetic models. *Phonology*, 9, s. 45–60.
- MACHEK, V. (1930): *Studie o tvoření výrazů expresivních*. Praha: Nákladem Filosofické fakulty Univerzity Karlovy.
- MACHEK, V. (1968): *Etymologický slovník jazyka českého*. Praha: Academia.
- MALÉCOT, A. (1974): Frequency of occurrence of French phonemes and consonant clusters. *Phonetica*, 29, s. 158–170.
- MATHESIUŠ, V. (1931): O výrazové platnosti některých českých skupin hláskových. *Naše řeč*, 15, s. 38–40.
- MATHESIUŠ, V. (1932): Cizí slova se stanoviska synchronického. *Časopis pro moderní filologii*, 18, s. 231–239.
- NĚMEC, I. (1971): Hláskoslovné předpoklady produktivity slovtvorných formantů. *Slovo a slovesnost*, 32, s. 217–224.
- NOVOTNÁ, R. – KARLÍKOVÁ, H. (2016): Zvukomalba. In: P. Karlík – M. Nekula – J. Pleskalová (eds.), *Nový encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny, s. 2131–2132. Dostupné online na adrese: <<https://www.czechency.org/slovník/ZVUKOMALBA>>.
- PARADIS, C. – LACHARITÉ, D. (2005): Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry*, 36, s. 223–258.
- POKORNÁ, E. (1979): O výslovnosti přejatých slov v češtině. *Typografia*, 82, s. 181.
- POPŁACK, S. (2018): *Borrowing*. Oxford: Oxford University Press.
- REJZEK, J. (2015): *Český etymologický slovník*. Praha: Leda.
- RECHZIEGEL, A. (1998): On one case of anomaly in Czech phonotactics. In: A. A. Barentsen – B. M. Groen – J. Schaecken – R. Sprenger (eds.), *Dutch Contributions to the Twelfth International Congress of Slavists*. Amsterdam – Atlanta: Rodopi, s. 329–349.
- RECHZIEGLOVÁ, A. (1998): K fonologii expresivity v češtině. In: *Slavica Pragensia ad tempora nostra: Sborník statí z mezinárodní vědecké konference ke 150. výročí založení stolice slovanské filologie na Karlově univerzitě, Praha, 8.–10. června 1998*. Praha: Euroslavica, s. 129–134.
- SŁAWSKI, F. (ed.) (1974): *Słownik prasłowiański, 1*. Wrocław: Zakład Narodowy imienia Ossolińskich.
- ŠTĚPÁNOVÁ, V. (2013): Výslovnost cizích slov, vlastních jmen, zkratk a některé další fonetické dotazy v jazykové poradně. *Naše řeč*, 96, s. 117–140.
- ŠTĚPÁNOVÁ, V. (2015): Vokální kvantita v přejatých slovech v češtině: Výsledky výzkumu výslovnostního úzu. *Naše řeč*, 98, s. 169–189.
- ŠTĚPÁNOVÁ, V. (2016): Kolísání znělosti souhlásek v přejatých slovech v češtině: Výsledky výzkumu výslovnostního úzu. *Naše řeč*, 99, s. 169–188.
- TRÁVNÍČEK, F. (1935): *Historická mluvnice československá: Úvod, hláskoslovi a tvaroslovi*. Praha: Melantrich.
- TRNKA, B. (1937): *Pokus o vědeckou teorii a praktickou reformu těsnopisu*. Praha: Nákladem Filosofické fakulty Univerzity Karlovy.

- TRNKA, B. (1942): O fonologických cizostech v češtině. *Slovo a slovesnost*, 8, s. 20–27.
- TRUBETZKOY, N. (1939): *Grundzüge der Phonologie*. Praha: Jednota československých matematiků a fyziků.
- VACHEK, J. (1940): Poznámky k fonologii českého lexika. *Listy filologické*, 67, s. 395–402.
- VACHEK, J. (1968): *Dynamika fonologického systému současné češtiny*. Praha: Academia.
- VALLEÉ, N. – ROSSATO, S. – ROUSSET, I. (2009): Favoured syllabic patterns in the world's languages and sensorimotor constraints. In: F. Pellegrino – E. Marsico – I. Chitoran – Ch. Coupé (eds.), *Approaches to Phonological Complexity*. Berlin – New York: Mouton de Gruyter, s. 111–140.
- WHALEN, D. H. – GIULIVI, S. – NAM, H. – LEVITT, A. G. – HALLÉ, P. – GOLDSTEIN, L. M. (2012): Biomechanically preferred consonant-vowel combinations fail to appear in adult spoken corpora. *Language and Speech*, 55, s. 503–515.
- ZIMA, J. (1961): *Expresivita slova v současné češtině*. Praha: Nakladatelství Československé akademie věd.

Ústav pro jazyk český AV ČR, v. v. i.

Veveří 97, 602 00 Brno

<bican@ujc.cas.cz>