



CZEDD a SLAN – online nástroje zaměřené na práci se slovtvorbou češtiny jako druhého/cizího jazyka¹

Ivana Kolářová – Adriana Válková

ABSTRAKT:

Nástroje CZEDD (Czech electronic derivational dictionary) a SLAN (Slovtvorný analyzátor pro výuku češtiny pro cizince) představují systematický způsob, jak přistupovat k výuce slovtvorby češtiny u nerodilých mluvčích. Oba nástroje vychází ze základní vlastnosti češtiny jakožto morfologického typu jazyka, jehož slovní zásobu tvoří více než 75 % slov odvozených. Principem obou nástrojů je učit studenty-cizince vztahovat odvozená slova (např. *vládce*) ke slovům základovým (např. *vládnout*) a současně uplatňovat princip kompozicionality (např. *samoplátce* souvisí se spojením slov *platit sám*, přičemž *-ce* indikuje osobu). Oba nástroje vychází z dat učebnicových korpusů cílených na jazykovou úroveň A2, B1 a B2, díky čemuž je možné uchopit frekventované (produktivní) slovtvorné modely, které odpovídají jazykové úrovni uživatele (studenta-cizince). Pravidelným užíváním aplikací lze předpokládat navýšení pasivní, popř. i aktivní slovní zásoby nerodilých mluvčích češtiny.

KLÍČOVÁ SLOVA:

slovtvorba, slovesné prefixy, sémantika, čeština jako druhý jazyk, korpusový výzkum

ABSTRACT:

The tools CZEDD (Czech Electronic Derivational Dictionary) and SLAN (a word-form analyzer for teaching Czech to foreigners) represent a systematic way of approaching the teaching of Czech word-formation to non-native speakers. Both tools are based on the basic properties of Czech as a morphological type of language, whose vocabulary consists of more than 75% of derived words. The principle of both tools is to teach foreign students to relate derived words (e.g. *vládce* 'ruler') to basic words (e.g. *vládnout* 'to rule'), at the same time applying the principle of compositionality (e.g. *samoplátce* 'self-payer,' which indicates a person, is related to the combination of words *platit sám* 'to pay for oneself'). Both tools are based on data from textbook corpora targeted at the language level A2, B1 and B2. Thanks to this it is possible to grasp frequent (productive) word-formation models specific to the language level of the foreign learner. Regular use of applications can be expected to increase the passive and active vocabulary of non-native speakers of Czech.

KEYWORDS:

Word formation, Verb prefixes, Semantics, Czech as a foreign language, Corpus research

¹ Text vznikl v rámci projektů TL03000293 Slovtvorný analyzátor pro výuku češtiny pro cizince, který je spolufinancován se státní podporou Technologické agentury ČR v rámci Programu ÉTA, a MUNI/A/1181/2020 Gramatika a lexikon češtiny.

1. VÝCHODISKA OBOU PROJEKTŮ

SLOVOTVORBA A JEJÍ POTENCIÁL V ČCJ

Čeština je flektivním jazykem se značně rozvinutou morfologií flektivní (flexí) a derivační (tj. derivací v širším smyslu). Flexi lze u takového typu jazyka chápat jako jeho schopnost vyjadřovat syntaktické vztahy mezi slovy a derivaci (v širším smyslu) jako jeho schopnost tvořit (nová) slova na základě slov v jazyce již existujících. Slovo-tvorba, pod kterou derivační morfologii řadíme, se zabývá tvořením jednoslovných pojmenování (např. *vést vodu* → *vodovod*, *les* → *lesík*, *Česká republika* → *ČR*) a v jazykovědném systému ji lze řadit mezi lexikologii (tj. nauku o slovní zásobě) a morfologii (tj. nauku o tvarech slov). Takového postavení lze velice dobře využít v oblasti češtiny jako cizího jazyka (dále jen ČCJ). Spjatost s lexikologií lze dle našeho mínění uchopit myšlenkou, že by si student-cizinec mohl vystačit se znalostí slov z jádra slovní zásoby, což jsou zpravidla slova ze současného pohledu neodvozená (např. *dům*, *les*, *jít*) a ostatní, tj. utvořená slova (např. *domovní*, *lesník*, *odejít*) by si mohl na základě znalosti neodvozených slov a slovo-tvorných pravidel odvodit. Ze starší literatury (Dokulil 1962) vyplývá, že odvozená slova tvoří asi 75 % z celé slovní zásoby, lze však předpokládat, že toto číslo je vyšší.² Spjatost s morfologií lze využít zejm. u sufixálně utvořených slov, která lze často na základě jejich zakončení přiřadit ke slovnímu druhu (např. slova zakončená na *-tel* a *-dlo* jsou podstatnými jmény) a v případech ohebných slovních druhů i k jejich paradigmatickým vzorům (např. slova zakončená na *-tel* se až na pár výjimek skloňují podle vzoru *muž*, slova s *-dlo* podle vzoru *město*).

SOUČASNÝ PŘÍSTUP K PREZENTACI SLOVOTVORBY V ČCJ

Ve srovnání s ostatními tematickými celky jazykového vyučování je slovo-tvorbě věnován malý prostor jak v češtině pro cizince, tak ve výuce češtiny jako mateřského jazyka na základní škole: ve stěžejních kurikulárních dokumentech — v *Rámcovém vzdělávacím programu pro základní vzdělávání* a ve *Standardech pro základní vzdělávání* slovo-tvorba je pouze částí jednoho z osmi výstupů: u žáků dosahujících minimální nebo optimální úroveň se na konci 9. třídy předpokládá schopnost poznat, zda bylo slovo utvořeno odvozováním, skládáním, zkracováním, schopnost doplnit předpony ke slovesům a poznat pak jejich význam nebo utvořit složená slova podle zadání, např. *název místa, kde se vaří pivo; název činnosti, při které se chytají ryby*. Pouze žáci, kteří dosahují excelentní úroveň, mají dokázat vytvořit podle významu a slovo-tvorné stavby např. *pojmenování člověka, který sbírá, maluje, plave, hlídá*.

V dokumentech MŠMT *Čeština jako cizí jazyk* (SERR 2001) pro úroveň A1 (Hádková-Línek-Vlasáková 2005), úroveň A2 (Čadská a kol. 2005) a úroveň B2 (Adamovičová a kol. 2005) se předpokládá, že na úrovni A1 se se student naučí tvořit názvy přechýlené — obecná jména (např. *profesor* — *profesorka*) včetně příjmení (např. *Dobeš* — *Dobešová*, *Starý* — *Stará*); jména obyvatelská, jména zemí a světadílů a pak další

² Ve slovenštině, tj. v jazyce češtině nejbližším, vyplývá z aktuálních dat, že zastoupení slovo-tvorně motivovaných slov je 83,6 % (Ološtiak 2015).





slovní druhy: adjektiva od názvů obyvatelských a slovesa typu *kupovat*, tj. s příponou *-ova(t)*. Na úrovni A2 se počítá s dovedností odvozovat substantiva, adjektiva, slovesa i adverbia různými slovtvornými způsoby a prostředky. Výrazně rozvinutá dovednost tvořit slova různými způsoby a procesy včetně tvoření abstraktních názvů dějů, vlastností, tvoření adjektiv různými sufixy a tvoření sloves včetně reflexivizace se předpokládá na úrovni B2. V materiálu MŠMT *Čeština jako cizí jazyk. Úroveň B2* jsou utvořená slova rozdělena již podle slovtvorných kategorií (přechýlené názvy, názvy osob podle zemí původu, jména konatelská, jména činitelská...).

Je zřejmé, že v současnosti nenacházíme žádné ucelené popisy slovtvorby nebo nástroj/e, jenž by vedly studenta-cizince k postupnému si osvojování slovtvorných modelů v češtině. Dle našeho mínění se popis slovtvorného systému pro potřeby nerodilých mluvčích češtiny může částečně překrývat se slovtvornými popisy určené pro rodilé mluvčí (např. při popisu hláskových alternací), ale na rozdíl od něj vyžaduje popis skutečností, které jsou rodilým mluvčím implicitně známy:

- polyfunkčnost některých sufixů: např. sufix *-tel* se využívá téměř výhradně pro tvoření názvů osob, sufix *-iště* téměř výhradně pro tvoření názvů míst, zatímco např. sufix *-ec* může být jednou součástí názvu osoby (např. *kupec*), podruhé součástí názvu objektu (např. *podstavec*); sufix *-ík/-ník* může tvořit název osoby (např. *pracovník, zedník*), název místa (např. *chodník*) i název předmětu (např. *popelník*);
- různé části základového slova, které se stávají slovtvorným základem; sufix se při tvoření slov připojuje k základu, kterým může být celé základové slovo zkrácené pouze o koncovku: *káv-a* — *kav-árna*, nebo také o příponu: *tisk-nou-t* — *tisk-árna*, popř. jinou část: *šir-oký* — *šířka*.

JAZYKOVÉ (SLOVOTVORNÉ) ZAMĚŘENÍ PROJEKTŮ

Oba projekty cílí primárně na slova utvořená derivací, neboť se jedná o nejčastější postup, kterým se nová pojmenování tvoří. Primárně jsou tedy zpracována (derivovaná) slova utvořená pomocí předpony (např. *dojít*), přípony (např. *stromek*), předponou a příponou současně (např. *pahorek*) nebo reflexivizací (např. *povídat si*). V podstatně menší míře jsou dále zpracovávána slova utvořená jinými slovtvornými postupy, např. kompozicí (např. *šéfkuchař*), dále morfy na pomezí derivate a kompozice, které se označují jako afixoidy (např. *samoplátce*) či postupy dalšími (např. *profík*).

V současné chvíli je v rámci obou projektů věnována pozornost jazykovým úrovním A2, B1 a B2. Dle výsledků vyplývajících z dotazníkových šetření (Sokolová-Válková) a z rešerší učebnic češtiny pro cizince (např. Malá 2021) nemá smysl se na nižší jazykové úrovni (tj. A1) slovtvorbou zabývat.³

³ Pochopitelně nepopíráme možnost pracovat se slovtvorbou zejm. u Slovanů (např. Rusů) již od úrovně A1, tedy jsme si vědomi výrazných podobností mezi slovanskými jazyky. Nicméně na rozdíl od úrovně A2 nenacházíme pro tuto jazykovou úroveň žádné podklady, jež by jednoznačně prokazovaly smysluplnost (potenciál) této výuky.

VÝCHOZÍ MATERIÁLY: UČEBNICOVÝ KORPUS UČEBKO A MATERIÁLY MŠMT



Pro oba projekty se stal primárním zdrojem dat specializovaný korpus UčebKo vytvořený z učebnic češtiny pro cizince (Válková, Válková 2021a). Korpus zahrnuje celkem tři subkorporusy: UčebKo-A2 (Válková, 2021b) vytvořený z učebnic pro úroveň A2, UčebKo-B1 (Válková, 2021c) vytvořený z učebnic pro úroveň B1 a UčebKo-B2 (Válková, 2021d) vytvořený z učebnic pro úroveň B2, díky čemuž je možné pracovat pouze s těmi slovy (a pouze s těmi jejich významy), s nimiž se studenti češtiny jako cizího jazyka na své úrovni skutečně setkávají.

Korpus UčebKo byl vytvořen přes rozhraní Sketch Engine (Kilgarriff a kol.), díky čemuž je možné pracovat s frekvenčními seznamy odvozených slov, které patří ke stejnému slovotvornému typu.⁴ Pomocí konkordancí je možné zjistit (a ověřit si) významy zpracovávaných slov a taktéž si ověřit výsledky automatických procesů — zejm. lemmatizace (Jakubíček-Kovář-Šmerk 2011; Šmerk 2008), která vykazuje určitou míru chybovosti (např. tvar *rovně* byl chybně lemmatizován jako *roven*). Učebnicový korpus čítá kolem 300 000 slov (dílčí subkorporusy čítají kolem 100 000 slov), čímž se řadí spíše k menším korpusům. Korpusy takové velikosti by měly postačovat k zachycení nejméně frekventovanějších jevů, nicméně se ukázalo, že tomu tak ve všech případech není. Důvodem je výchozí materiál korpusů, tedy učebnice. Nejedná se totiž o souvislý text, ale slova se v učebnicích vyskytují i izolovaně (tj. bez větného kontextu), např. jako součást popisů obrázků. Taková slova se pak nevyskytují v korpusech (neboť korpus tvoří jen celé věty), a tedy ani v našich výchozích datech. Právě z tohoto důvodu jsou seznamy slov z korpusů doplněny o slova z materiálů MŠMT pro referenční úrovně A1, A2 a B2, které jsou na stránkách ministerstva volně dostupné.⁵

METODOLOGIE PRO ZPRACOVÁNÍ ODVOZENÝCH SLOV

Při určování slovotvorné příbuznosti mezi slovy vycházíme ze stále převládající slovotvorné teorie představené v 60. letech 20. století Milošem Dokulilem (viz Dokulil 1962 a Daneš-Dokulil-Kučař 1967). Na základě této teorie jsou dvě slova (či více slov) slovotvorně příbuzná, jestliže je mezi nimi dodržen vztah formální příbuznosti, tzv. slovotvorné fundace (např. slovo *čekárna* se formálně zakládá na slově *čekat*), a vztah sémantické příbuznosti, tzv. slovotvorné motivace (význam slova *čekárna* lze vysvětlit skrze význam slova *čekat*). Ze současného pohledu proto nepokládáme za odvozená slova ta, u nichž je dodržen pouze formální vztah příbuznosti (např. slovo *silnice* nepokládáme ze současného pohledu za motivované slovem *silný*).

Cílem obou projektů je zprostředkovat uživatelům slova odvozená (např. *ředitel*, *pracoviště*) a slova, ze kterých byla tato slova utvořena (nebo se jeví, že byla utvořena)

4 Slovotvorný typ zahrnuje odvozená slova, která (a) patří k jednomu slovnímu druhu, (b) mají stejný významový vztah ke svým fundujícím slovům, (c) vznikla stejným slovotvorným postupem a (d) byla utvořena stejným slovotvorným formantem (Osolsobě 2017).

5 Dostupné z: <https://www.msmt.cz/mezinarodni-vztahy/referencni-urovne-pro-cestinu-jako-cizi-jazyk>.



(např. *řídít, pracovat*). V projektu CZEDD je cílem taktéž popsat a zprostředkovat slovo-
 tvorné prostředky a mechanismy, pomocí kterých byla odvozená slova utvořena
 (např. slovo *ředitel* bylo utvořeno od slova *řídít*, kdy připojením přípony *-tel* došlo
 k hláskové alternaci *í > e*). Při určování základových slov se opíráme o práce, které
 rovněž vychází z Dokulilovy teorie tvoření slov, jedná se zejména o *Slovník afixů uží-
 vaných v češtině* (Šimandl a kol. 2017) a o *Velkou akademickou gramatiku spisovné češ-
 tiny* (Štícha a kol. 2018). Při stanovení významové příbuznosti mezi slovy přihlížíme
 v případě potřeby do výkladového *Slovníku spisovného jazyka českého* (I-VIII /Havránek
 1989/, elektronická verze viz <https://ssjc.ujc.cas.cz/>) a pro rozlišení přejetých slov,
 u kterých funguje přípona jako čistě formální prostředek, od slov skutečně utvoře-
 ných nahlížíme do *Českého etymologického slovníku* (Rejzek 2012).

Z hlediska formy bývají slova odvozená (např. *zubatý, chudoba, dvakrát*) formálně
 složitější než slova základová (např. *zub, chudý, dva*). Ve výjimečných případech je
 tomu naopak (např. *lov* → *lovit*). Dále jsou případy, kdy směr odvozování není zřejmý
 (např. *blázen* → *bláznit* nebo *bláznit* → *blázen?*) nebo případy, kdy je určení slova zá-
 kladového problematické (např. *soudit* → *soudce* nebo *soud* → *soudce?*). V takových
 případech upřednostňujeme to slovo, které se dané úrovni vyskytuje (tj. v korpu-
 sech UčebKo či v materiálech MŠMT), a pokud se na příslušné úrovni vyskytují obě
 slova, volíme to, které je v obecných korpusech ČNK řady SYN (Křen 2019) frekven-
 tovanější.

Dále se setkáváme s případy, kdy odvozené slovo (např. *zákazník, sídliště*) je sou-
 částí lexika dané jazykové úrovně, ale základové slovo (např. *zakázka, sídlit*) již niko-
 liv. Na takové případy jsou uživatelé v obou aplikacích upozorněni.

2. CZECH ELECTRONIC DERIVATIONAL DICTIONARY (CZEDD)

CZEDD je součástí doktorského projektu realizovaného na Ústavu českého jazyka na
 Filozofické fakultě Masarykovy univerzity v Brně. Jedná se o online nástroj, jehož cí-
 lem je naučit uživatele porozumět významům odvozených slov na základě jejich (slo-
 votvorné) struktury. Nástroj tak pracuje explicitně vedle slovo-*tvorného* vztahu fun-
 dace (tj. formálního vztahu) i se slovo-*tvorným* vztahem motivace (tj. významovým
 vztahům). Z tohoto pohledu CZEDD představuje možnost, jak využít prediktability
 významu odvozených slov v oblasti češtiny jako cizího jazyka. CZEDD je určen jak
 studentům češtiny jako cizího/druhého jazyka pro samostudium, tak lektorům pro
 přípravu výuky zahrnující slovo-*tvorbu*. Veškeré popisy v aplikaci jsou psány anglicky
 a nevyžadují žádné lingvistické znalosti, neboť se domníváme, že se tak dosažitel-
 nost aplikace i pro běžné uživatele jazyka zvyšuje. Z hlediska funkcionality je CZEDD
 primárně specializovaným typem slovníku, jehož heslář tvoří pouze slova odvozená.
 Uživatel si jeho pomocí osvojuje základní slovo-*tvorné* vztahy mezi slovy, které jsou
 nezbytnou podmínkou slovo-*tvorné* příbuznosti mezi slovy. Další funkcí, kterou ná-
 stroj uživateli nabízí a která je propojena s funkcí slovníkovou, je automatické zpra-
 cování textu, ve kterém se uživateli označí všechna odvozená slova, která následně
 fungují jako slovníková hesla.



CZEDD JAKO SLOVNÍK

Funkce slovníku nabízí jeho uživatelům 1) definici odvozeného slova na základě jeho vztahu ke slovu základovému (např. *učitel* je *ten, kdo učí*), vč. její anglické verze, viz obr. 1, 2) popis utvářenosti slovníkového hesla (tedy to, jak bylo slovo utvořeno nebo se jeví, že bylo utvořeno), vč. hláskových alternací, viz obr. 2, 3) informaci o jazykové úrovni heslového slova (A2, B1 nebo B2), viz obr. 3, 4) sémantické rysy, které mají za cíl přiblížit obecnější slovtvorný význam specifičtějšímu významu lexikálnímu, viz taktéž obr. 3, a 5) ukázkové věty použití (dle jazykových úrovní A2, B1, B2), viz obr. 4. Dále má uživatel možnost zobrazit si u slov (např. *zákazník*), která jsou východiskem dalšího odvozování, jejich další bezprostřední deriváty (např. *zákaznice*, *zákazníkův*). Bezprostřední deriváty jsou graficky zobrazovány formou bublin a jsou setříděny dle jazykových úrovní A2–B2. Uživatel má tak možnost postupně se proklikat k celému slovtvornému hnízdu.

ředitel = ten, kdo řídí
(a director) (someone who directs)

OBRÁZEK 1: Slovníková definice (CZEDD)

řít(t) → ředitel
(í > e) + -tel

OBRÁZEK 3: Jazyková úroveň hesla a jeho sémantické rysy (CZEDD)

A2
#člověk #muž #profese

OBRÁZEK 2: Popis utvářenosti slovníkového hesla (CZEDD)

Example sentences:

A2
Pan ředitel má schůzku.
Ředitel společnosti KIKO je momentálně na dovolené.

B1
Pan učitel Konečný vyučuje češtinu a angličtinu.
Musíte se zeptat řídícího učitele.

OBRÁZEK 4: Ukázkové věty (CZEDD)



CZEDD – NÁSTROJ PRO AUTOMATICKÉ VYZNAČENÍ ODVOZENÝCH SLOV V TEXTU

CZEDD lze použít jako nástroj pro zpracování textu, ve kterém jsou automaticky vyznačena všechna odvozená slova, viz obr. 5. Funkce je propojena s funkcí slovníkovou, proto se může uživatel u každého vyznačeného (resp. odvozeného) slova prokliknout do slovníku.



OBRÁZEK 5: Ukázka automatického vyznačení odvozených slov v textu (CZEDD)

3. SLOVOTVORNÝ ANALYZÁTOR PRO VÝUKU ČEŠTINY PRO CIZINCE (SLAN)

SLAN je grantovým projektem podpořeným Technologickou agenturou České republiky (TAČR),⁶ který je realizován na Pedagogické fakultě Masarykovy univerzity v Brně a jehož výstup má podobu webové aplikace. Projekt se snaží nabídnout možnosti přístupu ke slovo tvorbě ve výuce češtiny jako cizího/druhého jazyka na různých referenčních úrovních a poskytnout:

- lektorům, popř. autorům učebnic češtiny jako cizího jazyka doporučení a metody pro výuku slovo tvorby v češtině pro cizince;
- studentům-cizincům i jazykový materiál — slovní zásobu a cvičení pro cílené a systematické procvičování slovo tvorby.

Na rozdíl od nástroje CZEDD pracuje SLAN explicitně pouze se slovo tvornou fundací (tj. formálním zakládáním se slovo tvorně příbuzných slov), slovo tvornou motivaci zohledňuje, ale pouze implicitně.

⁶ Článek představuje projekt TL03000293 Slovo tvorný analyzátor pro výuku češtiny pro cizince (SLAN). Informace o projektu a řešitelském týmu jsou zde: <<https://www.muni.cz/vyzkum/projekty/55407>>. Novinky můžete sledovat zde: <<https://www.researchgate.net/lab/SLAN-Word-formation-Analysis-Software-for-Teaching-Czech-for-Foreigners-Anastasija-Sokolova>>.

POPIS ČESKÉ SLOVOTVORNÉ SOUSTAVY V RÁMCI PROJEKTU SLAN
JAKO VÝCHODISKO VÝUKY ČEŠTINY PRO CIZINCE

Cílem projektu je vytvoření webové aplikace k praktickému procvičování slovo- tvorby, jejíž příprava je nyní v počáteční fázi. V hotové verzi bude studentům i lektorům k dispozici repertoár vytvořených slov, která by studenti češtiny jako cizího jazyka měli umět utvořit, protože jsou zařazena do učebnic češtiny pro cizince. Při výběru slov jsme dbali na to, aby šlo o slova se zřetelnou slovotvornou strukturou, kterou pochopí i student-cizinec, tj. aby u nich byla jasná hranice mezi slovotvorným základem a slovotvorným prostředkem. Webovou aplikaci vytvářejí spolupracující programátoři. Lingvisté z projektového týmu pro tento účel postupně zpracovávají seznamy slov, k nimž přiřazují relevantní informace nezbytné pro to, aby bylo ve webové aplikaci možné tato slova vyhledávat a seskupovat podle zvolených kritérií.

Slova jsou již rozdělena podle kategoriálního významu (podle toho, zda pojmenovávají osobu, místo, děj ...) a slovotvorných prostředků. K odvozeným slovům jsou vždy přiřazeny důležité informace o odvozeném slově, především typické kolokace, v nichž je slovo v korpusech doloženo, dále slovo základové a v neposlední řadě jazyková úroveň slova, viz Tab. 1.

odvozené slovo	častá spojení, poznámka	základové slovo	poznámka	úroveň
bydliště	příklady: <i>trvalé bydliště</i>	bydlet		A2
hřiště	příklady: <i>střed hřiště; domácí/foťbalové/golfové dětské hřiště; být blízko hřiště; chodit s kamarády na hřiště</i>	hrát		A2

TABULKA 1: Příklad zpracování informací o utvořených slovech (názvy míst s příponou *-iště*)

K důležitým informacím kromě kolokací patří upozornění o významech víceznačných slov, viz Tab. 2.

odvozené slovo	častá spojení, poznámka	základové slovo	poznámka	úroveň
tiskárna	příklady: <i>V kanceláři je počítač, myš, nabíječka, mobil a tiskárna.</i> význam: 1. místo, kde se tisknou knihy, noviny, časopisy; 2. prostředek/nástroj, tj. <i>tiskárna u domácího počítače</i>)	tisknout		A2

TABULKA 2: Ukázka zpracování polysémního slova tiskárna

Některá slova není možné ze současného pohledu považovat za slova odvozená, neboť není splněn vztah slovotvorné motivace (tedy význam odvozeného slova nelze vysvětlit skrze význam slova základového), viz Tab. 3. Tato slova jsou ponechána v apli-



kaci, neboť si díky nim student-cizinec uvědomí, že ne ve všech případech je možné vysvětlit si význam odvozeného slova skrze jeho strukturu.

odvozené slovo	častá spojení, poznámka	základové slovo	poznámka	úroveň
pokladna	příklady: <i>pokladna divadla; jít k pokladně; hledat pokladnu; platit u pokladny</i>	poklad	pozor: pomocí slova <i>poklad</i> nejde vysvětlit význam slova <i>pokladna</i>	A2

TABULKA 3: Ukázka zpracování neprediktabilního slova pokladna

V současné době je na webové aplikaci k dispozici částečně funkční sekce pro studenty, která jim umožní vyhledávat:

- **slova podle významu** (např. názvy osob, názvy prostředků, názvy vlastností, názvy míst...);
- **slova podle přípon** nebo **předpon**, např. názvy osob s příponou *-tel*, názvy prostředků s příponou *-dlo*, názvy míst s příponou *-iště...*, slovesa s předponou *na-*;
- **slova utvořená různými slovotvornými prostředky od jednoho slova základového**, např. *sportovat* — *sportov-ec*, *sportov-iště*; popř. **slova se společným kořenem:** *sport* — *sportovat* — *sportovní* — *sportovec* — *sportoviště*.

Současně vzniká i rozhraní pro lektory, kteří budou mít stejnou nabídku jako studenti, avšak rozšířenou o další možnosti:

- **vygenerovat věty a z nich vytvořit cvičení**, v nichž si studenti budou procvičovat slovotvorbu — tvořit slova podle různých kritérií, např. názvy osob vykonávajících činnost pojmenovanou slovesy: *Adamův dědeček pracoval jako (učit)*. *Kdo je váš oblíbený (sportovat)?*; názvy osob podle předmětu/objektu činnosti: *(Noviny) se ptá herce na jeho oblíbené jídlo*; názvy míst podle děje: *Letadlo přistává na (letět)*.
- **možnost ke každému slovu vyhledat ruský a anglický ekvivalent** bez ohledu na ekvivalenci/neekvivalenci způsobu tvoření a slovotvorných prostředků.

UKÁZKA VÝSLEDKŮ ČÁSTI VÝZKUMU VÝZNAMOVÉ SKUPINY NÁZVŮ MÍST

V materiálech MŠMT pro referenční úrovně A1, A2 a B2 názvy míst překvapivě nejsou vyděleny jako zvláštní skupina, přestože i cizinec potřebuje pojmenovávat místa ve všech komunikačních situacích a značná část jejich názvů patří k utvořeným slovům. Proto je pro potřeby projektu a webové aplikace jako speciální významovou skupinu vydělujeme.

Nejvíce frekventované jsou pro tvoření názvů míst přípony *-nice/-ice* a *-na*, kromě nich se pak uplatňují přípony *-iště* (včetně sufixů *-aliště*, *-iliště*, *-oviště*), *-árna*, *-ovna*, popř. *-ina*, *-ka*, *-ství* a periferní *-ník*, *-telna*.



Jako ilustrativní příklady jsme vybrali názvy s příponami *-iště* (s rozšířenými podobami *-iliště*, *-aliště*, *-oviště*) a *-árna*. Za vhodné k tomuto účelu je považujeme z toho důvodu, že právě těmito dvěma příponami se tvoří téměř výhradně názvy míst, u kterých je zřetelná slovtvorná struktura. Při jejich tvoření dochází v psané podobě pouze k několika běžným druhům alternací:

- h — ž: *trh* — *tržiště*
- r — ř: *hrát* — *hřiště*
- ř — r: *elektřina* — *elektrárna*
- á — a: *stát* — *stanoviště*, *káva* — *kavárna*
- a — á: *nastoupit* — *nástupiště*
- e — o: *účet* — *účetárna*
- í — i: *víno* — *vinárna*
- ou — u: *klouzat* — *kluziště*, *nastoupit* — *nástupiště*; *kouřit* — *kuřárna*, *poslouchat* — *posluchárna*

Názvy s oběma příponami jsme seřadili vždy do dvou typů tabulek:

- abecední přehled odvozených slov s přiřazenými slovy základovými zachycuje Tab. 4 a Tab. 6;
- rozdělení utvořených slov podle výskytu v textech úrovní A2, B1, B2; zvláště jsou vydělena slova, která jsou doložena v textech všech úrovní (viz Tab. 5 a Tab. 7).

odvozené slovo	základové slovo	odvozené slovo	základové slovo
bydliště	bydlet	pracoviště	pracovat
cvičiště	cvičit	přístaviště	přístav
hledišťe	hledět	rodiště	rodit
hřiště	hrát	schodiště	schod
jeviště	jevit	sídliště	sídlit
kluziště	klouzat	smetiště	smetít
kolejiště	kolej	sportoviště	sportovat
koupaliště	koupat	stanoviště	stát
letiště	letět	tržiště	trh
nástupiště	nastoupit	učiliště	učit
parkoviště	parkovat	výstaviště	výstava

TABULKA 4: Abecední seznam názvů s příponou *-iště* a jejich základovými slovy

V Tab. 5 prezentujeme:

- názvy míst s příponou *-iště* doložené v korpusech Učebko A2, B1 a B2 jsou na příslušném řádku, v závorce je zaznamenána frekvence.
- názvy, které nejsou v korpusech Učebko, ale jsou doloženy v materiálech MŠMT pro příslušnou referenční úroveň: *výstaviště* (A2), *kluziště* a *přístaviště* (B2); jsou označeny symbolem *.



— názvy v hranatých závorkách [*jeviště**], [*koupaliště**], [*tržiště**]; jsou doloženy v obou materiálových zdrojích, avšak v textech pro odlišné referenční úrovně.⁷

-iště	A2	bydliště (7/4/5), hřiště (2/6/6), letiště (29/49/3), nástupiště (3/2/2), parkoviště (20/11/1), [pracoviště*]	hledišťe (1), schodišťe (1), sportovišťe (4), stanovišťe (1), výstavišťe* [<i>jeviště*</i>], [<i>koupaliště*</i>], [<i>tržiště*</i>]
	B1	letiště (29/49/3), nástupiště (3/2/2), parkoviště (20/11/1), [pracoviště*]	schodišťe (2), pracovišťe (12), sídlišťe (9), učilišťe (2), kolejišťe (1), koupališťe (1), tržišťe (1)
	B2	parkoviště (20/11/1), [pracoviště*]	sportovišťe (1), stanovišťe (1), pracovišťe (6), sídlišťe (4), učilišťe (1), cvičišťe (1), jevišťe (3), rodišťe (1), smetišťe (1), kluzišťe*, přístavišťe*

TABULKA 5: Rozdělení názvů s příponou *-iště* podle výskytu v textech úrovní A2, B1, B2

Výrazy *bydliště*, *hřiště*, *letiště*, *nástupiště*, *parkoviště* jsou tedy doloženy v textech všech tří úrovní v korpusech Učebko i v materiálech MŠMT.

Rozdíly mezi názvy v korpusech Učebko a v materiálech MŠMT:

- Název *pracoviště* je doložen v materiálech MŠMT v textech pro všechny úrovně, zatímco v korpusech Učebko až pro texty úrovní B1 a B2.
- Názvy *koupaliště* a *tržiště* jsou v korpusech Učebko v textech úrovně B1 (*koupaliště*, *tržiště*) nebo B2 (*jeviště*), v materiálech MŠMT již v textech pro A2.

Je-li název místa doložen v textech všech úrovní kteréhokoliv ze dvou materiálových zdrojů, znamená to, že lektor češtiny může počítat s jeho znalostí u studentů všech úrovní studia češtiny (srov. též Sokolová-Válková). Tabulka ukazuje, že takových názvů s příponou *-iště*, popř. s jejími variantami je cca 27 %.

Současně však může lektor češtiny počítat u studentů všech úrovní se znalostí těch názvů míst, které se objevují v textech pro úroveň A2 (37 %, viz výčet dokladů níže), a to bez ohledu na skutečnost, jsou-li doloženy v našich učebních textech pro vyšší úrovně. Předpokládáme, že znalosti osvojené v počátcích studia, tj. z textů do úrovně A2, by měly být trvalé a přenášet se i do úrovní vyšších:

- slova *hledišťe*, *schodišťe*, *sportovišťe*, *stanovišťe* (18 %) jsou doložena buď pouze v učebnicích pro úroveň A2 z korpusu Učebko, nebo v učebnicích pro úroveň A2 a pro jednu z úrovní B1, nebo B; slovo *výstavišťe* je doloženo pouze v materiálu MŠMT pro A2;
- slova *jevišťe*, *koupališťe*, *tržišťe* (14 %) jsou doložena v učebních textech korpusu Učebko až pro úroveň B1 a B2, zatímco v materiálech MŠMT jsou již v textech A2.

Uvedená data ukazují, že lektor češtiny jako cizího jazyka může počítat se znalostí cca 64 % názvů s příponou *-iště* u studentů všech úrovní češtiny jako cizího jazyka.

⁷ Přebíráme formu grafického znázornění ze stati Sokolová-Válková 2021 (v tisku).



Naopak 36 % názvů s touto příponou je doloženo pouze v textech učebnic B2, tj. nejvyšší ze sledovaných úrovní (*cvičiště, kluziště, přístaviště, rodiště, smetiště*), v textech B1 a B2 (*sídlíště, učiliště*) nebo jen B1 (*kolejiště*).

odvozené slovo	základové slovo	odvozené slovo	základové slovo
cukrárna	cukr	obrazárna	obraz
čekárna	čekat	posluchárna	poslouchat
elektrárna	elektrína	směnárna	směnit
kavárna	káva	sušárna	sušit
kovárna	kovat	teplárna	teplo
kužárna	kouřit	tiskárna	tisknout
lékárna	lék	účetárna	účet
mlékárna	mléko	vinárna	víno

TABULKA 6: Abecední seznam názvů s příponou *-árna* a jejich základovými slovy

V Tab. 7 prezentujeme:

- názvy, které **nejsou v korpusech Učebko, ale jsou doloženy v materiálech MŠMT** pro příslušnou referenční úroveň: *posluchárna, vinárna (A2), kužárna, mlékárna (B2)*. Jsou značeny symbolem *.
- názvy v hranatých závorkách [*cukrárna**], [*směnárna**], které jsou doloženy v obou základních zdrojích, avšak v textech pro odlišné referenční úrovně.

-árna	A2	čekárna (3/2/2), kavárna (37/13/9),	cukrárna (6), elektrárna (1), kovárna (1), směnárna (1), tiskárna (1), vinárna*, posluchárna*
	B1	lékárna (5/13/5), [cukrárna*],	cukrárna (3), tiskárna (3), teplárna (1)
	B2	[směnárna*]	elektrárna (1), směnárna (2), obrazárna (1), sušárna (1), účetárna (1), kužárna*, mlékárna*

Tabulka 7: Rozdělení názvů s příponou *-árna* podle výskytu v textech úrovní A2, B1, B2

Výrazy *čekárna, kavárna, lékárna* jsou tedy doloženy v textech všech tří úrovní v korpusech Učebko i v materiálech MŠMT.

Rozdíly mezi názvy v korpusech Učebko a v materiálech MŠMT:

- Názvy *cukrárna, směnárna* jsou v korpusech Učebko doloženy vždy v textech pro úroveň A2 a pro jednu z úrovní dalších, v materiálech MŠMT v textech pro všechny úrovně.

Se znalostí názvů *cukrárna, čekárna, kavárna, lékárna, směnárna* (31 %) může tedy lektor češtiny počítat u studentů všech úrovní, stejně jako se znalostí názvů doložených v textech pro úroveň A2, popř. pro A2 a některou z vyšších úrovní: *elektrárna, kovárna, posluchárna, tiskárna, vinárna* (31 %).



Uvedená data ukazují, že většinu názvů s příponou *-árna* — 62 % — by si studenti měli osvojit hned v učebnicích pro úroveň A2

Naopak až na úrovni B1 nebo B2 se studenti seznámí s názvy *obrazárna, sušárna, teplárna, účtárna, kuřárna, mlékárna* (cca 38 %).

K tomu, aby studenti byli schopni tvořit názvy míst příponami *-iště* a *-árna* a dalšími slovtvornými prostředky, je důležitá znalost těch slov, která jsou považována za základová. Zjistili jsme, že v korpusech Učebko a v materiálech MŠMT některá z nich nejsou doložena vůbec, nebo jen ve velmi nízkém počtu. Překvapivě ani v korpusech Učebko, ani v materiálech MŠMT není doloženo sloveso *směnit* — základové slovo pro název *směnárna*, nenajdeme v nich ani slovo *směna* ve významu „výměna“. Slovo *směnárna* přitom rozhodně patří k těm, která by studenti měli umět vytvořit, protože se s ním běžně setkávají. Lektor by tedy měl na tuto situaci pamatovat a význam slov *směnit*, popř. *směna* studentům vysvětlit. Doložena nejsou slovesa *sušit* a *kovat* jako základová slova pro názvy *sušárna* a *kovárna*. Podle našeho názoru však ani slova *kovárna* a *sušárna*, ani jejich slova základová v současné době student-cizinec na rozdíl od slova *směnárna* pro komunikaci nezbytně nepotřebuje. Záleží tedy na lektorech, zda studenty nejen se slovy *kovárna* a *sušárna*, která jsou v učebních textech, ale též se slovesy *kovat* a *sušit* seznámí. Pouze ojedinele (v textech úrovně A2 z korpusů Učebko) je doloženo sloveso *sídlit* — základové slovo pro *sídlště*. Nabízí se tedy možnost považovat za základové slovo pro *sídlště* substantivum *sídlo*, doložené ve všech zdrojích. Ve velmi malém množství jsou doložena slovesa *hledět* — základové slovo pro *hledišťe* (2x v korpusech Učebko až v textech úrovně B2, 2x v materiálech MŠMT v textech vysvětlujících význam tohoto slova, nikoliv v textech reálné komunikace), *klouzat* — základové slovo pro *kluzišťe* (v korpusech Učebko ojedinele v A2 a zřídka v B1, v materiálech MŠMT ojedinele v B2), *přístav* — základové slovo pro *přístaviště* (v korpusech Učebko pouze několik dokladů v B2, v materiálech MŠMT vůbec), sloveso *tisknout* — základové slovo pro název *tiskárna*. Zejména slova *přístaviště* a *tiskárna* zřejmě studenti budou používat, lektor by tedy měl význam základových slov vysvětlit i přes jejich málo frekventovaný výskyt.

U slova *tiskárna* v současné době musíme počítat s dvěma významy: místo (dílna, firma...), kde se tisknou knihy, a prostředek/přístroj k tisku připojený k počítači nebo jinému zařízení; v korpusech i v materiálech MŠMT je doložen pouze jako název prostředku. Lektor by tedy měl studenty upozornit i na užívání slova *tiskárna* jako názvu místa.

Na případy nedoloženosti nebo nízké doloženosti základových slov členové řešitelského týmu pamatují a upozorní na ně i ve webové aplikaci, aby tak upozornili lektory na to, která slova je třeba studentům vysvětlit. Stejně tak budou studenti i lektoři vždy upozorněni na víceznačná slova, kterých je i mezi odvozenými slovy více.

4. SHRNUTÍ A BUDOUCÍ PRÁCE

CZEDD a SLAN jsou univerzitními projekty realizovanými na Masarykově univerzitě v Brně, které mají podobu online aplikací. Oba projekty nabízí nový pohled na práci s češtinou jakožto cizím jazykem, kdy cílí pozornost zejm. na slovtvornou formu odvozených slov, která tvoří minimálně 75 % celé slovní zásoby. Cílem projektů je na-



bídnout lektorům dostatečné informace pro uchopení slovtvorného systému češtiny a ten pak postupně přibližovat studentům-cizincům. Smyslem je tedy začlenit slovtvorbu jako běžnou součást každé lekce (upozorňovat studenty-cizince na slovtvornou příbuznost mezi slovy a strukturou odvozených slov), nikoliv jí věnovat speciální pozornost v předem určených lekcích. Popisy v obou aplikacích jsou oproštěny od jakékoliv lingvistické terminologie, a proto mohou být využity i studenty-cizinci zejm. při samostudiu. Výchozím materiálem se pro oba projekty stal specializovaný korpus, který byl pro tyto potřeby vybudován z učebnic češtiny pro cizince. Díky korpusovým datům obsahují obě aplikace jen ta slova, se kterými se studenti-cizinci skutečně setkávají.

V další práci na projektu budou postupně zpracovány další významové skupiny, jejichž slova budou roztríděna podle slovtvorných prostředků a opatřena informacemi relevantními pro jejich fungování ve webové aplikaci. Počítáme s nutností věnovat pozornost dalším jevům, které jsou při tvoření slov důležité pro cizince studujícího češtinu: popisu alternací samohláskových i souhláskových, variantám sufixů a upozornění na polyfunkčnost řady sufixů v češtině. Náročným úkolem pak bude srovnávání slovtvorné struktury a slovtvorných prostředků u českých utvořených slov a jejich ruských významových ekvivalentů. Na základě dosavadních ojedinělých zkušeností můžeme očekávat, že u slov významově ekvivalentních často nebudou ekvivalentní slovtvorné vlastnosti.

BIBLIOGRAFIE:

- Adamovičová, J. a kol. (2005): *Čeština jako cizí jazyk. Úroveň B2*. Praha: MŠMT.
- Čadská, M. a kol. (2005): *Čeština jako cizí jazyk. Úroveň A2*. Praha: MŠMT.
- Daneš, F. — Dokulil, M. — Kuchař, J. (1967): *Tvoření slov v češtině 2. Odvozování podstatných jmen*. Praha: Academia.
- Dokulil, M. (1962): *Tvoření slov v češtině 1. Teorie odvozování slov*. Praha: ČSAV.
- Hádková, M. — Línek, J. — Vlasáková, K. (2005): *Čeština jako cizí jazyk. Úroveň A1*. Praha: MŠMT.
- SERR (2001): *Společný evropský referenční rámec pro jazyky: Jak se učíme jazykům, jak je vyučujeme a jak v jazycích hodnotíme*. Olomouc: Univerzita Palackého.
- Jakubíček, M. — Kovář, V. — Šmerk, P. (2011): *Czech Morphological Tagset Revisited*, in: Horák, A. — Rychlý, P. (eds.), *Proceedings of Recent Advances in Slavonic Natural Languages Processing*, Brno: Tribun EU, s. 29–42.
- Kilgariff, A. a kol.: *Sketch Engine [Computer Software and Information Resource]*. [online] [cit. 2021-10-26]. Dostupné z [www: <http://www.sketchengine.co.uk>](http://www.sketchengine.co.uk).
- Křen, M. — Cvrček, V. — Čapka, T. — Čermáková, A. — Hnátková, M. — Chlumská, L. — Jelínek, T. — Kovaříková, D. — Petkevič, V. — Procházka, P. — Skoumalová, H. — Škrabal, M. — Truneček, P. — Vondříčka, P. — Zasina, A. (2019): *Korpus SYN, verze 8 z 12*. Ústav Českého národního korpusu FF UK, Praha. Dostupný z [www: <http://www.korpus.cz>](http://www.korpus.cz).
- Malá, R. (2021): *Sémantické zpracování slovesných prefixů pro potřeby češtiny jako cizího jazyka*. Magisterská diplomová práce. Brno. Masarykova univerzita.
- Ološtiak, M. (2015): *Kvalitatívne a kvantitatívne aspekty tvorenia slov v slovenčine*. Prešov: Filozofická fakulta Prešovskej univerzity v Prešove.
- Osolsobě, K. (2017): *Slovtvorný typ*, in: Karlík, P. — Nekula, M. — Pleskalová, J. (eds.), *CzechEncy — Nový encyklopedický slovník*



- češtiny. [online] [cit. 2021-10-26]. Dostupné z [www: <https://www.czechency.org/slovník/SLOVOTVORN%C3%9D%20TYP>](https://www.czechency.org/slovník/SLOVOTVORN%C3%9D%20TYP).
- Rejzek, J. (2012): *Etymologický slovník jazyka českého*. Praha: LEDA.
- Havránek, B. (ed.) (1989): *Slovník spisovného jazyka českého I–VIII*. Praha: Academia. [online]. Dostupné z [www: <https://ssjc.ujc.cas.cz/>](https://ssjc.ujc.cas.cz/).
- Sokolová, A. — Válková, A. (v tisku): *Slovtvorný analyzátor pro výuku češtiny pro cizince (SLAN) — didaktická pomůcka s korpusově podloženými daty*.
- Šimandl, J. a kol. (2017): *Slovník afixů užívaných v češtině*. Praha: Karolinum. [online]. Dostupné z [www: <http://www.slovníkafixu.cz/cil/>](http://www.slovníkafixu.cz/cil/).
- Šmerk, P. (2008): *K morfológické desambiguaci češtiny*. Advanced Master's thesis. Masaryk University, Faculty of Informatics [online] [cit. 2021-10-26]. Dostupné z [www: <https://is.muni.cz/auth/th/wteg5/teze.pdf>](https://is.muni.cz/auth/th/wteg5/teze.pdf).
- Štícha, F. a kol. (2018): *Velká akademická gramatika spisovné češtiny I. Morfologie: Druhy slov / Tvoření slov*. Praha: Academia.
- Válková, A. (v tisku): *Building Czech Textbook Corpora (UcebKo) for Word-formation Research of Czech as a Second Language*. Bratislava.
- Válková, A. (2021a): *Učebko*. Brno. [online] [cit. 2021-10-26]. Dostupné z [www: <https://ske.fi.muni.cz/#dashboard?corpname=user%2Fadriana%2Fucebko>](https://ske.fi.muni.cz/#dashboard?corpname=user%2Fadriana%2Fucebko).
- Válková, A. (2021b): *Učebko A2*. Brno: Masarykova univerzita.
- Válková, A. (2021c): *Učebko B1*. Brno: Masarykova univerzita.
- Válková, A. (2021d): *Učebko B2*. Brno: Masarykova univerzita.