

Słowiańskie przyjemności 3: Wielojęzyczność

pod red. Elżbiety Solak, Barbary Popiołek i Bojany Todorović



Adrian Jan Zasina

Univerzita Karlova

Korpusový přístup jako řešení jazykových
úskalí v osvojování češtiny jako cizího
jazyka: případ předpon s- a z-

Keywords

data-driven learning, foreign language teaching, language corpus, corpus linguistics, corpus instruments, Czech language, didactics, prefixes *s-* and *z-*

Abstrakt:

Studenti češtiny jako cizího jazyka poměrně často narážejí na problém s rozlišováním v psaní předpon *s-* a *z-*, jež v některých případech výrazně mění význam (srov. *směna* × *změna*). Této tematice věnují pozornost nejen učebnice češtiny pro cizince, ale i gramatiky a pravopisné příručky, neboť s tímto problémem se často potýkají i rodilí mluvčí. Tento příspěvek na zmíněnou problematiku nahlíží z hlediska studenta-cizince a učitele češtiny pro cizince, pro které korpusová data představují nový způsob řešení těchto problémů. Cílem této práce je představit korpusově založený přístup při řešení dané problematiky a také navrhnout způsob jejího procvičování ve formě konkrétních cvičení. Pozornost je věnována i možnostem automatického extrahování párů slov z korpusu s předponou *s-* a *z-* pomocí nástroje Morfio.



1. Úvod

Zapojení počítačů a nových technologií do různých oblastí života je v dnešní době zcela běžné, a není tedy divu, že se projevuje i ve výuce cizích jazyků, jako je např. e-learning, blended learning nebo vzdálená výuka v podobě webinářů či výuky po Skypu. Do metod jazykového vzdělávání tak v poslední době proniká i korpusová lingvistika (Reppen, 2010; Breyer, 2011; Thomas, 2016) využívající rozsáhlé soubory elektronicky uložených dat ke zkoumání přirozené řeči. Začlenění jazykových korpusů do běžné jazykové výuky může být řešením pro obtížné jevy, jež nejsou systematicky popsány z pohledu studenta-cizince, který, na rozdíl od rodilého mluvčího, jazyk vnímá odlišným způsobem. Korpusový přístup jako doplněk běžné výuky poskytuje možnost prozkoumat mj. přirozený kontext slov, typické kolokace, význam či stylové rozdíly v použití synonym.

Tento příspěvek si zaprvé klade za cíl seznámit čtenáře s korpusovým přístupem ve výuce cizího jazyka z teoretického hlediska, za druhé ukázat praktický přístup k řešení potíží s předponami *s-* a *z-* v češtině. Práce je rozdělena do tří částí. První, teoretická část shrnuje možnosti využití korpusů ve výuce a s tím související metody data-driven learning (pododdíl 1.1.); druhá část popisuje použitá data, nástroje a metodologii zpracování korpusových výsledků (oddíl 2.); třetí část nabízí cvičení pro přímou práci studentů s korpusem (oddíl 3.). V závěru jsou doporučeny další možné přístupy při řešení jazykových úskalí studentů češtiny pro cizince.

1.1. Úvod do problematiky data-driven learning (DDL)¹

Průkopníkem v oblasti metody data-driven learning (*daty řízená výuka*) je Tim Johns, jenž ve své studii zdůrazňuje, že *úkolem studenta je „objevovat“ cizí jazyk a úkolem učitele je zaručit kontext, ve kterém student bude schopen*

¹ Tento pododdíl se z velké části překrývá s polsky psanou pasáží publikovanou Zasinou (2018: 111–112) a částečně je jejím překladem.

*rozvinout strategii pro toto objevování*² (1991: 1). Johns (ibid.: 1–4) tím míní, že se u studenta vyvine schopnost „učit se, jak se učit“ (*learn how to learn*), a také zdůrazňuje, že student je v tomto případě zároveň vědcem, kterého korpusová data vedou k závěrům a popisu gramatických pravidel. Na rozdíl od běžných metod ve výuce je to přístup induktivní, který se nejdříve zakládá na korpusových datech, díky nimž student kategorizuje daný jev a generalizuje gramatická pravidla. Učitel přesně neví, jaké jevy a pravidla student objeví, proto je učitelovou rolí především být koordinátorem studentova výzkumu. Učitel již není autoritou, ale spíše partnerem. Neznamená to však, že tato pozice ubírá na učitelově významu, naopak zapojuje učitele do společné práce se studentem, kdy oba čelí jazykovým otázkám společně a snaží se přijít na vhodné řešení. Student díky této svobodě zvyšuje svou samostatnost a učitel pouze dohlíží na studentovy postupy, aby ho v případě obtíží nasměroval. Tento přístup umožňuje studentům objevovat nenásilnou formou nové věci, anebo přistupovat ke známým faktům z jiného úhlu pohledu, který příznivě působí na pochopení gramatických pravidel nerodilým mluvčím. Práce bývá koncipována tak, že učitel zadá studentovi problematickou úlohu, při jejímž řešení student sám objeví zákonitosti, aniž by mu byly předem vysvětleny učitelem³. DDL si tak zakládá na gramatických pravidlech objevených studentem, která jsou založena na autentickém jazyce. Díky tomu má student možnost skutečně prohloubit své jazykové znalosti a osamostatnit se ve výuce, což pozitivně ovlivňuje jeho studijní výsledky (Bernardini, 2004; Lewandowska, 2014).

DDL se tradičně dělí na dvě odvětví podle metody práce (Boulton, 2012; Vališová, 2012: 37), která je buď přímá (tzv. *hard DDL* nebo *hands-on*), kdy studenti pracují přímo s korpusovým vyhledávačem, anebo nepřímá (*soft DDL* nebo *hands-off*), kdy studenti řeší úkoly připravené na základě korpusových dat, např. ve formě vytištěných konkordancí (tj. příkladů hledaného slova v kontextu). Pod pojmem korpusy ve výuce si lze představit tři různé způsoby využití (Geist & Hahn, 2012: 124): a) výuka o korpusech (v rámci lingvistických kurzů na univerzitě); b) korpusový přístup ve výuce jazyků; c) výuka o využití korpusů. Nicméně, jak ukázaly výzkumy (Römer, 2009: 87; Frankenberg-Garcia, 2012: 37), i přes snahu korpusových lingvistů nejsou korpusy mezi učiteli středních škol

² [T]he task of the learner is to “discover” the foreign language, and (...) the task of the language teacher is to provide a context in which the learner can develop strategies for discovery (vlastní překlad).

³ Např. učitel nevysvětlí studentovi, jak se tvoří minulý čas, nýbrž student sám bude na základě pozorování korpusových dat schopen popsat pravidla tvoření, a tak si je lépe vštípit.

příliš populární a v akademickém prostředí jsou propagovány hlavně korpusovými specialisty. Frankenberg-Garcia (2012) zdůrazňuje nutnost zapojení korpusů do každodenní výuky cizích jazyků prostřednictvím přípravy učitelů k práci s korpusem v rámci jejich pedagogického studia, a také, jak uvádí Römer (2009), je v tomto ohledu nepostradatelná pomoc korpusových lingvistů, kteří by měli vyjít vstříc očekáváním učitelů a vytvářet nové, uživatelsky přívětivé nástroje nebo sestavovat ukázková cvičení pro studenty.

Pod vlivem rozvoje korpusové lingvistiky mimo anglicky mluvící země se zájem o začlenění korpusových metod do výuky cizích jazyků objevuje i v České republice a Polsku (srov. Zasina, 2018; Zasina 2019). Je pochopitelné, že vzhledem k rozšířenosti angličtiny a její výuky se pozornost v obou zemích nejprve zaměřila na ni a teprve poté rozšířila na národní jazyk.

2. Data a metodologie

Tato část se věnuje popisu dat, konkrétně pak využití korpusů a korpusových nástrojů k popisu problému. Pozornost je dále věnována metodologickému postupu při zpracování korpusových výsledků.

2.1. Data

Tento výzkum nejprve využívá data ze žákovského korpusu CzeSL-SGT (Šebesta et al., 2014), jenž zahrnuje texty nerodilých mluvčích češtiny a obsahuje automaticky provedenou anotaci. Velikost korpusu činí 1 147 477 pozic a jeho část obsahující texty slovanských mluvčích, na kterou jsme se zaměřili v našem výzkumu, zahrnuje 769 126 pozic. Korpus je kromě standardní morfologické anotace (Jelínek, 2008; Jelínek & Petkevič, 2011) vybaven i chybovou anotací zohledňující nejběžnější chyby studentů češtiny jako cizího jazyka⁴ (ČCJ). Žákovský korpus nám slouží jako východisko pro identifikaci a verifikaci obtíží s použitím předpon *s-* a *z-*. Následně je využit nástroj pro odhadování rozsahu produktivity slovotvorných modelů Morfio (Cvrček & Vondříčka, 2013). Tato aplikace je schopna vyhledávat v korpusu tvary se společnou a odlišnou částí, což umožnilo identifikaci párů s předponami *s-* a *z-* o stejném kořenu. V poslední řadě k určení významu na základě přirozeného kontextu analyzujeme data z nejnovějšího reprezentativního synchronního korpusu psané češtiny

⁴ <http://utkl.ff.cuni.cz/~rosen/public/SeznamAutoChybR0R1.html>.

SYN2015 (Křen et al., 2015; Křen et al., 2016) obsahujícího 120 748 715 pozic. Tento korpus je zároveň zdrojem pro sestavení cvičení.

2.2. Metodologie

Výchozím bodem byla analýza chyb v žákovském korpusu CzeSL-SGT u studentů, jejichž mateřštinou je nějaký slovanský jazyk. Žákovský korpus byl použit pro ověření, zda předpony *s-* a *z-* dělají studentům ČCJ potíže. Tento typ chyb⁵ nepatří k nejfrekventovanějším, nicméně z hlediska osvojování ČCJ se jeví důležitý, neboť změnou pravopisu může docházet k výrazné změně významu (viz dále). V korpusu lze tento jev sledovat dvěma různými způsoby.

Zprv, díky automatické lemmatizaci jak chybných tvarů (word), tak informací o tvarech správných (word1) je možné vyhledat nevhodně zvolené předpony pomocí dotazu [word="(?)z.+" & word1="(?)s.+" & tag1="[NV].*"] [word="(?)s.+" & word1="(?)z.+" & tag1="[NV].*"]. Tímto dotazem získáme slova s chybnou předponou *s-* a *z-* (např. **spůsob*, **zkončit*)⁶, pro něž neexistuje reálně v jazyce správný tvar s toutéž předponou (např. v češtině neexistují oba tvary jako **spůsob* a *způsob*). Na tomto základě byl vytvořen seznam chybných substantiv (Tabulka 1) a sloves (Tabulka 2).

Tabulka 1. Seznam chybně napsaných substantiv v korpusu CZeSL-SGT⁷

	Lemma	Frekvence výskytu
1.	zkouška	52
2.	způsob	37
3.	zkušenost	21
4.	ztráta, zpěvačka	3
5.	zpěvák, schůzka, skončení, stmívání, zpoždění, shromáždění, snížení, stravování	1
	celkem	124

⁵ Označen v korpusu jako spodoba znělosti *Voiced0* – chybně neznělá a *Voiced1* – chybně znělá.

⁶ Případy, ve kterých se buď nejednalo o předponu (např. *zub*, *sál*, *znalost*), anebo o předponu ze synchronního hlediska (např. *zboží*), byly z analýzy vyloučeny.

⁷ Tabulka neobsahuje chybný tvar vlastního jména Salzburg, který měl jeden výskyt.

Tabulka 2. Seznam chybně napsaných verb v korpusu CZeSL-SGT

	Lemma	Frekvence výskytu
1.	skončit	48
2.	zpívat	35
3.	ztratit	21 ⁸
4.	zkusit	18
5.	strávit	5
6.	zeptat, zkoušet, zkontrolovat, sbalit, ztrácet	2
7.	zbláznit, skrývat, zdát, snídat, splácet, zklamat, zplodit, sejít, svěřit, zvládnout, zhoršit, zvyknout, stlačovat	1
celkem		150

Dohromady bylo získáno 274 výsledků, 124 pro substantiva a 150 pro verba. Pokud jde o různé typy slov, našli jsme 13 typů pro podstatná jména a 23 pro slovesa, takže celkem 42 problematická slova. Podíváme-li se na nejfrekventovanější příklady, ve většině případů se jedná o předponu *z-* nebo *s-* stojící před neznělým konsonantem, který v důsledku artikulace působí na prefix, jež vyslovujeme nezněle. Lze to považovat za jeden z nejdůležitějších důvodů obtíží v této oblasti.

Za druhé, pomocí nástroje Morfio a dat z referenčního korpusu SYN2015 jsme sestavili potenciální seznam substantiv (Tabulka 3) a verb (Tabulka 4), pro něž (na rozdíl od Tabulek 1 a 2) reálně mohou v jazyce existovat správné tvary s předponou *s-* a *z-* (srov. *směna* a *změna*)⁹. Následně jsme v žákovském korpusu vyhledali slova z tohoto seznamu a chyby jsme určili ručně. Ve výsledku jsme našli pouze 5 typů takovýchto výrazů (celkem 15 výskytů): substantiva *správa/zpráva* a *směna/změna* a verba *směnit/změnit*, *svrhnout/zvrhnout*, *shlédnout/zhlédnout*:

- (1) Ted už můžu pochopit *spravy*, číst noviny a mluvit v bance, kde mě mohou rozumět.
- (2) Taký zkusím *směnit* svůj názor na život a na vztah s lidmi.
- (3) Když mi bylo 11 jsem *směnila* nazor, protože jsem rozuměla že žádná profese jako bisneslady ne existuje.

⁸ Ve výsledcích se chybně objevilo jedno adjektivum, které zde nezapočítáváme.

⁹ Kvůli tomu, že oba tvary jsou v jazyce správné, automatická chybová anotace reálné chyby nedetekuje a tímto přiřazuje chybný tvar i tam, kde bychom očekávali tvar správný (word1), proto nebylo možné najít tyto tvary automaticky v první kroku.

- (4) Oni budou chtít *zvrhnout* vládu.
 (5) Také řídící by měli komplikovaný život, když by neměli možnost *schlédnout* nebo poslechnout dopravní zpravodajství.

Tabulka 3. Seznam párů substantiv s předponami *s-* a *z-* v korpusu SYN2015¹⁰

	<i>s-</i>	<i>z-</i>
1.	sběh	zběh
2.	shlédnutí	zhlédnutí
3.	sjednání	zjednání
4.	sjednávání	zjednávání
5.	slézání	zlézání
6.	směna	změna
7.	správa	zpráva
8.	svážení	zvážení
9.	svolání	zvolání
10.	svolení	zvolení

Tabulka 4. Seznam párů verb s předponami *s-* a *z-* v korpusu SYN2015¹¹

	<i>s-</i>	<i>z-</i>
1.	sběhnout	zběhnout
2.	shlédnout	zhlédnout
3.	shlížet	zhlížet
4.	sjednat	zjednat
5.	sjednávat	zjednávat
6.	skapat	zkapat
7.	skopat	zkopat
10.	skosit	zkosit
11.	slézat	zlézat
12.	slézt	zlézt
13.	slíbat	zlíbat
14.	slíbit	zlíbit

¹⁰ Páry, ve kterých jeden z komponentů byl vlastním jménem, nejsou v tabulce zahrnuty.

¹¹ Tabulka nezahrnuje páry, jež jsou z hlediska sémantického stejné (srov. *skrápět* a *zkrápět*, *spráskat* a *zpráskat*, *schvátit* a *zchvátit*).

15.	slít	zlít
16.	smáčet	zmáčet
17.	smáčknout	zmáčknout
18.	směnit	změnit
19.	smotat	zmotat
20.	spravit	zpravit
21.	spravovat	zpravovat
22.	stěžovat	ztěžovat
23.	stlouci	ztlouci
24.	strhat	ztrhat
25.	sužovat	zužovat
26.	svažovat	zvažovat
27.	svolat	zvolat
28.	svolit	zvolit
29.	svrhávat	zvrhávat
30.	svrhnout	zvrhnout

I přesto, že jsme v žákovském korpusu našli málo příkladů tohoto typu chyb, patrně z důvodů úzce zaměřených témat studentských prací, stojí za to prozkoumat tento jev na datech korpusu SYN2015, kde si studenti-cizinci budou schopni v případě nouze ověřit význam slova a zvolit správnou variantu.

3. Interpretace korpusového materiálu

Seznamy získané na základě žákovského korpusu budou v této části analyzovány z hlediska fonologického a z hlediska jazykové úrovně podle Společného evropského referenčního rámce (SERR), ve které se vyskytlo nejvíce chyb. Následně ukážeme možnosti práce s korpusem SYN2015, ve kterém lze ověřovat význam slov, a nakonec nabídneme cvičení pro procvičování této látky přímo se studenty.

3.1. Analýza dat žákovského korpusu

Získané seznamy slov (Tabulka 1 a 2), jak jsme už zmínili, lze interpretovat z hlediska fonologického. Pro každou předponu se vyskytují tři případy: a) za

předponou následuje znělý konsonant, b) za předponou následuje vokál, c) za předponou následuje neznělý konsonant. Na tomto principu jsme substantiva a verba s předponami *s-* a *z-* rozdělili do jednotlivých skupin prezentovaných v Tabulkách 5 a 6 obsahujících chybné tvary nalezené v žákovském korpusu a tvary z hlediska jazykové normy správné.

Tabulka 5. Seznam slov s předponou *s-* z hlediska fonologického

Správný tvar		Chybný tvar
s- + znělý konsonant		
1.	shromáždění	zhromáždění
2.	sbalit	zbalit
3.	snížení	znižení
4.	svěřit	zvěřit
s- + vokál		
1.	sejít	zejít
s- + neznělý konsonant		
1.	schůzka	zchůzka
2.	stmívání	ztmívání
3.	strávit	ztrávit
4.	skončit	zkončit
5.	stravování	ztravování
6.	splácet	zplácet
7.	stlačovat	ztlačovat
8.	skrývat	zkrývat

Tabulka 6. Seznam slov s předponou *z-* z hlediska fonologického

Správný tvar		Chybný tvar
z- + znělý konsonant		
1.	zbláznit	sbláznit
2.	zdát	sdát
3.	zvládnout	svládnout
4.	zhoršit	shoršit
5.	zvyknout	svyknout

z- + vokál		
1.	zeptat	septat
z- + neznělý konsonant		
1.	zkouška	skouška
2.	způsob	spůsob
3.	zkušenost	skušenost
4.	ztráta	stráta
5.	zpěvačka	spěvačka
6.	zpěvák	spěvák
7.	zpoždění	spoždění
8.	zpívat	spívat
9.	ztratit	stratit
10.	zkusit	skusit
11.	zkoušet	skoušet
12.	zkontrolovat	skontrolovat
13.	ztrácet	strácet
14.	zklamat	sklamat
15.	zplodit	splodit

Jak tomu nasvědčují informace z českých gramatik (Havránek & Jedlička, 1960; Karlík et al., 1997; Cvrček, 2010), při psaní zkoumaných předpon hraje roli spodoba znělosti, což vysvětluje chyby u předpon *s-* a *z-*, po kterých následují neznělé souhlásky. Jsou také početnější oproti ostatním skupinám. Nicméně některé chybné tvary mohou být výsledkem interference z mateřského jazyka. Například v případě tvaru **spůsob* lze na základě metainformací určit, že chyby se vyskytly u rusky, polsky a ukrajinsky mluvících studentů (rus. *cnoco6*, pol. *sposób*, ukr. *cnoci6*).

Další informaci, kterou lze získat na základě dat žákovského korpusu, je úroveň studentů, u nichž se vyskytla chyba. Tyto výsledky představuje Tabulka 7.

Tabulka 7. Úroveň studentů podle SERR, u nichž se vyskytla chyba

úroveň	frekvence	ipm
A2	115	461,23
B2	38	438,43
B1	55	395,80
A1	90	369,31
C1	2	114,32

Jelikož počet textů pro jednotlivé úrovně je odlišný, pro srovnání jsme použili relativizovanou míru ipm (*instances per milion*), která měří výskyt jednotlivého jevu v hypotetickém milionovém korpusu. Je jasné vidět, že největší problémy nastávají pro úroveň A2, jelikož dané téma je tradičně probíráno se studenty až na vyšší úrovni (srov. Adamović & Hrdlička, 2016; Cvejnová, 2016). Toto korpusové zjištění naznačuje, že by této problematice ve výuce češtiny pro cizince měla být věnována pozornost mnohem dříve. Především by měly být zdůrazněny rozdíly mezi češtinou a mateřským jazykem, aby se mohlo předcházet případným interferenčním chybám.

3.2. Analýza dat korpusu SYN2015

Pro samostatnou práci jsou pro studenty vhodné korpusy textů rodilých mluvčích, jakým je např. SYN2015. Množství kontextů poskytovaných rozhraním KonText (Machálek & Křen, 2013) umožňuje studentům nahlédnout do přirozeného jazyka, ve kterém mohou zkoumat význam slov anebo si ověřit jejich psanou podobu.

Dříve zjištěný seznam problematických slov s předponami *s-* a *z-* majících stejný kořen (substantiva: *s/zpráva* a *s/změna* a verba: *s/změnit*, *s/zvrhnout*, *s/zhlédnout*) lze v korpusu jednoduše ověřit a zjistit správnou psanou podobu pro vhodný kontext. Student, který si např. není vědom rozdílu mezi slovy *směna* a *změna*, může zadat do korpusu dotaz pro jednotlivá slova a na základě konkordance či kolokací určit význam a spojit si ho se správnou předponou.



Obrázek 1. Konkordance lexému *směna*

mi líto . Zkuste se ale zeptat portýra . Měl	směnu	i v sobotu odpoledne . " Birkin ho vyhledal a
nealokují zdroje racionálně , zkrachují . V tomto pojetí se	směna	práce za mzdu neliší od směny jakéhokoli jiného zboží .
hrst . Táta většinu dní pracoval v restauraci na desetihodinových	směnách	, které začínaly v šest ráno . „ Volala jsi
a s ním spojeného vychloubání , od slasti z povzbuzující	směny	mezi získáváním a spotřebou objektů – , jež postihuje spíše
Willie McCovey po boku Willieho Mayse . Na začátku osmé	směny	McCovey stáhl dva body petelící , která přeplula místa k
před třemi týdny v Pittsburghu , když dotáhl do sedmé	směny	pětibodový náskok a promrhal ho , než se Mets dokázal

Obrázek 2. Konkordance lexému *změna*

2012 byl podle ní pro každého náročný , odehrály se	změny	jak v partnerství , tak v pracovní oblasti , mnoho
v některých jednotlivých případech se dokončují revize , popř.	změny	norem . Jak bude v rámci harmonizace Eurokódů naloženo s
vědomí , které jsou způsobeny náhlou , vůlí neovlivnitelnou ,	změnou	činností mozku . V období mezi záchvaty pacient bez patologického
a Severní Ameriky . Jeho zánik patrně souvisel s rozsáhlými	změnami	životního prostředí na počátku svrchního miocénu , kdy v důsledku
prošla mnoha změnami . Střídali se hudebníci , nástroje a	změn	doznal také styl . Něco však zůstalo stejné , a
utvořit , jestliže výše uvedené skutečnosti uvedeme do souvislosti se	změnami	velikosti mozku u rodové linie hominidů , jak znázorňuje obrázek
vodidla objevit i na silnicích Moravskoslezského kraje . Nejpodstatnější	změnou	je radikální snížení hmotnosti systému , kdy nový systém je

Na základě pozorování korpusových dat lze zjistit, že lexém *směna* souvisí s prací, kde se střídá pracovní doba, anebo se vyskytuje ve významu výměny nějakého zboží. Zatímco kontexty lexému *změna* naznačují, že se jedná o přechod do jiného stavu, vystřídání něčeho nebo nabytí nových vlastností.¹²

Tímto způsobem můžeme také přistupovat ke slově s předponou *s-* a *z-*, u nichž studenti nejsou schopni určit správnou předponu. Stačí tak, např. pro sloveso *skončit*, položit dotaz do korpusu¹³ s použitím seznamu v podobě [sz]končit. Díky tomu v souhrnu výsledků získáme nejfrekventovanější vzorek obou slov, jak jsou v jazyce používány.¹⁴ Tento přístup umožní studentům osamostatnit se ve svém studiu cizího jazyka a motivuje je k zamyšlení nad fungováním jazyka, proto tento typ cvičení může být použit jako doplněk běžné jazykové výuky.¹⁵

¹² Ukázkové cvičení tohoto typu se nachází v Příloze 1.

¹³ Typ dotazu je třeba určit jako *lemma* nebo *slovní tvar*.

¹⁴ Samozřejmě může dojít k tomu, že najdeme obě předpony, z nichž jedna, nesprávná z hlediska normativního, bude méně zastoupena. Do korpusu vstupují texty rodilých mluvčích, kteří se rovněž dopouštějí chyb. Nicméně tyto chyby jsou zanedbatelné ve srovnání s četností správného tvaru.

¹⁵ Ukázkové cvičení tohoto typu se nachází v Příloze 2.

4. Závěr

Korpusový přístup ve výuce češtiny jako cizího jazyka je poměrně novou metodou. Z tohoto důvodu ještě není systematicky popsáno, jak ke korpusovým datům přistupovat. Tento příspěvek ukazuje na konkrétním příkladu problematických předpon *s-* a *z-* u substantiv a sloves, jak lze zacházet s korpusovým materiálem a jeho interpretací. Pomocí různých nástrojů a různých korpusů lze výsledky jednoduše setřídit a interpretovat. Následně na základě těchto dat je možné nabídnout jazyková cvičení, jejichž cílem je prohloubení znalostí studentů v této oblasti a poskytnutí materiálů k samostudiu.

Tato práce se zaměřuje jen na úzkou část spektra možností korpusového přístupu ve výuce češtiny. Nicméně prezentované postupy mohou být pro cizince studující češtinu podnětem a návodem pro další zkoumání problematických oblastí. Jsou také příkladovým materiálem, který může posloužit jak učitelům, tak i studentům v další práci a zároveň je motivovat k dalšímu objevování zákonitostí českého jazyka z úhlu pohledu studenta-cizince a jeho učitele.

Literatura

- Adamović A. & Hrdlička M. (2016), *Basic Czech III* (2. vydání). Praha: Karolinum.
- Bernardini S. (2004), *Corpora in the classroom: An overview and some reflections on future developments* [in:] *How to Use Corpora in Language Teaching*, J. McH. Sinclair (red.), 15–36. Amsterdam/Philadelphia: John Benjamins.
- Boulton A. (2012), *Hands-on / hands-off: Alternative approaches to data-driven learning* [in:] *Input, Process and Product: Developments in Teaching and Language Corpora*, J. Thomas & A. Boulton (red.), 152–168. Brno: Masaryk University Press.
- Breyer Y. (2011), *Corpora in Language Teaching and Learning: Potential, Evaluation, Challenges*. Frankfurt am Main: Peter Lang.
- Cvejnová J. (2016), *Česky, prosím III*. Praha: Karolinum.
- Cvrček V. (2010), *Mluvnice současné češtiny*. Praha: Karolinum.
- Cvrček V. & Vondříčka P. (2013), *Morfio*. Praha: Univerzita Karlova. Dostupné z <<http://morfio.korpus.cz>>.
- Frankenberg-Garcia A. (2012), *Integrating corpora with everyday language teaching* [in:] *Input, Process and Product: Developments in Teaching and Language Corpora*, J. Thomas & A. Boulton (red.), 36–53. Brno: Masaryk University Press.

- Geist M. & Hahn A. (2012), *Using a corpus for written production: A classroom study* [in:] *Input, Process and Product: Developments in Teaching and Language Corpora*, J. Thomas & A. Boulton (red.), 123–135. Brno: Masaryk University Press.
- Havránek B. & Jedlička A. (1960), *Česká mluvnice* (2. uprav. rozšíř. vyd.). Praha: Státní pedagogické nakladatelství.
- Jelínek T. (2008), *Nové značkování v Českém národním korpusu*. „Naše řeč“ 91 (1): 13–20.
- Jelínek T. & Petkevič V. (2011), *Systém jazykového značkování současné psané češtiny* [in:] *Korpusová lingvistika Praha 2011, sv. 3: Gramatika a značkování korpusů*, V. Petkevič & A. Rosen, (red.), 143–170. Praha: Nakladatelství Lidové noviny.
- Johns T. (1991), *Should you be persuaded: Two samples of data-driven learning materials*. „Classroom Concordancing: ELR Journal“ 4: 1–16.
- Karlík P., Nekula M. & Rusínová Z. (1997), *Příruční mluvnice češtiny* (vyd. 2., opravené). Praha: Nakladatelství Lidové noviny.
- Křen M., Cvrček V., Čapka T., Čermáková A., Hnátková M., Chlumská L., Jelínek T., Kovářiková D., Petkevič V., Procházka P., Skoumalová H., Škrabal M., Truneček P., Vondříčka P. & Zasina A. (2015), *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z WWW: <http://www.korpus.cz>
- Křen M., Cvrček V., Čapka T., Čermáková A., Hnátková M., Chlumská L., Jelínek T., Kovářiková D., Petkevič V., Procházka P., Skoumalová H., Škrabal M., Truneček P., Vondříčka P. & Zasina A. J. (2016), *SYN2015: Representative Corpus of Contemporary Written Czech* [in:] *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, N. Calzolari et al. (red.), 2522–2528. Portorož: ELRA.
- Lewandowska A. (2014), *Using corpus-based classroom activities to enhance learner autonomy*. „Konińskie Studia Językowe“ 2 (3): 237–255.
- Machálek T. & Křen M. (2013), *Query interface for diverse corpus types* [in:] *Natural Language Processing, Corpus Linguistics, E-learning*, K. Gajdošová & A. Žáková (red.), 166–173. Lüdenscheid: RAM Verlag.
- Reppen R. (2010), *Using corpora in the language classroom*. New York: Cambridge University Press.
- Römer U. (2009), *Corpus research and practice. What help do teachers need and what can we offer?* [in:] *Corpora and Language Teaching*, K. Aijmer (red.), 83–98. Amsterdam: John Benjamins.
- Šebesta K., Bedřichová Z., Šormová K., Štindlová B., Hrdlička M., Hrdličková T., Hana J., Petkevič V., Jelínek T., Škodová S., Poláčková M., Janeš P., Lundáková K., Skoumalová H., Sládek Š., Pierscieniak P., Toufarová D., Richter M., Straka M., Rosen A. (2014), *CzeSL-SGT: korpus češtiny nerodilých mluvčích s au-*

- tomaticky provedenou anotací, verze 2 z 28. 7. 2014. Praha: Ústav Českého národního korpusu FF UK. Dostupný z WWW: <http://www.korpus.cz>*
- Thomas J. (2016), *Discovering English with Sketch Engine: A Corpus-Based Approach to Language Exploration*. Brno: Versatile.
- Vališová P. (2012), *Data-driven learning a výuka češtiny jako cizího jazyka*. „CASALC Review“ 2: 22–39.
- Zasina A. J. (2018), *Korpusy językowe w nauczaniu języków obcych – metoda, narzędzia, praktyka* [in:] *Specyficzne potrzeby studentów szkół wyższych a nauczanie języków obcych: Tom II: Praktyczne narzędzia*, M. Jedynak (red.), 110–123. Wrocław: Uniwersytet Wrocławski.
- Zasina. A. J. (2019). *Podejście korpusowe w nauczaniu języka polskiego jako obcego na przykładzie rzeczownikowych alternacji ó:o* [in:] *Panorama glottodydaktyki polonistycznej. Wyzwania, pytania, kierunki*, K. Ziolo-Pużuk (red.), 181–199. Warszawa: Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego.

Příloha 1. Ukázkové cvičení *hands-on* na slova s předponami s- a z- mající stejný kořen.

Zadání:

Pomocí korpusových dat vysvětlete rozdíl ve významu uvedených slovních dvojic a dopište příkladovou větu.

a) směna × změna

.....

b) správa × zpráva

.....

c) svrhnout × zvrhnout

.....

d) shlédnout × zhlédnout

.....

Klíč:

- a) směna – střídání pracovní doby
změna – přechod do jiného stavu, vystřídání něčeho nebo nabytí nových vlastností
- b) správa – vedení; oprava
zpráva – informace
- c) svrhnout – shodit; zbavit vlády
zvrhnout – převrátit se; změnit se k horšímu
- d) shlédnout – podívat se směrem dolů
zhlédnout – spatřit



Příloha 2. Ukázkové cvičení *hands-on* na slova s předponou s- a z-, u nichž studenti nejsou schopni určit správnou předponu.

Zadání:

Pomocí korpusu SYN2015 ověřte, která z předpon s- nebo z- je správná pro jednotlivá slova (z nabídky vyberte typ dotazu: lemma).

- | | |
|---------------|---------------|
| a) [sz]balit | f) [sz]plácet |
| b) [sz]kouška | g) [sz]pěvák |
| c) [sz]chůzka | h) [sz]ejít |
| d) [sz]končit | ch) [sz]tráta |
| e) [sz]působ | i) [sz]horšit |

Klíč:

- | | |
|------------|------------|
| a) sbalit | f) splácet |
| b) zkouška | g) zpěvák |
| c) zboží | h) sejít |
| d) snažit | ch) ztráta |
| e) způsob | i) zhoršit |