



Pedagogická
fakulta
Faculty
of Education

Jihočeská univerzita
v Českých Budějovicích
University of South Bohemia
in České Budějovice

Jihočeská univerzita v Českých Budějovicích
Pedagogická fakulta
Katedra slovanských jazyků a literatur
Oddělení českého jazyka a literatury

Diplomová práce

Využití Českého národního korpusu ve výuce českého jazyka

Vypracovala: Bc. Žaneta Zemanová
Vedoucí práce: PhDr. Milena Nosková, Ph.D.

České Budějovice 2021

Poděkování

Ráda bych poděkovala vedoucí své práce PhDr. Mileně Noskové, Ph.D. za odborné vedení, cenné rady a ochotu při vypracování diplomové práce. Dále bych chtěla poděkovat PhDr. Marii Kopřivové, Ph.D. za odbornou spolupráci a pomoc při práci s Českým národním korpusem. Také velmi děkuji gymnáziu, zvláště třídním učitelům a žákům ze sekundy a tercie, kteří se ochotně podíleli na vyplňování pracovního listu.

Prohlášení

Prohlašuji, že svoji diplomovou práci jsem vypracovala samostatně, pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledky obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích, 7. července 2021

.....

Bc. Žaneta Zemanová

Anotace

Využití Českého národního korpusu ve výuce českého jazyka

Diplomová práce je zaměřena na využití Českého národního korpusu ve výuce českého jazyka na víceletém gymnáziu, především na nižším stupni. Práce se dělí na teoretickou a výzkumnou část. Cílem teoretické části je vymezit základní přehled o didaktice českého jazyka, korpusové lingvistice, Českém národním korpusu a lexikologii. Empirická část popisuje tvorbu pracovního listu zaměřeného na lexikologii, zvláště na frazeologii. Dále shrnuje výsledky žákovských prací. Při tvorbě pracovního listu byla použita data z psaného korpusu KonText a WaG.

Klíčová slova: didaktika českého jazyka, lexikologie, korpus, pracovní list.

Abstract

Use of the Czech national corpus in teaching the Czech language

The diploma thesis is focused on the use of the Czech National Corpus in the teaching of the Czech language at a grammar school, especially at lower-secondary school. The work is divided into theoretical and research part. The purpose of the theoretical part is to define a basic overview of the didactics of the Czech language, corpus linguistics, the Czech National Corpus and lexicology. The empirical part describes creation of a worksheet focused on lexicology, especially phraseology. It also summarizes the results of student work. Data from the written corpus KonText and WaG were used to create the worksheet.

Key Words: Czech language didactics, lexicology, corpus, worksheet.

Obsah

Úvod.....	10
Teoretická část	11
1 Didaktika českého jazyka	11
1.1 Obecná didaktika.....	11
1.2 Vymezení didaktiky českého jazyka	12
1.3 Jazykové cíle	13
1.4 Jazykový systém ve vyučování	15
1.5 Motivace.....	16
1.6 Vztah učitele a žáka	18
1.7 Vyučovací metody	19
2 Český národní korpus	22
2.1 Vznik ČNK.....	22
2.2 Členění ČNK.....	23
2.3 Reprezentativnost ČNK	24
2.4 František Čermák	25
2.5 Korpusové aplikace	26
2.6 Pro školy.....	29
2.7 Korpusy ve vyučování.....	31
3 Korpusová lingvistika.....	33
3.1 Korpus a jazyk.....	33
3.2 Korpusová lingvistika	34
3.3 Historie korpusové lingvistiky	36
3.3.1 Hlavní světové korpusy	36
3.4 Metodologie	37
3.4.1 Korpusová data	38
3.4.2 Zpracování dat	39
3.5 Typy korpusů	40
3.6 Přístup corpus-based a corpus driven.....	42
3.7 Reprezentativnost	43
4 Lexikologie.....	45
4.1 Předmět lexikologie	45
4.1.1 Metoda sémaziologická a onomaziologická	46
4.1.2 Sémantické vztahy	46
4.2 Slovo, lexém.....	48
4.3 Rozšiřování slovní zásoby.....	48
4.4 Frazeologie.....	50
4.4.1 Obecné vlastnosti.....	50
4.4.2 Idiom a frazém	52
4.4.3 Frazémy nevětné.....	53
4.4.4 Frazémy větné.....	54
5 Výzkumná část	56
5.1 Cíle a organizace práce	56
5.2 Metodika práce.....	56
5.3 Forma a metoda.....	58
5.4 Klíčové kompetence.....	58
6 Pracovní list.....	59
6.1 Tvorba pracovního listu	59
6.1.1 Přechylování pomocí přípon.....	59
6.1.2 Synonyma	63
6.1.3 Slovní spojení	66

6.1.4	Zkratky	71
6.1.5	Práce s textem, přejatá slova	72
6.1.6	Význam frazémů	75
6.1.7	Přirovnání	76
6.1.8	Doplňování číslovky do frazému	78
6.2	Vyhodnocení pracovních listů	79
6.2.1	První cvičení	80
6.2.2	Druhé cvičení	80
6.2.3	Třetí cvičení	81
6.2.4	Čtvrté cvičení	82
6.2.5	Páté cvičení	82
6.2.6	Šesté cvičení	83
6.2.7	Sedmé cvičení	84
6.2.8	Osmé cvičení	84
7	Závěr	86
8	Seznam literatury a zdrojů	88
9	Přílohy	93
9.1	Příloha č. 1	93
9.2	Příloha č. 2	96
9.3	Příloha č. 3	99

Seznam použitých značek

Poziční atributy:

col_lemma (collocation lemma): lemma víceslovné jednotky v podobě slovníkového hesla v základním tvaru (nominativ singuláru, infinitiv apod.);

col_type (collocation type): dvoupísmenná značka, jejíž první písmeno určuje druh víceslovné jednotky a druhé slouží k rozlišení hlavního (H) a závislého (Z) slova v ní;

M: přísloví, okřídlená rčení, citace apod.;

tag: gramatická informace automaticky přiřazená ke každému tvaru/wordu;

Regulární výrazy:

ampersand (&): a zároveň (and), platí všechny podmínky zároveň;

hvězdička (*): představuje libovolný počet (0 a více) opakování předchozího znaku nebo celku;

seznam ([]): představuje alternativu. Nabízí možnost vybrat jeden libovolný znak z těch, které jsou uvedeny v seznamu uvnitř hranatých závorek;

tečka (.): představuje jeden libovolný znak;

Morfologické značky:

A: 1. pozice, slovní druh, adjektivum;

B: 1. pozice, zkratka;

C: 1. pozice, numerál (číslovka nebo číselný výraz s číslicemi);

N: 1. pozice, slovní druh, substantivum;

V: 1. pozice, slovní druh, verbum;

A: 2. pozice, adjektivum obyčejné;

B: 2. pozice, tvar přítomného nebo budoucího času;

f: 2. pozice, infinitiv;

l: 2. pozice, číslovka základní 1-4, „půl“, ...; sto a tisíc v nesubstantivním skloňování;

N: 2. pozice, substantivum obyčejné;

r: 2. pozice, číslovka řadová;

F: 3. pozice, femininum (ženský rod);

M: 3. pozice, maskulinum animatum (rod mužský životný);

P: 4. pozice, plurál (množné číslo);

Značení v korpusu:

discussions: webové diskuse (pod články na zpravodajských serverech);

doc.t xtype_group: skupina textových typů;

Facebook: posty včetně reakcí (pouze do prosince 2020);

Filter: filtrování výsledku;

forums: samostatná webová fóra (nezávislá na zpravodajských portálech);

Instagram: k dispozici jsou data pouze v některých obdobích

i.p.m.: relativní frekvence;

p/n: pozitivní/ negativní filtr;

news: internetová žurnalistika (analyticko-investigativní, antisystémové weby, bulvární média, hlavní proud, market-driven média, názorové deníky, ostatní, politický bulvár, stranické weby, web instituce);

lemma: základní (slovníkový) tvar automaticky přiřazený ke každému tvaru/wordu;

Twitter: posty, včetně reakcí;

word: konkrétní slovní tvar.

Zdroje:

Cvrček, Václav - Richterová, Olga (eds). cnk: syn2015 [online]. Příručka ČNK; 2016 [cit. 2021-05-31]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=cnk:syn2015&rev=1476702845>

Cvrček, Václav - Richterová, Olga (eds). kurz:pokrocile_dotazy [online]. Příručka ČNK; 2018, [cit. 2021-05-31]. Dostupné na: http://wiki.korpus.cz/doku.php?id=kurz:pokrocile_dotazy&rev=1535546576

Cvrček, Václav - Richterová, Olga (eds). seznamy:frazemy [online]. Příručka ČNK; 2018 [cit. 2021-05-31]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=seznamy:frazemy&rev=1534164400>

Cvrček, Václav - Richterová, Olga (eds). kurz:zobrazeni_dotazu [online]. Příručka ČNK; 2018 [cit. 2021-05-31]. Dostupné na: http://wiki.korpus.cz/doku.php?id=kurz:zobrazeni_dotazu&rev=1534344877

Cvrček, Václav - Richterová, Olga (eds). pojmy:dotazovaci_jazyk [online]. Příručka ČNK; 2016 [cit. 2021-05-31]. Dostupné na: http://wiki.korpus.cz/doku.php?id=pojmy:dotazovaci_jazyk&rev=1458408131

Cvrček, Václav - Richterová, Olga (eds). cnk:online [online]. Příručka ČNK; 2021 [cit. 2021-05-31]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=cnk:online&rev=1611752>

Úvod

Pro diplomovou práci jsem si vybrala téma Českého národního korpusu a jeho využití ve výuce českého jazyka se zaměřením na lexikologii. S korpusem jsem již pracovala v bakalářské práci, kde jsem analyzovala frekvenci frazémů za pomoci korpusového manažeru KonText. Rozhodla jsem se v tématu pokračovat a uplatnit korpus ve vyučování. Výsledky bakalářské práce ukázaly, že frazémy se v textu objevují, a to nejčastěji v podobě přísloví. Proto i tato práce bude obsahovat cvičení na frazémy, neboť frazémy obohacují naši slovní zásobu.

Kvalifikační práce se dělí na teoretickou a výzkumnou část. Teoretická část se člení do několika kapitol a podává základní přehled o didaktice českého jazyka; dále o vzniku Českého národního korpusu, jeho členění, korpusových aplikacích a využití ve škole. Za pomoci odborných publikací a článků představuji korpusovou lingvistiku, která zkoumá jazyk prostřednictvím korpusů. V poslední kapitole uvádím lexikologii, obohacování slovní zásoby, zvláště frazeologii.

Výzkumná část popisuje tvorbu pracovního listu, který je vytvořen za účelem diplomové práce, stanovuje cíle, organizaci a formu výzkumu. Při tvorbě pracovního listu používám korpusový manažer KonText a WaG, kde pokládám základní a pokročilý dotaz, tzv. CQL. Získaná data analyzuji z psaných korpusů syn2015, syn v8, online_now a syn2020, která poté třídím a vybírám vhodná slova do pracovního listu tak, aby odpovídala přiměřenosti věku a intelektu žáků. Pracovní list obsahuje cvičení na slovní zásobu, zvláště na problematiku frazémů. Dále shrnuji výsledky vyplněných pracovních listů žáků z nižšího stupně gymnázia.

Hlavním cílem empirické části je zjistit, zda má korpus své uplatnění ve výuce českého jazyka a je vhodnou didaktickou pomůckou pro učitele.

Teoretická část

1 Didaktika českého jazyka

1.1 Obecná didaktika

Pedagogika se zaměřuje na vychovávání a vzdělávání. Výchova se v různých odborných publikacích vymezuje různě. Obecně se uvádí jako nejobecnější pojetí činnosti, díky ní se předává „duchovní majetek“ společnosti každé další generaci. Dochází k zprostředkování vzorců a norem chování, komunikačních rituálů, hygienických návyků apod. V pedagogickém pojetí je výchova doménou v procesu socializace společnosti. *„Výchova je v pedagogickém pojetí považována za záměrné působení na osobnost jedince s cílem dosáhnout změn v různých složkách osobnosti.“* Pedagogové hovoří o tzv. formování osobnosti, s tím rozlišují výchovu rozumovou, mravní, citovou, tělesnou, aj. Pojem vzdělávání se obecně formuluje jako proces záměrného a organizovaného osvojování poznatků, dovedností, postojů aj. ve školním prostředí.¹

Obecná didaktika studuje obecné otázky, které se vztahují ke vzdělávání a procesu vyučování a učení. Kromě obecné didaktiky se zvláště oddělily i oborové didaktiky jako samostatné disciplíny. *„Pěstování oborových didaktik předpokládá jak dobré znalosti v oboru pedagogiky a psychologie, tak také solidní znalosti příslušných vědních oborů, ke kterým se oborová didaktika vztahuje.“* Didaktika poskytuje oborovým didaktikám teoretický základ. Oborové didaktiky jsou důležitým zdrojem pro didaktické výzkumy. Jejich vzájemná spolupráce je přínosná a žádoucí.²

Rozlišujeme dvě roviny při vyučování a učení, jedna se vztahuje k **obsahu** a druhá k **vlastnímu procesu**. Obsah se vyvíjí stabilně, tzv. konstantně. Naopak procesní složka se stává dynamičtější v interakci učitele a žáka. V této složce dochází k vyučovacímu procesu. Právě tento proces může být vnímán několika způsoby, obecná didaktika ho nazývá jako systém, který zahrnuje **cíl, obsah, podmínky a prostředky výuky, včetně výsledků**. Ve vyučovacím procesu rozlišujeme aktivity subjektů: *a) činnost učitele*

¹ PRŮCHA, Jan. *Přehled pedagogiky: úvod do studia oboru*. 4., aktualizované vydání. Praha: Portál, 2015, s. 16-17.

² SKALKOVÁ, Jarmila. *Obecná didaktika: vyučovací proces, učivo a jeho výběr, metody, organizační formy vyučování*. Praha: Grada, 2007. Pedagogika (Grada), s. 18.

(označujeme jako vyučování); b) činnost žáků, jejímž prostřednictvím získávají své vědomosti (označujeme jako učení).³

Obecně se předpokládá, že daný jedinec si osvojí určité vědomosti a dovednosti, chápe vztahy mezi poznatky, při řešení úkolů uplatní své vědomosti a bude se dále vzdělávat. Také má docházet k osvojení si estetických a morálních hodnot. Jedinec si má utvářet své životní postoje ke světu, společnosti i sebe samému. Jeho postoje se odráží v kritickém myšlení, praktických názorech a jednání. Obecné cíle vzdělávání jsou formulovány společností i jedinci. Předpokládá se, že se jedinec začlení do společenského, kulturní i pracovního procesu. Vzdělání je historicky podmíněno a je nutné cíle, obsah i funkce vzdělávání analyzovat vzhledem k měnící se společnosti.⁴

1.2 Vymezení didaktiky českého jazyka

Jazykové vyučování se zakládá na specifických názorech a přesvědčeních. Mohou být vytvořeny 4 principy:

1. tradice – Učitel jedná podle svého přesvědčení, jak je jeho zvykem, jak on sám to zažil jako žák či začínající učitel. Sám přejímá určité postoje, hodnocení a názory. Později se k tomu přidá jeho osobní zkušenost z praxe, která se zakládá na osvědčených postupech. Tradice se také uplatňuje při přípravě vysokoškolského studenta na jeho budoucí profesi učitele. Při hospitacích se současně s přímou osobní zkušeností uplatňuje i analýza. Metodika také může být zprostředkovatelem tradice.⁵

2. autorita – V souvislosti s autoritou mluvíme o určitém modelu vyučování, který se doporučuje na základě vnější autority – ředitel, osnovy, metodická příručka, učebnice apod. Tento princip se vedle tradice řadí mezi nejrozšířenější. Uplatní se zde i metodika, která dané metody předepisuje, nebo naopak odmítá. Důraz je především kladen na důkladnější poznání vzdělávacích postupů. Působení autority se pojí s funkcí zprostředkovatele tradice. Základem se stává studium a hodnocení praxe.⁶

3. intuitivní náhled – V tomto principu učitel jedná daným způsobem, vychází ze své intuice, má jistý předpoklad (ale bez uvědomovaných důvodů). Může tak dosáhnout pozitivních výsledků. Intuitivní náhled je nezbytný v situacích, kdy se učitel nemůže tzv.

³ JANIŠ, Kamil. *Obecná didaktika - vybraná témata*. Vyd. 4. Hradec Králové: Gaudeamus, 2010, s. 12.

⁴ SKALKOVÁ, Jarmila. *Obecná didaktika: vyučovací proces, učivo a jeho výběr, metody, organizační formy vyučování*. Praha: Grada, 2007. Pedagogika (Grada), s. 27-28.

⁵ ŠEBESTA, Karel. *Od jazyka ke komunikaci: didaktika českého jazyka a komunikační výchova*. Vyd. 2. Praha: Karolinum, 2005. Acta Universitatis Carolinae, s. 15.

⁶ Tamtéž, s. 15-16.

o něco opřít, co má vyzkoušené nebo doporučené. Metodika nám v tomto ohledu nepomůže.⁷

4. teoretické poznání – Nejpodstatnějším principem při procesu jazykového vyučování a učení je právě vědecké poznání. Učitelé poskytují pojmy a teorii, nástroje a analýzy k racionální analýze, plánování, hodnocení výsledků a posuzování vyučovacího procesu. Nejedná se o bezprostřední vztah k vyučovací praxi. Představuje teoretickou část didaktiky češtiny, která má základ v teoretických východiscích jazykového vyučování.⁸

1.3 Jazykové cíle

Na systém vzdělávacích cílů se nahlíží jako na nedělitelný celek. Úkol učitele spočívá v pochopení principů vzdělávacích cílů a aplikování je do vlastní pedagogické praxe. Kritické myšlení vychází z kořenů ze systému vzdělávacích cílů, které sestavil Benjamin Bloom. Rozlišil systém taxonomie vzdělávacích cílů pro rozvoj kognitivních funkcí do šesti fází. V systému zahrnuje znalost, pochopení, aplikaci, analýzu, syntézu a v poslední fázi hodnocení. Bloomova teorie byla zcela výjimečná, že i v současnosti ovlivňuje úhel pohledu na vzdělávací cíle. Na jeho základě rozeznáváme dvě úrovně kognitivních funkcí: **poznávací funkce nižšího řádu** (znalost faktů, pojmů a postupů) a **poznávací funkce vyššího řádu** (abstraktní myšlení a schopnosti srovnávací, hodnotící, aplikační).⁹ „*U Bloomovy taxonomie by měl učitel dbát na to, aby jednotlivé činnosti, kterým odpovídají akční slovesa, postupovaly od nejnižší úrovně k nejvyšší.*“ Ke každé fázi náleží tzv. aktivní slovesa. Jedná se o pomůcku k vysvětlení každé fáze. Některá slovesa mohou být využitelná i pro více fází.¹⁰

Učitelé českého jazyka by měli volit takové metody, které žáky rozvíjejí v mnoha směrech (např. rozvoj mechanické a logické paměti). Žák má být schopen reprodukovat ověřené poznatky, samostatně přemýšlet o problému, rozvíjet svou kreativitu a abstraktní myšlení, uplatnit své teoretické znalosti v praxi. Obecným cílem učitele je dynamické střídání činností, metod a použití různých myšlenkových postupů ve výuce. Důležitou součástí se stává oblast prověřování, hodnocení a klasifikace. Cílem však není, aby učitel

⁷ ŠEBESTA, Karel. *Od jazyka ke komunikaci: didaktika českého jazyka a komunikační výchova*. Vyd. 2. Praha: Karolinum, 2005. Acta Universitatis Carolinae, s. 16.

⁸ Tamtéž, s. 16.

⁹ SIEGLOVÁ, Dagmar. *Konec školní nudy: didaktické metody pro 21. století*. Praha: Grada, 2019, s. 20-23.

¹⁰ ČAPEK, Robert. *Moderní didaktika: lexikon výukových a hodnotících metod*. Praha: Grada, 2015. Pedagogika (Grada), s. 434-438.

na každé hodině aplikoval několik metod. Vždy by měl být počet metod vyvážený a přiměřený.¹¹

Mezi obecné cíle patří informativní, formativní a instrumentální cíl. Informativní cíl se snaží předat žákovi určité množství faktografických informací pro poznání oboru a jeho další aplikaci. Především se jedná o žákovy znalosti. Formativním cílem rozvíjíme žákovy postoje a návyky (rozvoj pečlivosti ve studiu, pracovitosti, aktivity a samostatného úsudku). Instrumentální cíl se zaměřuje na rozvoj v oblasti aplikačních dovedností a schopností. Automatizace dovedností se později stává návykem. Při kladení didaktických cílů uvažujeme nad vhodnými metodami, promyšlíme strukturu vyučovací hodiny a působíme svou osobností pedagoga. Z obecných cílů si stanovujeme dílčí cíle výuky.¹²

Hlavním cílem výuky v češtině je rozvoj komunikačních schopností spisovným jazykem. Schopnost komunikace se rozlišuje podle stupně školy a s postupným jazykovým vzděláváním. Žák se s českým jazykem seznamuje postupně v každém ročníku školy. V procesu poznávání jazyka se žákovi zvyšuje úroveň jazykového povědomí. Ani absolventi škol nezapomenou své osvojené znalosti, pokud projdou kvalitní úrovní výuky, mohou jejich znalosti být na vysoké úrovni.¹³

Pokud máme nastavenou výchovnou složku našeho pedagogického působení, osobnostní složku pedagoga, informativní, formativní a instrumentální požadavky navrhuje J. Kostečka tuto alternativu cílových kvalit pro jazykovou část českého jazyka:

- *relativně ucelené a teoretické poučení o vzniku, vývoji, struktuře a funkci jazyka a češtiny zvláště;*
- *seznámení s institucemi a příručkami, v nichž je možno načerpat další, mimoškolní poučení o češtině pro potřeby budoucí praxe žáka;*
- *aktivní zvládnutí spisovné normy českého jazyka v projevech mluvených i psaných, především v rovině fonetické, gramatické, lexikální a pravopisné s akcentováním požadavku na komunikativnost s ohledem na budoucí, mimoškolní potřeby absolventů;*
- *solidní zvládnutí jazyka mateřského jako opora pro studium jazyků cizích;*

¹¹ KOSTEČKA, Jiří. *Do světa češtiny jinak: (didaktika českého jazyka a slohu pro vyučovací praxi)*. Jinočany: H & H, 1993, s. 11-12.

¹² Tamtéž, s. 12.

¹³ ČECHOVÁ, Marie a Vlastimil STYBLÍK. *Čeština a její vyučování: didaktika českého jazyka pro učitele základních a středních škol a studenty učitelství*. Praha: SPN - pedagogické nakladatelství, 1998, s. 10.

- *získání dostatečného jazykového základu pro uvědomělou a poučenou ideově-estetickou interpretaci literárních textů, a tím i pro hlubší estetický zážitek a obohacení jeho osobnosti při četbě (srov. obdobný cíl u složky slohové).*¹⁴

Stanovené cíle do hodin českého jazyka jsou náročné, ale ve spojení s podporou učebnic, metodických knih, autorů učebních osnov, úsilí učitelů a dobré vůle žáků se dají cíle naplnit. Někteří starší a zkušenější učitelé musí překonat své stereotypní návyky a osvojit si nové přístupy k učivu, seznámit se s novými metodami a formami práce. Hlavní reflexí pro učitele jsou výsledky práce žáků, zvláště v českém jazyce – jazykové dovednosti, vědomosti, návyky, schopnosti, postoje a názory. Z těchto výsledků by měla být posuzována úspěšnost učitele. „*Učitel má v rámci stanoveném učebními osnovami dostatek prostoru pro tvořivé uplatnění své osobnosti, komplexu daného vzděláváním, životními zkušenostmi, nadáním i charakterovými vlastnostmi uplatňovanými při učitelské praxi.*“ Každý učitel si nese své osobnostní rysy. Nelze stanovit dokonalý obraz učitele českého jazyka, ideální učitel má různé verze, jaký by měl být, ale některé základní rysy by měli mít všichni učitelé nehledě na obor, který vyučují.¹⁵

1.4 Jazykový systém ve vyučování

Z hlediska slovní zásoby vyplývá pro učitele, aby u žáků nejen rozšiřoval slovní zásobu (aktivní i pasivní), ale také je i vedl k chápání významu a užívání slov. Náročnější se stává pro žáky pochopení obsahu slova, tj. poznání složek významu. Zaměřit se ve výuce na pochopení pojmů: **věcný význam** (informační obsah), **symbolický význam** (např. slovo *jedle* je symbol štíhlosti a výšky) a **slovotvorný význam**. Při učivu slovní zásoby sledujeme významové vztahy mezi slovy. Proměnu slovní zásoby vztahujeme k vývoji společnosti. Nová slova vznikají s návazností na společenské změny. K obohacení slovní zásoby patří znalost frazémů. Z výzkumů o znalosti frazémů u žáků vyplývá, že žáci užívají nejfrekventovanější, jednoduché frazémy (přirovnání typu *Je chytrá jako liška*, rčení typu *Má za ušima*). Chápání frazeologie podporuje rozvoj myšlení. Ve školním prostředí užíváme běžnou tradiční frazeologii. V hodinách českého jazyka můžeme frazémy interpretovat, analyzovat vztah slova a frazému, zkoumat stupeň frazeologie některých slovních spojení atd. Začleněním frazeologie do výuky jazyka

¹⁴ KOSTEČKA, Jiří. *Do světa češtiny jinak: (didaktika českého jazyka a slohu pro vyučovací praxi)*. Jinočany: H & H, 1993, s. 13.

¹⁵ ČECHOVÁ, Marie a Vlastimil STYBLÍK. *Čeština a její vyučování: didaktika českého jazyka pro učitele základních a středních škol a studenty učitelství*. Praha: SPN - pedagogické nakladatelství, 1998, s. 11.

můžeme ovlivnit atmosféru a zároveň přinést do hodin vtip i humor, zároveň posilujeme výchovný aspekt, neboť frazeologie obsahuje řadu životních moudrostí.¹⁶

1.5 Motivace

Při zkoumání motivace ve výchovně vzdělávacím procesu se na motivaci nahlíží dvěma způsoby: a) jako prostředek zvyšování efektivity žáků – motivovanost žáků ve vyučování; b) jako jeden z cílů výchovně vzdělávacího působení školy – rozvoj motivace žáků. Nelze ale tyto dva způsoby rozdělovat, neboť se vzájemně ovlivňují a doplňují. Motivace učební činnosti závisí na motivaci osobnosti a naopak. V průběhu učení, zejména v aktivním učení, příznivé atmosféře, použití vyučovacích metod a dobrých vzájemných vztazích se motivace žáka může změnit. Žák má zájem o předmět, ale ještě více působí dílčí motivy. Stoupá jeho zvědavost, potřeba činnosti a nastavení dlouhodobých cílů ve svém životě.¹⁷ Zkušení a začínající učitelé kladou důraz na motivaci jako předpoklad úspěšnosti v učení. Největším úkolem je motivovat žáky, aby se chtěli učit. Pokud se žáci nechtějí učit, stane se pro ně výuka neefektivní a případně se vůbec nic nenaučí. Jestliže učitel žáky dokáže motivovat, může se od toho odvíjet tempo výuky.¹⁸

Pravidla pro probuzení zájmů žáků:

- *Projevujte zájem o svůj obor, buďte pro něj zapálení.*
- *Ukazujte jeho spojitost se skutečným světem. Noste do hodin předměty z praxe, použijte instruktážní filmy, hovořte o konkrétní aplikaci učiva, začleňte do vyučování návštěvy odborníků a exkurze.*
- *Využívejte tvořivosti a sebevyjadřování žáků.*
- *Přesvědčujte se, že se žáci aktivně zapojují do výuky.*
- *Pravidelně obměňujte činnosti žáků.*
- *Využívejte překvapení a neobvyklých činností.*
- *Zadávejte třídě soutěživé a problémové úlohy.*
- *Dávejte žákům „hádky“, na které jim později sdělíte správnou odpověď.*

¹⁶ ČECHOVÁ, Marie a Vlastimil STYBLÍK. *Čeština a její vyučování: didaktika českého jazyka pro učitele základních a středních škol a studenty učitelství*. Praha: SPN - pedagogické nakladatelství, 1998, s. 94-99.

¹⁷ HRABAL, Vladimír, František MAN a Isabella PAVELKOVÁ. *Psychologické otázky motivace ve škole*. 2., upr. vyd. Praha: Státní pedagogické nakladatelství, 1989. Knižnice psychologické literatury, s. 24.

¹⁸ PETTY, Geoffrey. *Moderní vyučování*. 6., rozš. a přeprac. vyd. Přeložil Jiří FOLTÝN. Praha: Portál, 2013, s. 53.

- *Propojte učení s tím, co žáky zajímá mimo školu.*
- *Dodejte svému oboru „osobní rozměr“.¹⁹*

Motivaci lze rozdělit na **dlouhodobou** a **aktuální** (krátkodobou), přičemž každá z nich je buď vnitřní, nebo vnější. Dlouhodobá motivaci sleduje trvalé či průběžné studijní činnosti žáka. Vnější motivaci může ovlivnit např. osobnost učitele, rodičů, obava z neúspěchu; vnitřní motivace se pojí s talentovou stránkou žáka a jeho vnitřním zájmem o daný předmět. Motivace často souvisí i s věkem, s předchozím studiem, se zpětnou vazbou, ale také i s odměnou a trestem.²⁰

U dlouhodobé motivace se stává silnou motivací apel učitele na užitečnost bohemistického základu pro studium cizích jazyků. Téměř všichni absolventi jakékoliv školy budou formulovat písemné projevy různého charakteru. Nezáleží na budoucím povolání absolventa, ať už se stane lékařem, právníkem, historikem apod. Absolvent má být schopný se vyjádřit a neměl by mít jazykové nedostatky. Dlouhodobým úkolem při budování motivace u žáka je přesvědčení významu a důležitosti studia jazyka. Jazykové znalosti uplatní v životě každý a je naší nezbytnou součástí. Učitel proto musí zvolit vhodnou metodu. Motivací pro žáky může být zařazení volné hodiny přibližně jednou za dva měsíce. Pokusit se připravit zajímavou jazykovou či slohovou hodinu formou například hry nebo soutěže, žáci se na takové hodiny těší a zvýší se jejich aktivita při vyučování. Velmi žáci oceňují poutavost a odlehčenost tématu, takové hodiny se i pojí s úspěchem netradičního pojetí hodiny. Zařazení motivačních hodin se vyplatí pro udržení dlouhodobé motivace v dalších hodinách. Svou roli v motivaci má role a osobnost učitele, který by měl umět žákům naslouchat a ptát se. Nezbytnou součástí motivace je pochvala žáků. Výrazný vliv, který na žáky působí, má klasifikace a hodnocení.²¹

„Základ aktuální motivace tvoří vhodně zvolená metoda a forma výuky, přehledná struktura hodiny s rytmicky se střídajícími činnostmi a dodržování didaktických zásad.“ Věnujme čas zajímavé myšlence žáka na úkor probírané látky v hodině a rozvíjejme společně tuto myšlenku ve třídě, ať se každý žák zamyslí. Podporujeme tak rozvoj fantazie, myšlení a smyslů. Formulujme správně téma hodiny, vazba na praktičnost

¹⁹ PETTY, Geoffrey. *Moderní vyučování*. 6., rozš. a přeprac. vyd. Přeložil Jiří FOLTÝN. Praha: Portál, 2013, s. 60.

²⁰ KOSTEČKA, Jiří. *Do světa češtiny jinak: (didaktika českého jazyka a slohu pro vyučovací praxi)*. Jinočany: H & H, 1993, s. 16-17.

²¹ Tamtéž, s. 17-21.

tématu v životě přispívá k motivaci. Vhodně vyjádřeme otázku či daný úkol k tématu hodiny, žáci potřebují znát smysl úkolu.²²

1.6 Vztah učitele a žáka

Jak učitelé vyučují není z pedagogických výzkumů zcela objasněno. Výzkumy však rozpracovaly, že se učitelé ve své realizaci vyučování liší. Tak to se nazývá koncepce o pedagogickém myšlení učitelů, resp. o učitelově pojetí výuky. Tyto koncepce se snaží o vysvětlení, jak učitel přistupuje ke své roli vyučujícího, jakou má odpovědnost za výsledky svých žáků, jak přistupuje k plánování cílů vyučování a k výběru učiva, jaké postoje zastává ke svým nadřízeným, k rodičům a žákům a zejména i ke svému dalšímu vzdělávání v oboru atd.²³

Jedním z významných prostředí, které žáka formuje a ovlivňuje, se nachází ve škole. Z hlediska didaktiky jazyka probíhá průběh procesu socializace žáka ve školním prostředí. Škola představuje pro žáka nový typ prostředí, kde přichází do kontaktu s novými požadavky na společenskou interakci a komunikaci. Žák se musí naučit, jak spolupracovat s ostatními žáky ve třídě, s učebními materiály, ale i spolupráci s učiteli. Výchovně-vzdělávací působení se cíleně zaměřuje na přípravu žáků v komunikaci ve škole a ve společnosti.²⁴ Při seznamování s žáky záleží na prvním dojmu, který ovlivňuje vztahy a vytváří úsudky. Z úvodních informací si utváříme dojem, který v nás osoba zanechala. Záleží na faktorech, jako je oblečení, účes, výraz tváře, gestikulace, věk, postava, co osoba říká a jak to říká. Učitel se zaměřuje na rukopis žáka, jeho dosavadní prospěch, pověst apod. Stejně tak jako většina populace i učitel má své stereotypy. Pokud zastáváme tento postoj, vede nás k tomu, že nevnímáme rozdíly mezi členy skupiny. Očekávání učitele žáka ovlivňuje, když si učitel myslí, že žák je dobrý, žák bude mít skvělé výsledky. Naopak když je žák špatný, bude se zhoršovat. „*Všichni bez výjimky potřebují mít pocit důstojnosti, smysluplnosti a sebeúcty.*“ Pokud toto docílíme u žáků, budou se snažit vydat ze sebe to nejlepší.²⁵

V současné době se mladá generace učitelů v mnohém liší. Narodili jsme se do jiných podmínek, kde disponujeme s větším množstvím znalostí, protože máme mnohem

²² KOSTEČKA, Jiří. *Do světa češtiny jinak: (didaktika českého jazyka a slohu pro vyučovací praxi)*. Jinočany: H & H, 1993, s. 21-22.

²³ PRŮCHA, Jan. *Přehled pedagogiky: úvod do studia oboru*. 4., aktualizované vydání. Praha: Portál, 2015, s. 140.

²⁴ ŠEBESTA, Karel. *Od jazyka ke komunikaci: didaktika českého jazyka a komunikační výchova*. Vyd. 2. Praha: Karolinum, 2005. Acta Universitatis Carolinae, s. 46-47.

²⁵ PETTY, Geoffrey. *Moderní vyučování*. 6., rozš. a přeprac. vyd. Přeložil Jiří FOLTÝN. Praha: Portál, 2013, s. 86-89.

dostupnější přístup k informačním zdrojům. Musíme uvažovat efektivně, abychom se uplatnili v komunikačních technologiích a digitalizaci. V řadě případů jsou proto kladeny vyšší nároky a požadavky. Nijak se neliší význam učitele v procesu vzdělávání, který se podílí na charakteru a postoje žáků. Zcela je respektován, je-li kvalifikovaným a inspirativním učitelem. Učitel má nejen žáky vzdělávat, předávat jim teoretické znalosti, ale „provázet“ je životem a pomáhat jim se orientovat v dnešní době. Úkol učitele náleží na příznivé atmosféře a vzájemné důvěře. Žáky by měl motivovat, vést je ke spolupráci, aby se neobávali svých omylů a chyb. Učitel jim je vzorem, partnerem, ale ne kritikem, mentorem či soudcem.²⁶ Schopnost komunikace je základní vlastností učitele českého jazyka. K vytváření dobrých vztahů mezi učitelem a žákem je třeba žáka poznat, znát jeho zájmy, rodinné zázemí, úroveň řečové komunikace. Do osobnosti žáka se dostáváme rozhovorem, díky tomu zjišťujeme jeho názory, obavy, zkoumáme zkušenosti a cíle. Rozhovor závisí na důvěře a jak se dovedeme ptát. Komunikace je základem k prohloubení dobrých vztahů mezi učitelem a žákem, klíč k jeho vnitřnímu duševnímu světu, teprve poté můžeme pracovat na výchovném efektu.²⁷

„Didaktický princip orientace na žáka koresponduje s pedocentrickým přístupem k výuce češtiny a s celkovou humanizací školy.“ Princip orientace na žáka například zohledňuje všeobecné didaktické zásady přiměřenosti výuky věku, respektuje kognitivní, sociální, emoční a jazykový vývoj žáka. Reflektuje aktuální jazykové teorie a poznatky z oblasti lingvistiky v daném jazykovém vývoji žáka ad. Nastává-li, že žáka vyučujeme v jazykově-komunikační edukaci, co je součástí jeho jazykového vědomí a v praktické komunikaci vše zvládá, stává se pro něho výuka nepotřebná. Spíše žáka demotivuje a narušuje jeho vztah k danému předmětu, zvláště jeho vztahu k jazyce.²⁸

1.7 Vyučovací metody

Pojem metoda se odvozuje od řeckého slova „*meta hodos*“ a znamená cesta směřující k cíli. Tímto pojmem označujeme dané prostředky, postupy a návody, s kterými můžeme dosáhnout cíle. Prostřednictvím výukové metody rozvíjíme vzdělání žáků a dosahujeme výchovně-vzdělávacích cílů. Důležitou roli však zastává vzájemná spolupráce učitele s žáky. Při aplikování metody vede učitel žáka k vytvoření si určitého

²⁶ SIEGLOVÁ, Dagmar. *Konec školní nudy: didaktické metody pro 21. století*. Praha: Grada, 2019, s. 29-31.

²⁷ NELEŠOVSKÁ, Alena. *Pedagogická komunikace v teorii a praxi*. Praha: Grada, 2005. Pedagogika (Grada), s. 82.

²⁸ ŠTĚPÁNÍK, Stanislav, Eva HÁJKOVÁ, Klára ELIÁŠKOVÁ, Ludmila LIPTÁKOVÁ a Marta SZYMAŃSKA. *Školní výpravy do krajiny češtiny: didaktika českého jazyka pro základní školy*. Plzeň: Fraus, 2020, s. 38-39.

učebního stylu a osamostatnění se. Výukové metody se řadí mezi základní didaktické kategorie a plní funkci zprostředkování vědomostí a dovedností, funkci aktivizační, učitel tak žáka motivuje k učení, funkci formativní, funkci výchovnou a funkci komunikační.²⁹

Vybrané výukové metody:

Aktivizační metoda – Mapa pojmů

Mapa pojmů je kognitivní metoda, která se zaměřuje na logické vazby v souboru vybraných pojmů napsaných na samostatných kartičkách. Žáci mohou pracovat s hesly, pojmy, klíčovými slovy, terminologií, názvy, jmény, ale také na kartičkách mohou být napsané věty, definice nebo natištěné obrázky. Metoda pomáhá žákům v učení se znalostí, získání základního přehledu v tématu, ujasnění si pojmů a jejich významy, chápání kontextu. Aktivitu lze zařadit do jazykové hodiny a zaměřit ji na rozvoj slovní zásoby. V postupu práce učitel rozdělí žáky do menších skupin, aby všichni aktivně pracovali. Každá skupina dostane sadu kartiček, učitel zdůrazní, že není pouze špatné či správné řešení, záleží na žakově obhajobě řešení. Po skončení aktivity následuje společná diskuse, která dále rozvíjí argumentaci žáků. Učitel diskusi řídí a zpětně reaguje.³⁰

Metoda práce s textem – Doplnování

Tyto metody spojené s prací s textem mají několik cílů. Mezi jedny nejdůležitější patří schopnost čtení s porozuměním. Navýší se komunikativní dovednosti a dovednosti analytické. Čtenářské dovednosti rozvíjí doplňování do textu.³¹

Příklad

Zadání: Dopln do příběhu slova: vládl, nejstarší, poručník, trůn.

Příběh jedné daleké země

V nedávné historii země Farland ___ král Richard III. Moudrý. Vládu zdědil jako ___ syn předchozího vládce, krále Richarda II. Krásného, svého otce. Nejprve byl sice příliš malý, takže zemi spravoval jeho ___ (strýc Matěj), ale v den své plnoletosti byl korunován a usedl na ___. Takovému státnímu zřízení říkáme konstituční monarchie.³²

²⁹ ZORMANOVÁ, Lucie. *Výukové metody v pedagogice: tradiční a inovativní metody, transmisivní a konstruktivistické pojetí výuky, klasifikace výukových metod*. Praha: Grada, 2012. Pedagogika (Grada), s. 13.

³⁰ SIEGLOVÁ, Dagmar. *Konec školní nudy: didaktické metody pro 21. století*. Praha: Grada, 2019, s. 110-111.

³¹ ČAPEK, Robert. *Moderní didaktika: lexikon výukových a hodnoticích metod*. Praha: Grada, 2015. Pedagogika (Grada), s. 300-301.

³² Tamtéž, s. 301.

Názorné metody mají v jazykovém vyučování nejzásadnější postavení, kromě učebnic a pracovních sešitů lze v hodinách aplikovat další didaktické pomůcky: akustické pomůcky (magnetofon), optické pomůcky (dataprojektor), kombinované pomůcky (počítač, televize), tabelární pomůcky (nástěnné tabule, pracovní listy, soubory obrazů, ...).³³

³³ ČECHOVÁ, Marie a Vlastimil STYBLÍK. *Čeština a její vyučování: didaktika českého jazyka pro učitele základních a středních škol a studenty učitelství*. Praha: SPN - pedagogické nakladatelství, 1998, s. 69-70.

2 Český národní korpus

2.1 Vznik ČNK

Český národní korpus je akademický projekt, který byl založen v roce 1994. Zpřístupňuje jazyková data pro výzkum, ale i výuku ve formě elektronických korpusů. Rozsah ČNK zajišťuje přístup k více než 3 miliardám slov, které jsou utříděny v korpusech synchronních, diachronních, mluvených, psaných, paralelních a jednojazyčných. Na projektu ČNK pracují pracovníci zejména z ústavů Filozofické fakulty Univerzity Karlovy: Ústav Českého národního korpusu a Ústav teoretické a počítačové lingvistiky. Při sběru dat a koordinaci úkolů se na práci v ČNK podílejí i externisté z celé České republiky.³⁴

Počátky směrem ke korpusu u nás byly počítačově zpracované slovníky (Retrográdní slovník současné češtiny – Academia 1986, Anglicko-český výkladový slovník výpočetní techniky – SNTL 1990). Mnohem přínosnější pro projekt ČNK byly zkušenosti lingvistů při tvorbě slovníkových databází, které vznikaly v 80. a 90. letech 20. století. Autoři počítačových programů vytvořili slovníkovou část automatických korektorů. J. Hajič z Výzkumného ústavu matematických strojů a J. Drózd ze Střediska výpočetní techniky ČSAV sestavili ve svém korektoru databázi s původně 100 tis. českých lexikálních jednotek. Následoval je K. Pala z Katedry českého jazyka Filozofické fakulty Masarykovy univerzity v Brně, který uspořádal databázi o 170 tis. českých kmenů. Už v té době katedra vlastnila 200 tis. elektronizovaných položek kompletního hesláře SSJČ. Morfologicky a syntakticky byl tagován korpus 180 textů, obsahoval texty o velikosti 540 tis. slovních výskytů a jeho rozsah obsahoval téměř 5 mil. údajů. Při Kybernetické společnosti vznikla dne 3. 10. 1988 „Iniciativní skupina pro přípravu počítačových korpusů textů a slovníků“. Hlavním cílem této aktivity bylo sjednocení metodiky lexikografické práce se zahraničím.³⁵

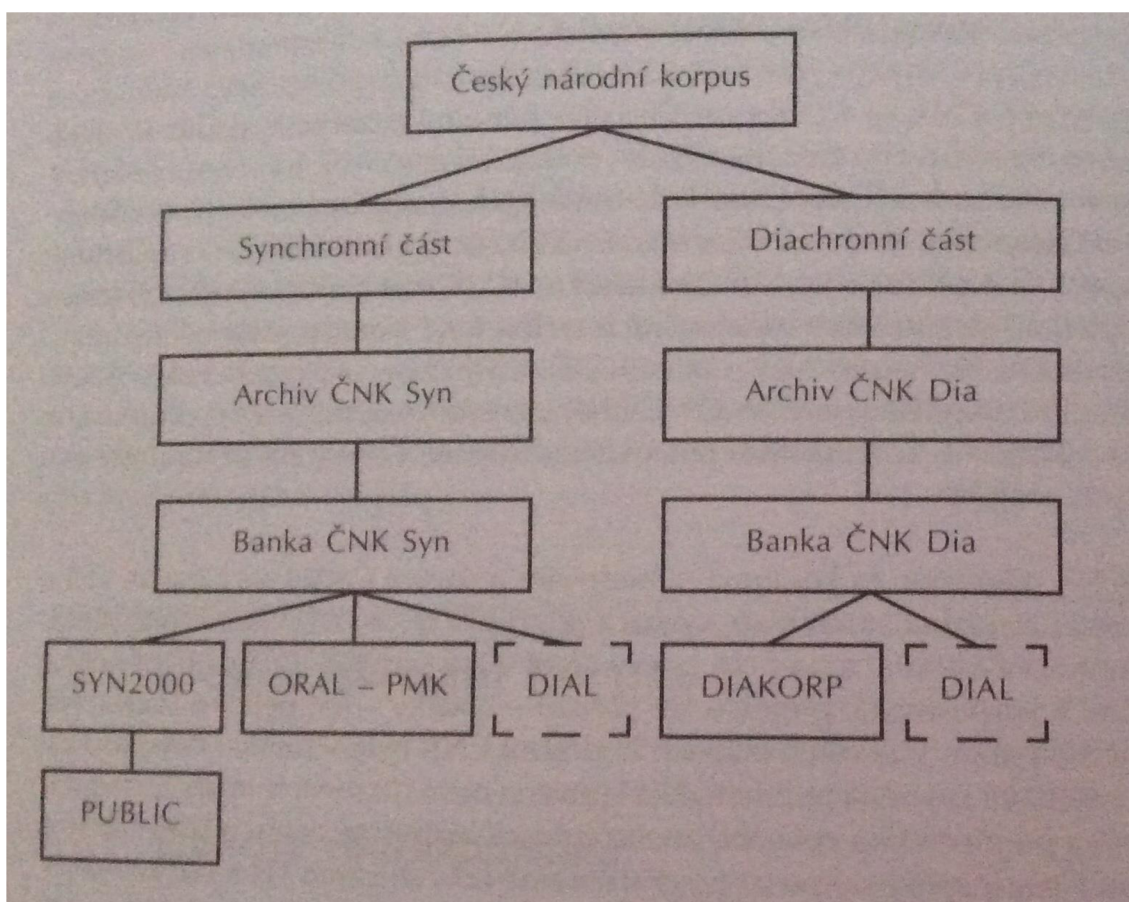
Po změně politického režimu se lingvisté rozhodli společně vybudovat „Počítačový fond češtiny“. Databáze měla být plně elektronická a modernizovaná podle světového trendu. Na projektu spolupracovali pracovníci z FF UK Praha, FF MU Brno, MFF UK Praha, ÚJČ a původně VÚMS (Výzkumného ústavu matematických strojů) Praha. Stanoveným úkolem bylo vytvořit fungující projekt, standardizaci kódů záznamu,

³⁴ Cvrček, Václav - Richterová, Olga (eds). start [online]. Příručka ČNK; [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=start&rev=1602059248>.

³⁵ ŠULC, Michal. *Korpusová lingvistika: první vstup*. Praha: Karolinum, 1999, s. 45.

standardizaci uložení a zpracování, standardizaci editorů pro komentáře, vyřešení právních vztahů a v poslední řadě i financování. První jednání o vybudování ČNK vedl lingvista **F. Čermák**, který je i tvůrce projektu, s vedením Ústavu pro jazyk český AV ČR. Úspěšné jednání se konalo **9. 9. 1994**, kde také vznikl Ústav Českého národního korpusu. ÚČNK získal v roce 1996 grant s názvem „Počítačový korpus českého jazyka“ v rámci programu Posílení výzkumu na vysokých školách.³⁶

2.2 Členění ČNK



KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 12.

Znázorněné členění ČNK má různý rozsah. Velikostně největším se stává synchronní psaný korpus (SYN2000, rozsah přes 100 mil. slovních tvarů), z něho vychází korpus PUBLIC (zhruba 20 mil. slovních tvarů). V menším rozsahu je diachronní psaný korpus (DIAKORP, zhruba 1 750 000 tvarů) a synchronní mluvený korpus ORAL-PMK, 700 tis. tvarů). DIAL označuje nářeční korpusy obojího typu. Každý korpus je výhradně tvořen z celých textů, nikoliv z vybraných vzorků. V případě některých starších děl

³⁶ ŠULC, Michal. *Korpusová lingvistika: první vstup*. Praha: Karolinum, 1999, s. 46.

vývojových fází češtiny jsou dochovány části textů. Nekorpusové složky ČNK se skládají z archivů, kde jsou uchovány všechny získané texty ve výchozí podobě. V bankách se nachází všechny texty v konečném formátu SGML.³⁷

Mimo budování korpusů a jejich specifické metodologie a mapování jazykové situace a s tím spojeného vývoje češtiny se také ČNK zabývá zpřístupňováním dat pro výzkum. Jedná se o oblast automatického zpracování přirozeného jazyka (*NLP, natural language processing*). V současnosti se ČNK natolik vývojově posunul, že může nabídnout bohatou databázi: *skládá se z řady korpusů vázaných časově a žánrově a hostí i některé cizí korpusy (zvláště korpusy obou lužických srbštin)*. Přístupným se stal i mnohojazyčný paralelní korpus InterCorp (2008) pro další lingvistické obory.³⁸

2.3 Reprezentativnost ČNK

V průběhu několika let se u psaných korpusů jejich reprezentativnost a vyváženost v řadě SYN vyvíjela. Korpusy publikované od roku 2010 se opírají o řadu průzkumů, které se snaží kvantifikovat recepci psaných textů u čtenářů. Ze zjištění vyplynuly závěry, že při zkoumání žánrů a textových typů tímto způsobem má několik metodologických problémů. Korpus SYN2015 již obsahuje revizi klasifikace textů a byla přepracována i vyváženost korpusu. U mluveného korpusu (Oral, Ortofon) máme mnohem méně dostupných dat. Pokud byla k dispozici data ze sociolingvistické oblasti (pohlaví, věk, vzdělání a oblast pobytu mluvčího v dětství), byly podle toho korpusy řazeny. Specializované korpusy nejsou reprezentovány jazykovým územím jako celek, pouze nám dokládají specifické užití jazyka.³⁹

Pojem reprezentativnost se nejčastěji objevuje ve statistice a empiricky orientovaných vědách. Vykládá se jako vztah mezi **vzorkem** (zkoumaná množina jednotek) a **populací** (množina jednotek jako základ poznatků, která se zobecňuje). Musíme především vymezit, co je vzorek a co populace. V korpusové lingvistice představuje vzorek určitý korpus, problematické se stává vymezení populace. Musíme určit faktické hranice populace (například všechny texty publikované v angličtině v roce 1961). Nutné je zjistit stratifikaci populace, jaké žánry a registry se v ní používají.

³⁷ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 12-13.

³⁸ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 69.

³⁹ Cvrček, Václav - Richterová, Olga (eds). *pojmy:reprezentativnost* [online]. Příručka ČNK; 2019. [cit. 2021-01-24]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:reprezentativnost&rev=1571233796>.

Na reprezentativnost nahlíží V. Cvrček a D. Kovaříková jako na „*vyváženost korpusu s ohledem na různé typy textů, žánry a témata*“. Pojmu se také věnuje M. Křen, který tvrdí, že pojem je terminologicky nejasný. Ve svém konstatování se přiklání k pojetí Sinclairově: „*Korpus je pouze více či méně věrným odrazem, vzorkem jazyka nebo některé jeho variety, ale nikdy ho nezastoupí celý... Hlavním problémem, který se k reprezentativnosti váže, je však právě nemožnost objektivně měřit, nakolik korpus jazyku odpovídá.*“⁴⁰

Obecně se nedá tvrdit, že by existovala univerzální, všeobecně uznávaná představa o využití korpusů. Výchoziskem pro ČNK se staly odhady očekávání potenciačních zájemců o práci s korpusem. Průzkumy prokázaly několik stupňů poznání. Členění textových zdrojů ČNK se stalo výchozím bodem výsledku průzkumu o poměru čtení knih, časopisů a novin českou populací. V členění nenovinových textů na imaginativní a informativní byly výsledky vyvážené ve třech různých průzkumech, které se týkaly zájmu o literaturu naučnou a literaturu krásnou u různých skupin čtenářů. Ve třetím průzkumu (vnitřní členění informativních a imaginativních textů) se výsledky průzkumu zakládaly na struktuře zájmu o naučnou literaturu, průzkumu struktury zájmu o literaturu ze strany čtenářů nových knih, a průzkumu struktury katalogu domácích periodik.⁴¹

2.4 František Čermák

František Čermák, prof. PhDr., DrSc. (1940)

Svá studia na Filozofické fakultě Univerzity Karlově dokončil v roce 1962. Vystudoval obor češtiny, angličtiny a holandštiny. Docentem obecné jazykovědy se stal v roce 1991 a poté v roce 1994 získal titul profesor pro obor český jazyk. V letech 1991-93 se věnoval práci v Ústavu pro jazyk český v Praze jako vedoucí lexikografického oddělení. Od založení Ústavu Českého národního korpusu na filozofické fakultě v Praze byl řadu let jejím ředitelem. Svou vědeckou činnost vyvíjí nejen v rámci republiky, ale i mezinárodně. Za svou vědeckou činnost byl několikrát vyznamenán Akademií věd a Karlovou univerzitou. Jeho vědecké zaměření se řadí do oblastí obecné lingvistiky, jazykovědné teorie a do jazykovědné bohemistiky, zvláště

⁴⁰ CHROMÝ, Jan (ed.). Korpus a reprezentativnost. *Naše řeč – Základní informace* [online]. 2014. roč. 97, č. 4-5 [cit. 2021-01-24]. Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8337>

⁴¹ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 16-17.

lexikologie a lexikografie. Mezi jeho publikacemi bychom našly i dvojjazyčné slovníky a učební texty.⁴²

Stal se členem několika lingvistických sdružení (Jazykovědné sdružení, Pražský lingvistický kroužek, Societas linguistica Europea aj.), zároveň je členem redakční rady *International Journal of Corpus Linguistics*. Mezi jeho publikace se například řadí: Česká lexikologie, Český národní korpus: 100miliónový SYN 2000, Jazyk a jazykověda, Slovník české frazeologie a idiomatiky, Stručná mluvnice češtiny pro zahraniční studenty, Základy lingvistické metodologie, Základní učebnice češtiny I-V a mnoho dalších.⁴³

2.5 Korpusové aplikace

Korpusový manažer či prohlížeč (angl. corpus manager) umožňuje práci s korpusovými daty. Tyto aplikace jsou komplexním nástrojem k vyhledávání jazykových forem. Dotaz zadáváme ve speciálním dotazovacím jazyce. Korpusové aplikace nabízejí i další funkce (např. vyhledávání kolokací ad.). Manažer pracuje dvěma typy, buď **lokálně**, nebo častěji jako **klient-server**. Lokálně musí být manažer na počítači při práci s korpusem i celý textový korpus. Při principu klient-server je v počítači pouze klientská část, která se zaměřuje na uživatelské rozhraní. Samotné vyhledávání se odehrává na vzdáleném serveru, kde se nachází korpus. Klientský program může být nainstalován na osobním počítači nebo může uživatel používat webovou aplikaci. Serverovou aplikaci CWB (Corpus Workbench) vytvořili hlavní vývojáři současné doby S. Evert a A. Hardie. Manatee je modernější korpusový manažer, jehož základ vychází z CWB a jeho autorem je P. Rychlý. Mezi klientské aplikace řadíme Bonito, Sketch Engine, NoSketch Engine, Park, KonText a další specializované aplikace.⁴⁴

Pro práci s korpusovými daty jsou volně dostupné aplikace na hlavní webové stránce www.korpus.cz v horní liště. Najdeme zde aplikace: WaG, KonText, SyD, Morfio, KWords, Treq, Calc, Lists, Mapka, KorpusDB a QuitaUp.⁴⁵

⁴² VEČERKA, Radoslav. *Biografickobibliografické medaloinky českých lingvistů: bohemistů a slavistů (2. vydání)* [online]. *Linguistica online*. Brno: electronic journal of the Department of Linguistics and Baltic Languages at Masaryk University, Czech Republic, 2008, s. 22 [cit. 2021-01-23]. ISSN 1801-5336. Dostupné z: <http://www.phil.muni.cz/linguistica/art/vecerka/vec-medailonky2.pdf>

⁴³ TŘEŠTÍK, Michael, ed. *Kdo je kdo: osobnosti české současnosti : 5000 životopisů*. Praha: Agentura Kdo je kdo, 2002, s. 84.

⁴⁴ Cvrček, Václav - Richterová, Olga (eds). *pojmy:korpusovy_manazer* [online]. Příručka ČNK; 2016. [cit. 2021-01-24]. Dostupné na: http://wiki.korpus.cz/doku.php?id=pojmy:korpusovy_manazer&rev=1470210588.

⁴⁵ ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019, s. 60.

KonText zpřístupňuje práci s korpusy ČNK, nahrazuje předchozí aplikaci Bonito. Jeho technické vybavení má řadu funkcí, které může uživatel potřebovat při své práci, včetně nástrojů statistických, identifikace kolokací a další, s nimiž souvisejí programy, které ČNK nabízí k dispozici.⁴⁶ Poslední upravená verze KonTextu (18. 12. 2020) byla rozšířena o další funkce, kde byly provedeny změny v typů dotazu na dva: *pokročilý (stejně funkce jako CQL) a jednoduchý – viceslovné hledání, volitelné regulární výrazy a volitelné výchozí atributy (v závislosti na korpusu)*. Dále došlo ke změnám *u nových prvků pro výběr časové intervalu v „Omezení hledání“ založený na klikacím kalendáři, rozšíření plug-inu syntax_viewer - podpora nových vlastností v SYN2020, nových plug-in query_suggest poskytující nápovědu k aktuálně zapsanému dotazu* a v řadě dalších funkcích.⁴⁷

WaG: aplikace **Slovo v kostce** (Word at a Glance) udává základní přehled, jak se dané slovo či sousloví používá. Získáme zde informace, které jsou dostupné v korpusech ČNK a uživateli poskytuje přehledný profil slova z různých perspektiv. V současné době se WaG uplatňuje pouze pro české zdroje. Zahrnuje tři základní moduly: *profil slova či viceslovné jednotky v češtině (modul Vyhledat slovo), komparace profilů dvou (či více) českých slov či sousloví (modul Porovnat slova), informace o možných překladech zadaného slova do jiných jazyků (modul Hledat ve dvou jazycích)*. Uživateli se výsledky daného slova zobrazí do uspořádaných tzv. dlaždic. U každé dlaždice jsou k dispozici podrobné informace o tom, z jakého zdroje údaje pocházejí a zároveň aplikace odkazuje do dalších aplikací korpusů ČNK, kde uživatel může dále pracovat s detailnějšími daty.⁴⁸

SyD (Synchronní a diachronní analýza variant) udává informace k průzkumu vzájemně se konkurujících jazykových jevů. Napomáhá přehledně zprostředkovat uživatelům výsledky korpusu. Má dvě oddělené části: *synchronní, v níž je možné porovnávat varianty podle jejich frekvence a distribuce v textech; diachronní, v níž lze sledovat vývojové tendence v užívání variant*.⁴⁹ Při porovnávání dvou a více jevů se může jednat o varianty pravopisné, morfologické, lexikologické, syntaktické a slovosledné. Při analýze zkoumáme výsledky, jak je výše uvedeno, ze dvou hledisek: *synchronního (z hlediska současného jazyka) a diachronního (z hlediska vývoje jazyka v čase)*.

⁴⁶ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 70.

⁴⁷ Cvrček, Václav - Richterová, Olga (eds). seznamy:kontext_verze [online]. Příručka ČNK; 2021. [cit. 2021-01-25]. Dostupné na: http://wiki.korpus.cz/doku.php?id=seznamy:kontext_verze&rev=1610643179.

⁴⁸ Cvrček, Václav - Richterová, Olga (eds). manualy:wag [online]. Příručka ČNK; 2020. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manualy:wag&rev=1600851035>.

⁴⁹ Cvrček, Václav - Richterová, Olga (eds). manualy:syd [online]. Příručka ČNK; 2017. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manualy:syd&rev=1506346256>.

Pro výuku žáků je vhodné použít synchronní část. V zadaných úkolech žáci mohou pracovat na analýze rozdílů mezi variantami pravopisnými, morfologickými i lexikálními. Zpracované výsledky se zobrazují v několika variantách; a to buď v koláčovém grafu, kde jsou zastoupeny varianty psaného i mluveného jazyka; další varianty jsou uloženy v záložkách pod grafem.⁵⁰

Morfio slouží k pozorování slovtvorné pravidelnosti češtiny. Z korpusových dat můžeme získat produktivitu slovtvorných modelů se specializací na odvozování. Princip práce Morfia je založený na poznacích slovtvorby češtiny. Vyhledává dvojice nebo vícečetné kombinace slov, které mají stejné vlastnosti jako zadaný model.⁵¹ Obecně můžeme tvrdit, že každý slovtvorný vztah je tvořen dalšími dvěma faktory: *formální shodou/podobností v určitých částech slova, tzv. báze* (např. dřev- je část společná pro slova dřevo i dřevěný); *formálními odlišnostmi v částech specifických, tzv. formantech* (morfy -o a -ěný v předchozím příkladu). Konečné výsledky Morfia nemusí být zcela bezchybné, spíše je tato aplikace pomůckou pro badatele, aby při své práci získali data rychleji a jednodušeji. Cílem je snadné dosažení úspěšnosti při hledání daného typu, relevantnost je ponechána na badateli.⁵² Ve výuce lze zahrnout i práci s Morfiem pro odhalování slovtvorných pravidel češtiny, pro uvědomění si jazykového systému a osvojení si nových slovtvorných postupů. Při učení se slovtvorby si žáci osvojují a obohacují slovní zásobu.⁵³

KWords (tzv. klíčová slova, angl. keywords), tato aplikace slouží k analýze textů, jejím základem je srovnávání textů s obecným územ. Cílem se stává zjištění klíčových slov neboli slovní tvary, které se ve zkoumaném textu vyskytují častěji než v referenčním korpusu. Klíčová slova jsou základem pro textovou analýzu a interpretaci. Aplikace byla vytvořena pro analýzu politických projevů a na jejím rozvoji spolupracuje Brownova univerzita. KWords zjišťuje dva typy prominentních slov: *klíčová slova (keywords)*, *slova nesoucí tematickou koncentraci (TC) textu*. Identifikace klíčových slov se srovnává s relativní frekvencí slov v textu s relativní frekvencí daného slova v referenčním korpusu.⁵⁴ Texty může uživatel nahrát v češtině nebo angličtině, avšak uživatel musí

⁵⁰ ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019, s. 61-62.

⁵¹ Tamtéž, s. 65.

⁵² Cvrček, Václav - Richterová, Olga (eds). *manualy:morfio* [online]. Příručka ČNK; 2017. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manualy:morfio&rev=1506346481>.

⁵³ ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019, s. 65.

⁵⁴ Cvrček, Václav - Richterová, Olga (eds). *manualy:kwords* [online]. Příručka ČNK; 2017. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manualy:kwords&rev=1506346748>.

vhodně zvolit referenční korpus, aby výsledky byly relevantní. Výsledky analyzovaných textů se zobrazují mnoha způsoby. Signifikantní slova jsou označena červeně. Ve výuce jazyka má i KWords své uplatnění. Analyzovat se mohou textové druhy a žánry, nebo mohou sami žáci vytvořit vlastní text a porovnávat je s ostatními žáky se stejným zadaným tématem.⁵⁵

Treq se soustředí na prohledávání česko-cizojazyčných, anglicko-cizojazyčných a španělsko-cizojazyčných slovníků, a to v obou směrech. Data získává z paralelního korpusu **InterCorp**. Uživatel si může dohledat překladové ekvivalenty nebo mu jsou přínosem nová synonyma.⁵⁶ Vyhledávání probíhá zvolením výchozího a cílového jazyka. Dále se zadává hledaný výraz. Aplikace provádí hledání ve všech textech InterCorpu. Treq se dá využít v hodinách jazyka při práci s cizojazyčnými ekvivalenty. Nelze ho však zařadit mezi slovníky, neboť v něm chybí řada lexikografických informací. Výhodou však je jeho propojení s autentickými texty.⁵⁷

2.6 Pro školy

Na webové stránce ČNK najdeme odkaz určený pro školy, kde se nachází repozitář cvičení pro výuku jazyků. Tento **repozitář slouží pro učitele na ZŠ i SŠ**, kteří chtějí mít své jazykové hodiny zajímavější a pestřejší. Cvičení jsou založena na aktuálním jazykovém úzu a učitelé v nich mohou nacházet nejmodernější poznatky o jazyce.⁵⁸

⁵⁵ ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019, s. 67-70.

⁵⁶ Cvrček, Václav - Richterová, Olga (eds). *manualy:treq* [online]. Příručka ČNK; 2020. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:treq&rev=1597072085>.

⁵⁷ ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019, s. 70-71.

⁵⁸ Cvrček, Václav - Richterová, Olga (eds). *manualy:proskoly* [online]. Příručka ČNK; 2017. [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:proskoly&rev=1487855476>.

Výhody korpusových cvičení:

1. vycházejí ze skutečného jazykového úzu (oproti učebnicovým cvičením, která často obsahují vymyšlené modelové věty);
2. jsou pro žáky i studenty atraktivní (vzhledem k využití počítače a online aplikací, které využívají i názornou grafiku, např. SyD);
3. umožňují přemýšlet o jazyce a jeho zákonitostech jinak (třeba díky srovnání rozdílností mluveného a psaného jazyka nebo s ohledem na typická slovní spojení, tzv. kolokace);
4. jsou v souladu s nejnovějšími trendy ve výuce, tzv. DDL (data-driven learning);
5. využívají mnoho typů textu, od dětské literatury přes publicistiku až po odbornou literaturu;
6. zahrnují i příklady ze spontánní mluvy v různých regionech (na základě přepisů autentických rozhovorů nahraných po celé zemi).⁵⁹

Učitel může korpusy využít ve výuce různými způsoby a formami. Pokud má k dispozici počítač/tablet/mobil nebo projektor a názorně tak s pomůckami pracuje v hodině. Nebo si sám vytvoří předem připravené korpusové cvičení na určitou jazykovou problematiku. Cvičení lze stáhnout a vytisknout žákům na papír, tzv. žáci pracují s pracovním listem. Repositář poskytuje zmíněné typy cvičení.⁶⁰

Typy cvičení

1. *pozorování*: cvičení zaměřené na rozpoznání významu a užití slova či více slov v kontextu na základě vybraného souboru vět z jazykového korpusu;
2. *porovnávání*: cvičení zaměřené na srovnání významu a užití dvou a více slov (např. synonym), případně více významů jednoho slova (na základě paralelních korpusů);
3. *doplňování*: cvičení zaměřené na správné určení vhodného chybějícího slova nebo slov na základě kontextu;
4. *spojování*: cvičení zaměřené na vyhodnocení kontextu, v němž se slovo používá, na základě spojování levého a pravého okolí slova;
5. *mix*: cvičení kombinující více typů;
6. *psaní*: cvičení zaměřené na písemný projev;
7. *mluvení*: cvičení zaměřené na ústní projev;

⁵⁹ Cvrček, Václav - Richterová, Olga (eds). manualy:proskoly [online]. Příručka ČNK; 2017. [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manualy:proskoly&rev=1487855476>.

⁶⁰ Tamtéž, [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manualy:proskoly&rev=1487855476>.

8. *zařazování: cvičení zaměřené na volbu adekvátního užití slova či slovního spojení v kontextu, typu jazyka či určitém textovém typu;*
9. *hledání: cvičení využívající přímo online jazykový korpus s cílem najít v textech zvolené slovo či slovní spojení a zjistit tak co nejvíce informací o jeho použití.*⁶¹

Stránka Pro školy se zaměřuje na výuku mateřského jazyka, tedy českého jazyka na českých školách, ale i na výuku cizího jazyka, zejména češtiny pro cizince a dalších jazyků, které se u nás vyučují. Učitelé si mohou na stránkách vyhledat připravená cvičení a tipy, jak využít korpusy ve vyučování. U cvičeních je vždy znázorněno, co je nezbytně vyžadováno (počítač/tablet/mobil nebo papír a tužka). Vhodným domácím úkolem může být on-line korpusová cvičení, která učitel žákům zadá a oni sami pracují s korpusovými aplikacemi. Žáci interaktivní typ úloh baví a motivuje k učení. Při vyhledávání vhodného cvičení jsou k dispozici štítky, které filtrují obsah. Lze si tak vybrat nejen jazyk, ale i věkovou kategorii, jazykovou oblast, úroveň a typ cvičení. Korpusová cvičení pochází od P. Vališové a její klasifikaci cvičení pro L2, zároveň jsou obohaceny o specifická cvičení pro výuku mateřského jazyka.⁶²

2.7 Korpusy ve vyučování

V zahraničí učitelé velmi často aplikují při vyučování **Computer Assisted Language Learning (CALL)**, mimo další aplikace zahrnují do výuky také použití jazykových korpusů, kde žáci mohou získat praktické zkušenosti s korpusy prostřednictvím zadaných úkolů a materiálů. Tzv. induktivní přístup se v tomto případě opírá o schopnost najít problematiku daného jazykového jevu, formulovat obecné pravidlo a prakticky aplikovat. Obecně se tato aktivita nazývá „učení založené na datech“ (**DDL, data driven learning**). Proces DDL je vnímán jako „učení se s údaji“ a z žáka se stává „lingvistický výzkumník“. DDL se stalo mezi učiteli jazyků velmi populárním a využívaným. Mnozí z nich své materiály zveřejňují a zpřístupňují ostatním. S DDL je spojována konstruktivistická teorie učení, komunikativní přístup k výuce jazyků a vývoj v oblasti autonomie studentů. Studenti si při procesu mnohem více rozšiřují slovní zásobu.⁶³

⁶¹ Cvrček, Václav - Richterová, Olga (eds). *manuály:proskoly* [online]. Příručka ČNK; 2017. [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manuály:proskoly&rev=1487855476>.

⁶² Tamtéž, [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=manuály:proskoly&rev=1487855476>.

⁶³ CARTER, Ronald, O' KEEFFEE, Anne, McCARTHY, Michael. *From corpus to classroom: Language use and language teaching*. UK: Cambridge university press, 2007, s. 24.

Když jsou studenti seznámeni s korpusem, pracují v korpusem s autentickými texty a setkávají se tak se „skutečným“ jazykem. Pro pedagogické účely se ve většině kontextů výuky a studia jazyka vyučuje pomocí korpusem, kde jsou autentické příklady, které nejsou vymyšlené nebo nereálné, jako to bývá v učebnicích.⁶⁴

Korpusová data mohou učitelé aplikovat při svém vyučování, např. pro výuku slovní zásoby. Vhodným úkolem je zaměření na frekvenční analýzu. Ve slovní zásobě jazyka najdeme spojení mezi frekvencí výskytu slova a jeho pořadím.⁶⁵

Dle výzkumů vyplývá, že nástroje a metody korpusové lingvistiky jsou úspěšné ve výuce jazyků. Korpusy mají potenciál při uplatnění učení se mateřského jazyka na základních nebo středních školách. Sealey a Thompson (2004) vedli výzkum v rámci projektu CLLIP (Korpus učení o jazyce na základních školách), kde závěry uvádějí, že korpusy se mohou úspěšně použít i u velmi malých dětí.⁶⁶

⁶⁴ CARTER, Ronald, O' KEEFFEE, Anne, McCARTHY, Michael. *From corpus to classroom: Language use and language teaching*. UK: Cambridge university press, 2007, s. 26-27.

⁶⁵ ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019, s. 75.

⁶⁶ ČERMÁKOVÁ, Anna, CHLUMSKÁ, Lucie. *What are the suitable corpus data to support L1 teaching (not only) in Czech?* [online]. In: Teaching and language corpora conference 20. - 23. July, Univerzita Giessen, Německo, 2016, s. 20 [cit. 2021-01-24]. Dostupné z: <https://www.uni-giessen.de/faculties/f05/engl/ling/talc/home/programme/abstracts>

3 Korpusová lingvistika

3.1 Korpus a jazyk

Korpus, latinsky psané *corpus*, znamená „*tělo, těleso*“, později se však užívalo i pro jiné věci mimo obor lingvistiky (např. v armádě, diplomacii, historii, kulinářství aj.). V lingvistice se poprvé tento termín začal užívat od 50. let 20. století, a to v angličtině jako označení pro soubor elektronických textů, který je zpracovaný softwarovými nástroji v počítači. Pro odlišení významu se také používá výraz **jazykový či textový korpus**.⁶⁷ Z hlediska lingvistiky označuje korpus soubor autentického materiálu přirozeného jazyka. Jedná se o **materiální základnu**, která má své využití v lingvistice při analýze a popisu jazyka psaného, tak i mluveného. Mohli bychom tak označit řadu sbírek lingvistického materiálu, které dříve vznikaly ručně, obvykle se užíval výraz např. lístkový katalog. Při vstupu počítačů do vědy a běžného využití v domácnosti se mnohem rychleji snížila náročnost ručního zpracování a texty se začaly elektronizovat. Tím vznikl základ pro lingvistický výzkum. Vývojem programu bylo snadné během pár vteřin nalézt požadované slovní tvary. Elektronický korpus tak postupně získal své uplatnění ve výzkumné i aplikované lingvistice. Z kartotéčních lístků přešli lingvisté na počítačové zpracování a využili jeho potenciál. Elektronické korpusové projekty našly řadu uplatnění, dnes se termín korpus převážně užívá pro elektronicky uložené soubory textů (počítačově zpracované).⁶⁸

Korpus se nachází v **jednotném formátu**, aby se v něm mohl uživatel jednoduše orientovat a vyhledat jazykové jevy, slova a slovní spojení. Poskytuje dané jazykové jevy v přirozeném kontextu a umožňuje na základě reálných dat jazykový výzkum. Kromě poskytnutí jazykového materiálu je k dispozici i frekvence jevů.⁶⁹

Od běžného elektronického archivu se korpus odlišuje tím, že nemá omezené možnosti prohledávání. Součástí korpusu je korpusový manažer, který je schopen zpracovat miliardy slov. Korpusový manažer se zaměřuje na určitý cíl, tj. *zpravidla na celý jazyk v odpovídajících proporcích, se záměrem ho textově maximálně pokrýt*. Mluví se tedy o tzv. **reprezentativnosti korpusu**. Většina korpusů má svůj vlastní korpusový manažer, v ČNK uživatel vyhledává za pomoci KonTextu (nahrazující Bonito,

⁶⁷ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 37.

⁶⁸ ŠULC, Michal. *Korpusová lingvistika: první vstup*. Praha: Karolinum, 1999, s. 9-10.

⁶⁹ Cvrček, Václav - Richterová, Olga (eds). *pojmy:korpus* [online]. Příručka ČNK; 2017. [cit. 2021-02-11]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:korpus&rev=1496929366>.

který má základ z manažeru Manatee). Pro angličtinu je to Xaira (BYU-BNC) nebo němčinu Corpus WorkBench nebo Cosmas I a II, některé z nich používají ostatní země i pro své jazyky.⁷⁰

Problematika jazykového korpusu spočívá ve značkování, které zásadně zvyšuje užítí. Základní typ je tzv. vnější anotace, která vychází z praktické informace, z jakého typu, z kterého roku, od jakého autora pochází jednotlivé citace korpusu. Tato data se ukládají do textu a programy při práci s korpusem vše uvádějí v konkordančním řádku hledaného slova. V některých programech si mohou uživatelé zvolit možnosti vyhledávání podle specifických parametrů. Formát, kde jsou potřebné informace ke korpusu, je standardizovaný. Mezinárodně se užívá formát **SGML** (Standard Generalized Markup Language – formální jazyk a metoda pro obecné značkování korpusových textů).⁷¹ Korpusový přístup se uplatňuje v lingvistice, a to ve všech rovinách jazyka. V současné době vznikají výzkumy pro fonetické databáze, které analyzují jevy mluveného jazyka. V gramatice najdeme mnoho studií z morfologie a syntaxe, které vznikly vlivem přístupu **corpus-based** i **corpus-driven**. V lexikologii a lexikografii převládá v korpusovém přístupu převážně metodologie. Kromě tradičních výkladových a překladových slovníků se začaly vytvářet specializované slovníky (frekvenční, kolokační). V současnosti se výzkumy zaměřují na vyšší roviny, jako CADS – Corpus Assisted Discourse Studies (studium diskurzu založené na korpusu) nahlíží na jazyk z přístupu synchronního, tak i diachronního. Při kvalitativní metodě studia diskurzu se uplatňuje CDA – Critical Discourse Analysis. Kromě lexikografie a cizojazyčné pedagogiky se korpusový přístup prosadil v dialektologii, sociolingvistice, psycholingvistice, forenzní lingvistice, při studiu jazykové variability, v jazykové didaktice a studiu akvizice jazyka ad.⁷²

3.2 Korpusová lingvistika

Korpusová lingvistika studuje jazyk prostřednictvím korpusů. Tato poměrně mladá disciplína závisí na rozvoji korpusů tak i na metodologii, je spjatá s reálnými texty (psanými i mluvenými). Korpusy se neustále rozrůstají, dnes se již počítají ve stamilionech a bilionech slov. ČNK se řadí mezi největší a nejlépe zpracované korpusy

⁷⁰ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 38-39.

⁷¹ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 6-7.

⁷² Cvrček, Václav - Richterová, Olga (eds). pojmy:korpus [online]. Příručka ČNK; 2017. [cit. 2021-02-11]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:korpus&rev=1496929366>.

v Evropě. V současné době má korpusová lingvistika k dispozici relativní množství dat, které mohou ovlivnit paradigmatický pohled na lingvistiku k pohledu syntagmatickému.⁷³ Jako obor se začala korpusová lingvistika rozvíjet od 80. let 20. století, s tím souvisí vývoj počítačové techniky, který tomuto oboru přispěl vznikem souborů jazykových dat. Některé menší korpusy vznikly již dříve, **Brown Corpus**, jehož tvůrcem byl lingvista **Henry Kučera**. Teprve rychlejší rozvoj počítačové techniky umožnil širší přístup k informačním možnostem. Korpusově-lingvistický přístup zkoumá texty, kdy texty jsou chápány jako produkty jazykového systému a jazykových schopností jejich tvůrců, prostřednictvím nich dochází k poznání obecnějších jazykových zákonitostí a pravidel. *Od jiných jazykovědných směrů se liší jen důsledným využíváním jazykových dat.* Korpusová data lze charakterizovat jako: *1. typická, nenáhodná a věrná ve vztahu k tomu, jak lidé užívají jazyka; 2. dále jako aktuální, resp. skutečně odrážející svou dobu; 3. neselektivní a objektivní; 4. dostatečná; 5. s pomocí počítače snadno získatelná a rychle přístupná.*⁷⁴

Lingvisté využívají nejmodernější prostředky ke svému jazykovému výzkumu. Během jejich práce s velkými korpusy mohou nalézt jevy, které nebyly popsány. Zájem o korpusovou lingvistiku a korpusy mají především mladí lingvisté a vědci. Zvláště v 2. polovině 20. století poskytla korpusová lingvistika řadu přínosů do všeobecné lingvistiky. Velké množství materiálu poukazuje nejen na to, co všechno je v jazyce možné (variantnost a variabilita jazyka), ale díky počítačovému zpracování (vyhledávání, třídění, počítání) také základní rysy a jevy jazyka, na jejichž základě můžeme s jazykem pracovat i s určitou pravděpodobností výskytu.⁷⁵

Především korpusová lingvistika pokračuje v rozvoji metodologie, *tj. výstavby korpusů i jejich vytěžování, propracování souvisejících oblastí, zvl. aplikovaných (lexikografie, jazyková propedeutika a další), i oblastí nejazykových, a to ve vazbě na úzce související disciplíny, jako je počítačová lingvistika a informatika.* Jedná se tedy o metodologickou disciplínu, která vnáší vhléd a metodologické zkušenosti pro analýzu spolehlivých dat a jejich další studium. Studium poskytuje všem, ať už je zájem odborný, či neodborný. Dostupnost korpusových objektivních informací se stává v lingvistice populárnější. *Korpusová lingvistika navazuje zčásti na poznání a aplikace počítačové*

⁷³ ČERMÁK, František a Renata BLATNÁ, ed. *Korpusová lingvistika - stav a modelové přístupy*. Praha: NLN, Nakladatelství Lidové noviny, 2006, s. 9.

⁷⁴ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 4.

⁷⁵ ŠULC, Michal. *Korpusová lingvistika: první vstup*. Praha: Karolinum, 1999, s. 10-11.

lingvistiky. Důležitou činností se stala automatická anotace, tagování textů. Zakládá se na manuálním označování malého výchozího korpusu. Větší část jazykových jevů a forem se objevila již v malém korpusu (etalonu) a dá se zobecnit pro celý velký korpus. Ve velkých korpusech je značkování přeneseno automaticky. Tato dlouhodobá činnost, která kvůli přínosu nových slov do korpusů nikdy nekončí. Dodnes není úplná shoda co do výhodnosti anotace korpusových textů. Ta odráží zvláště syntaktickou teorii a její nedostatky.⁷⁶

3.3 Historie korpusové lingvistiky

Korpusová lingvistika, v elektronické podobě, se datuje od 60. let 20. století. Předěšlá generace lingvistů pracovala s autentickými texty, ale jejich zpracování bylo ve formě kartotéky, archivu či jiné. Již v 17. století citaci využívali zejména lexikografové. Samuel Johnson se svými spolupracovníky vytvořil 40 tis. slovníkových hesel, *Dictionary of English Language*. Další velký soubor citací shromáždil Noah Webster ve svém slovníku *Dictionary of the English Language* (1828). V 19. století se lexikografové ve Velké Británii zabývali dialekty, kde zkoumali psanou podobu slova a výslovnost, své práce sestavovali do korpusů a publikovali pod názvem *English Dialect Dictionary* (1898-1905) a *The Existing Pfonology of English Dialects* (1898). Na začátku 20. století měla zájem o korpusy jazyková pedagogika. Sestavený korpus E. L. Thorndika obsahoval 4,5 mil. slovních výskytů a své výsledky publikoval v knize *Teacher's Wordbook* (1921).⁷⁷

3.3.1 Hlavní světové korpusy

Klasifikace korpusů se rozděluje na korpusy první, druhé a třetí generace. **Brown Corpus** (*Brown Corpus University Standard Corpus of Present-Day American English*) se stal jedním z prvních počítačově zpracovaných korpusů, vytvořen byl N. Francise a H. Kučerou v 60. letech 20. století v USA. Jeho kapacita obsahuje 500 vzorků z textů různých žánrů, které byly publikovány v roce 1961. Tento korpus dodnes používají lingvisté jako zdroj dat a je základem pro tvorbu dalších korpusů. Inspirací se stal pro britský korpus **LOB** (Lancaster-Oslo-Bergen Corpus), který vznikl v 70. letech 20. století. Představitel druhé generace korpusů je **British National Corpus** (BNC), který reprezentuje současnou psanou a mluvenou britskou angličtinu. **Bank of English**,

⁷⁶ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 90-91.

⁷⁷ ŠULC, Michal. *Korpusová lingvistika: první vstup*. Praha: Karolinum, 1999, s. 28-29.

původní jméno korpusu Cobuild, založil J. Sinclair. Korpus se stal nereprezentativním a synchronním, avšak v roce 2005 dosáhl rozsahu 525 mil. slov, poté se přestal rozvíjet. Novější **Corpus of Contemporary American English (COCA)** působící v letech 1990- 2011 je největším americkým nereprezentativním synchronním psaným korpusem s rozsahem 450 mil. slov. Skládá se z 5 žánrů, mechanicky vymezených (jazyk mluvený, próza, časopisy, noviny a akademická próza). **International Corpus of English** mapuje mluvený jazyk z 24 zemí, kde se angličtina používá jako mateřský či úřední jazyk. **DWDS** (Digitales Wörterbuch der Deutschen Sprache) má kapacitu 5 slovníků s vazbou na 15 podkladových korpusů. **DeReKo** (Deutsches Referenzkorpus) se svou velikostí 24 mld. slovních tvarů se stal největším korpusem německých textů na světě. Obsahuje beletristické, odborné, populárně naučné i publicistické texty, z důvodu autorských práv jsou zpřístupněny pouze některé texty. **Frantext**, smíšený korpus francouzštiny, se skládá z více než 4 000 literárních neanotovaných textů z období středověku až po současnost. Tento korpus se začal budovat v 60. letech 20. století a stal se základem pro slovník *Trésor de la langue française*, nyní má i elektronickou podobu. Ve třetí generaci (po roce 2000) vznikly velmi velké korpusy pro více jazyků. Mohli bychom jmenovat: **Oxford English Corpus**, **Slovenský národní korpus**, **Slovinský národní korpus FIDAPlus**, **Polský národní korpus**, **Ruský národní korpus**, italský korpus **CORIS/CODIS**, španělský korpus **CREA**, **Referenční korpus současné portugalštiny**, **Korpus současné litevštiny**, **Maďarský národní korpus**, **Chorvatský národní korpus**, **Korpus současné srbštiny** a celá řada dalších evropských i mimoevropských korpusů.⁷⁸

3.4 Metodologie

Pro rozvíjející metodologii korpusové lingvistiky je charakteristické, že se v ní nachází pracovní zásady a postupy reagující na novou problematiku spjatou s analýzou korpusu. V začátcích musela metodologie reagovat na limitující kapacitu hardwaru. Lingvisté mají značně velkou možnost výzkumu korpusových dat, výzkumy se převážně orientují na postupy, které jsou metodologicky alespoň zčásti prověřené. V korpusové lingvistice je nezbytnou součástí statistika. Korpus můžeme zkoumat několika formálními způsoby a metodami, do jisté míry to ovlivňuje povahu dat a užití softwaru (korpusový manažer). Korpusová data mají široký rozsah, mohou být nepřehledná

⁷⁸ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 84-88.

a potřebují být řazena. K primárnímu přístupu dat má kdokoliv možnost se dostat, jedná se o metodu statistickou, formální povaha, od které se odvíjí analýza kvalitativní.⁷⁹

Velikost dnešních korpusů stále roste, a tím vzrůstá potřeba hledat a rozvíjet nástroje, které dokážou výsledky vyhledat a dále zpracovat. Uživatel je schopen určité množství výsledků analyzovat, pokud ale výsledná konkordance obsahuje tisíce a více výskytů, jeho vyhodnocování bude zdlouhavé. Dochází tedy k formálnímu problému korpusové metodologie. Další problematikou se stává technické a lingvistické zpracování textů. V technické podpoře se texty označují do žánrů a svých jednotlivých složek (od názvu, resp. hlavičky přes slova až po interpunkci) v rámci mezinárodní normy **TEI** (Text Encoding Initiative) a značkovacím jazykem **XML** (Extensible Markup Language, vzniklým z SGML). Jednou z hlavních chyb jsou textové chyby (autorské či redaktorské), které se v textu vyskytují a značkovací analyzátor (tagger) není schopen tyto chyby řešit. Opravovat tyto chyby nelze kvůli rozsahu korpusu. Problémem v lingvistické části metodologie je přístup k textovému tvaru. Lemma, tj. zobecněný tvar lexému, které je přiřazováno automaticky ke všem tvarům, může význam a speciální funkcí specifického tvaru potlačit. Proto korpusová lingvistika klade větší důraz na studium tvarů.⁸⁰

Korpusová metodologie musí řešit kromě formálních témat i témata lingvistická, kam patří:

1. *kombinace a víceslovnost jednotek;*
2. *jejich variabilita;*
3. *škálovitost výsledků a v oblasti interpretace;*
4. *problém vzorků;*
5. *perifernost některých jevů;*

všechna jsou svou povahou zásadně nová, dřívější lingvistika je neřešila, popř. nemusela řešit.⁸¹

3.4.1 Korpusová data

Lze odlišovat povahu jazykových dat (data vnější a hrubá) před jejich vložením do korpusu a jejich vložení do něj (data vnitřní a strojově čitelná). Zdrojem korpusových dat jsou texty mluvené i psané. Záznam mluveného jazyka a jeho převod do elektronické

⁷⁹ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 92-93.

⁸⁰ ČERMÁK, František a Renata BLATNÁ, ed. *Korpusová lingvistika - stav a modelové přístupy*. Praha: NLN, Nakladatelství Lidové noviny, 2006, s. 12-13.

⁸¹ Tamtéž, s. 13.

podoby je velmi nákladný. Tedy psaných textů se nachází v korpusu mnohem více. Zdroje dat se člení do třech druhů, nejlevnější a nejsnadnější jsou data v podobě **elektronických sazeb textů**, kterou uplatňují v centrálních novinách, časopisech a některých nakladatelství. Další možností je načítání textů, jejich **skenování** pomocí scannerů. Vyžaduje se velmi specializovaný software. Složité texty obsahují řadu chyb a je nutno mnoho oprav. Třetí možností je **manuální přepisování** textů do počítače. Ve všech těchto možnostech je zapotřebí kontroly a oprav. Součástí sběru dat bývá jeho uvedení s copyrightem, autorskými právy vydavatele či autora.⁸²

Korpusová data jsou pro lingvistické výzkumy velmi hodnotné, a to z důvodů: data pochází ze skutečného úzu (mluveného i psaného); rozsah materiálu v korpusu je bezprecedentní; korpus není budován selektivně; korpus je nespécifickým zdrojem informací o jazyce; mohou se provádět výzkumy, které by s jiným jazykovým materiálem nebylo možno provést; dostupnost dat je k dispozici každému uživateli; mohou se provádět i složité dotazy; nespécializované korpusy v ČNK jsou statické, uživatel má možnost dotazy zpětně ověřit v řádu několika měsíců i let; data z subkorpusů lze rozčlenit podle údajů textu; mapování jazykového vývoje slouží jako archiv.⁸³

3.4.2 Zpracování dat

Před vstupem textů do elektronického zpracování a následně do korpusu musí texty projít několika fázemi:

1. *konverze a standardizace textů;*
2. *selekce (výběr) textů;*
3. *tokenizace dat*
4. *propojení dat s (klasifikačními) atributy vnějšími (vnější anotace)*
5. *propojení dat s (anotačními) atributy vnitřními (vnitřní anotace, značkování a lemmatizace)*
6. *propojení dat, už „vybavených“, s indexy.*⁸⁴

⁸² ČERMÁK, František (ed.). Jazykový korpus: Prostředek a zdroj poznání. *Slovo a slovesnost* [online]. 1995, roč. 56, č. 2 [cit. 2021-02-14] Dostupné z: <http://sas.ujc.cas.cz/archiv.php?art=3627>

⁸³ CVRČEK, Václav, KOVÁŘÍKOVÁ, Dominika (eds.). Možnosti a meze korpusové lingvistiky. *Naše řeč – Základní informace* [online]. 2011, roč. 94, č. 3. [cit. 2021-02-14] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8191>

⁸⁴ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 30.

Při konverzi probíhá sjednocování, resp. standardizace a další zpracování, částečně automaticky, ale i manuálně. Texty různého druhu se musí konvertovat, tak i sjednotit z různých formátů do jednotného formátu. Podle kritérií se selektují texty do zvolených korpusů. Nutné je vytřídit texty podle zvolených procentuálních proporcí a následně je sestavit tak, jak má korpus vypadat a sloužit uživateli.⁸⁵ V tokenizaci dat se jedná o rozčlenění na jednotky (pozice), které jsou předmětem další analýzy.⁸⁶ S tím souvisí čištění dat od nežádoucích příměsí (obrázků, tabulek apod.). Proces tokenizace můžeme nazvat jako jeden z druhů segmentace. Vnější anotace textů se zakládá na informacích o textu, autorech a typech textů v podobě bibliografických a sociolingvistických metadat. Při automatickém značkování se přidává k textu vnitřní anotace, tj. lingvistická informace. Anotace se skládá z morfologického značkování a lemmatizace. *Morfologické značkování (tagging) přiřazuje optimálně tvarům textu podle jejich podoby a funkce jejich morfologickou platnost v daném místě textu. Přiřazení neutrální a zpravidla konvenční reprezentativní podoby flektovaným slovům (lemma) nazýváme lemmatizace, vychází z morfologického značkování (např. nominativ sg. pro jména a infinitiv pro slovesa).* Zpracovaná data jsou připravena pro indexaci, tj. zpracování korpusových textů do podoby korpusového manažeru.⁸⁷

3.5 Typy korpusů

Korpusy vznikají podle **typologie** a **potřeb uživatelů**. Podle jazyka a situace se dají dále modifikovat.⁸⁸ Počet korpusů lze dělit podle charakteristik, např. na korpusy velké, střední a malé, nebo také na korpusy budované s určitým cílem a korpusy obecně reprezentativní aj. Základní členění se dělí na lingvistické, kam patří korpusy psané a mluvené, a na synchronní a diachronní.⁸⁹

Dělení korpusů podle následujících hledisek:

1. *Počet jazyků – a) jednojazyčný korpus, b) vícejazyčný, paralelní korpus*
2. *Téma – a) obecný korpus, b) specializovaný korpus*
3. *Modus – a) psaný korpus, b) mluvený korpus*
4. *Časový záběr – a) synchronní korpus, b) diachronní korpus*

⁸⁵ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 30.

⁸⁶ OSOLSOBĚ, Klára. *Česká morfologie a korpusy*. Praha: Karolinum, 2014, s. 9.

⁸⁷ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 31-33.

⁸⁸ Tamtéž, s. 74.

⁸⁹ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 7.

5. *Dokumentace – a) synchronní korpus, b) archivní korpus*
6. *Účel – různý.*⁹⁰

Psaný korpus je nepoužívanější typ korpusů, poskytuje velký vzorek pro psaný jazyk a pro stanovení vědeckého cíle. Tento typ korpusu se stal nejrozšířenějším z důvodu, že psaný jazyk je dostupný v elektronické podobě (např. tisk, knihy). Snadno se tyto materiály získávají a jsou technicky snadno zpracovatelné. Psané korpusy mají nejrozsáhlejší kapacitu a umožňují různé typy jazykovědného výzkumu. Dělí se na synchronní a diachronní, popř. i na žánrově specifické.⁹¹

Mluvený korpus se řadí k novějšímu typu korpusu oproti psanému korpusu. Jejich počet není příliš vysoký a objem dat je mnohem menší než u psaného korpusu. Identita mluveného korpusu není specificky daná. Co všechno může mluvený korpus být, se v různých názorech lingvistů mění. Hlavním zřetelem je vymezení cíle a funkce. Nejde tedy pouze o nahrávku z rádia, televize apod. U získané nahrávky (např. přednášky, veřejného projevu atd.) nás nezajímá pouze intonace a výslovnost, není tedy cíl obecný a taková nahrávka do mluveného korpusu patří. Cíle výstavby mluveného korpusu mohou být různé. Typy mluvených korpusů dělíme na specializované a obecné. Zdroje mluvených dat nejsou tak snadno dostupné. Rozhlas, televize nabízí typy mluveného textu, ale většina není spontánní a v češtině není ani charakteristická lingvisticky. Záleží to však na povaze jazyka a konkrétním médiu. Náročnost vytěžování dat závisí na plánované funkci korpusu, jeho struktuře a strategii sběru dat.⁹²

Označení **synchronní korpus** se často používá v oblasti psaného korpusu (obecné i speciální). Tento typ korpusu zachycuje jazyk současnosti různého typu (včetně mluveného). Cílem synchronního korpusu je zaznamenat živý jazyk, tak jak se skutečně užívá ve společnosti. Korpus obsahuje záznam psaného jazyka, tak i záznam mluveného jazyka. Současně dochází ke změně jazyka a jeho uživatelů, tím se mění v čase a ve své povaze, a přechází do diachronie a do diachronního korpusu. Charakteristikou **diachronního korpusu** je mapování a zaznamenávání jazyka v minulosti, minulé stavy jazyka. Na něho navazuje psaný korpus synchronní. V diachronním korpusu najdeme pouze psané texty (akustické zdroje starého jazyka neexistují). Data tak pokrývají celou

⁹⁰ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 74.

⁹¹ Tamtéž, s. 74-75.

⁹² ČERMÁK, František a Renata BLATNÁ, ed. *Korpusová lingvistika - stav a modelové přístupy*. Praha: NLN, Nakladatelství Lidové noviny, 2006, s. 53-55.

jazykovou historii. Žánrově nejsou diachronní texty stejné. Mnohem širší nabídku textů nabízí synchronní korpus. Některé žánry textů z počátku písemného záznamu chybí, popř. vznikly v průběhu času jako texty publicistické. Grafická podoba diachronních textů se v průběhu času měnila a má tedy více podob. Při přepisu těchto textů do diachronního korpusu se přepisovatel snaží o zachování dobové podoby. Součástí textu jsou bibliografické značky (údaje o autorovi, edici, čase apod.).⁹³

Paralelní korpus je tvořen ze dvou a více jazyků vytvářených z překladů. Obsahuje texty originální i jejich jinojazyčné mutace.⁹⁴ Dříve pojem paralelní korpus označoval všechny typy vícejazyčných korpusů. V současnosti se užívá ve spojitosti s vícejazyčným korpusem, ve kterém jsou obsaženy originální texty v jednom jazyce a jejich překlady do jiného jazyka či jazyků. Rozlišuje se paralelní korpus dvoujazyčný nebo vícejazyčný. V závislosti na překladu mluvíme o korpusem jednosměrném (např. jen anglické originály a jejich české překlady) a obousměrném (např. i české originály a jejich anglické překlady).⁹⁵ Práce s těmito korpusy se stává více žádoucí, výzkum prokázal, nakolik jsou různorodější cizojazyčné ekvivalenty používané v porovnání s klasickými dvoujazyčnými slovníky. Zároveň výsledky ukazují přínos do budoucna v obohacení těchto slovníků v paralelním korpusem. Autentický překladový materiál je další vhodnou pomůckou pro studenty při výuce jazyka.⁹⁶

3.6 Přístup corpus-based a corpus driven

Lingvistický výzkum korpusu má dvojí charakter, jak a do jaké míry chce záviset na datech získaných z korpusu. Buď pracuje pouze s daty z korpusu, jde tedy o přístup **corpus-driven** (korpus řízený). Jedná-li se však o výzkum závislý na datech jen podpůrným způsobem, kdy uživatel data z korpusu pouze doplňuje, mluvíme o přístupu **corpus-based** (na korpusem založený). Oba přístupy se však často vzájemně doplňují. První přístup je náročnější, avšak poskytuje celá data. Druhý přístup může být více subjektivní.⁹⁷

⁹³ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 76-81.

⁹⁴ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 9.

⁹⁵ Cvrček, Václav - Richterová, Olga (eds). pojmy:paralelni [online]. Příručka ČNK; [cit. 2021-02-18]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:paralelni&rev=1566383701>.

⁹⁶ KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000, s. 9.

⁹⁷ ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017, s. 100.

V přístupu corpus-based již uživatel pracuje s daty s předem vytvořenou hypotézou a hledá důvody pro jejich potvrzení či vyvrácení. S přístupem corpus-driven uživatel formuluje hypotézu o jazyce na základě jazykových dat. Tradičně se u nás více setkáváme s přístupem corpus-based, kde uživatel pracuje především se svými znalostmi. Při druhém přístupu však uživatel je ochoten svou hypotézu pozměnit či upravit a na základě získaných dat vytvořit hypotézu znovu a jinak.⁹⁸

3.7 Reprezentativnost

Pojem reprezentativnost se používá ve statistice a v empirických vědách. Základem se stává vztah mezi vzorkem (reálná množina jednotky) a populací (množina jednotek), na kterou se poznatky o vzorku zobecňují. Můžeme tvrdit, že vzorek, který se v určitých charakteristikách podobá populaci, je reprezentativnější. V korpusové lingvistice představuje vzorek určitý korpus nebo jeho část. Složitější je však určení populace. **Douglas Biber** tvrdí, že je nutné při tvorbě korpusu určit faktické hranice populace (např. všechny texty publikované v angličtině v roce 1961). Musíme také zjistit stratifikaci populace, jaké žánry či registry se v ní vyskytují. Z tohoto pak vybíráme určitý počet textů o stanovené délce. Biber však odmítá koncept, že by korpusy měly být tvořeny na základě demografických proporcí. „*Říká, že jazykové korpusy potřebují jiné vymezení reprezentativnosti a že badatelé potřebují vzorky, které jsou reprezentativní v tom smyslu, že zahrnují úplné spektrum jazykové variace, které v jazyce existuje.*“ Není důležitá statistická reprezentativnost, ale zastoupení textových zdrojů v korpusu. S touto teorií nesouhlasí **Váradí**, mluvící o reprezentativním vzorku, pro který je nutností znalost celé populace, jež není zcela k dispozici. „*Reprezentativní vzorek z populace můžeme získat pouze prostřednictvím charakteristik této populace, které jsou doloženy z nějakého nezávislého zdroje. Takovým zdrojem nezávislých poznatků jsou dostupná data o mluvčích.*“ Pojmem reprezentativnosti se u nás v poslední době zabýval **Křen**, který se přiklání k teorii **Sinclaira**. „*Korpus je pouze více či méně věrným odrazem, vzorkem jazyka nebo některé jeho variety, ale nikdy ho nezastoupí celý. Není to ani nutné, je-li ovšem korpus reprezentativní, a odpovídá tak v mnohem menším měřítku celku jazyka (variety). Hlavním problémem, který se k reprezentativnosti váže, je však právě nemožnost objektivně měřit, nakolik korpus jazyku odpovídá.*“⁹⁹

⁹⁸ CVRČEK, Václav, KOVÁŘÍKOVÁ, Dominika (eds.). Možnosti a meze korpusové lingvistiky. *Naše řeč – Základní informace* [online]. 2011. roč. 94, č. 3. [cit. 2021-02-14] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8191>

⁹⁹ CHROMÝ, Jan (ed.). Korpus a reprezentativnost. *Naše řeč – Základní informace* [online]. 2014. roč. 97, č. 4-5. [cit. 2021-02-15] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8337>

Vliv žánrů není však jediný problém, který může ovlivnit možnosti výzkumu a jeho následnou interpretaci výsledků. V korpusové translatologii se v posledních letech mluví o vlivu, zda je jazyk zkoumaných textů jazykem originálu, nebo zda je jazyk pozměněn v důsledku překladu. Výsledky výzkumu překladovosti ukázaly, že skutečně překladový jazyk může mít odlišné vlastnosti (např. tendence k explicitaci, ilustrovanou na výzkumu norštiny a angličtiny). Rysy, které jsou charakteristické pro překlad, se liší v závislosti na zkoumaném jazyce překladů, ale i v originálu nebo v kombinaci s vlivem žánrů.¹⁰⁰

¹⁰⁰ CHLUMSKÁ, Lucie, RICHTEROVÁ, Olga (eds.). Překladová čeština v korpusech. *Naše řeč – Základní informace* [online]. 2014, roč. 97, č. 4-5. [cit. 2021-02-15] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8343>

4 Lexikologie

4.1 Předmět lexikologie

Lexikologie se zaměřuje na slovní zásobu, popisuje lexikální jednotky různých typů, slova, slovní spojení a frazémy, jejich vztahy a jednotlivé systémy. Termín slovní zásoba se užívá v užším slova smyslu. Základní dělení lexikologie se člení na synchronní, historickou, a to lexikologii jednoho jazyka a více jazyků (historickosrovnávací), a předhistorickou (etymologii). V lexikologii synchronní mluvíme o lexikální sémantice. Sémaziologie, která se jinak nazývá teorie označování a pojmenování. Onomaziologie a další části plynou z jejich spojení a týkají se vztahů lexikálních jednotek, jejich tříd a dílčích systémů a lexikální stylistiky. Vztah mezi lexikologií a lexikografií je dialektický. Každá teorie lexikální zásoby nachází verifikaci v slovnících daného typu. Metajazyk lexikologie se buduje na jazyku lexikografie.¹⁰¹ Zájem lexikologie se vztahuje k **vnitrojazykovým sémantickým vztahům** mezi lexikálními jednotkami, jako je polysémie s homonymií, synonymie, hyponymie s hyperonymií, různé druhy antonymie, konverzívnost, kompatibilita s inkompatibilitou; a k **pragmatickým vztahům** mezi lexikálními jednotkami (zejm. synonymními).¹⁰² Lexikologie, nauka o slovní zásobě, se dělí na dílčí disciplíny: sémaziologie, sémantika – nauka o významu; onomaziologie – nauka o pojmenování; frazeologie – nauka o frazeologických jednotkách; etymologie – nauka o původu slov; onomastika – nauka o vlastních jménech; lexikografie – nauka o slovníku a další.¹⁰³

Úkoly české lexikologie plynou z vývoje této lingvistické oblasti a z jejího postavení v dnešním kontextu lingvistiky společně s jejími pomocnými vědami. Názory na lexikologii se různí, časté protikladné názory se týkají hlavních otázek, např. otázka na lexikální jednotky, a to v aspektu výpovědním (výskyt) i systémovém, paradigmatickém, abstrakčním. Pro některé lingvisty je základní jednotkou morfém a výskyt, což má své důsledky ve výpovědní sémantice a pro pojetí sémantických rovin. V synchronním aspektu výzkumu české lexikologie se výsledky prací člení na čtyři zdroje: *1. materiálové zkušenosti získané zpracováním rozsáhlého materiálu pro jednojazyčné (a na jiném úseku pro dvojjazyčné) slovníky; 2. teoretická lingvistická*

¹⁰¹ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 13-14.

¹⁰² ČEJKA, Mirek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992, s. 5.

¹⁰³ HAUSER, Přemysl. *Nauka o slovní zásobě*. 2. vyd. Praha: Státní pedagogické nakladatelství, 1986. Učebnice pro vysoké školy (Státní pedagogické nakladatelství), s. 7.

tradice reprezentovaná především pražskou školou; 3. kritické využití podnětů lingvistické, především strukturně orientované teorie (ženevská škola, kodaňská škola, deskriptivismus, generativní gramatika); 4. kritické kontakty s významnými díly zahraniční lexikologie a lexikografie, především sovětské a francouzské.¹⁰⁴

4.1.1 Metoda sémaziologická a onomaziologická

Termín sémaziologie se užívá pro lexikální sémantiku. Předmětem tohoto zkoumání je lexikální význam. Metodologický postup se zakládá **od formy k významu**, zodpovídá, jaké významy mají formy (morfy, lexémy) v jazyce. Onomaziologie zkoumá teorii označování, její postup je **od významu, popř. funkce k formě**. Lingvista může například analyzovat, jak je pole barev, které známe ze své vlastní mimojazykové zkušenosti, formulováno podsystémem významů tzv. barevných termínů v jistém jazyce. Onomaziologický přístup má své uplatnění u jazykových rovin, jejichž jednotky mají význam. S tímto termínem se nejčastěji setkáme u problematiky lexikologické a slovotvorné.¹⁰⁵

4.1.2 Sémantické vztahy

Mezi lexikálními jednotkami jsou externí jazykové vztahy, na jejich základě tyto jednotky tvoří skupiny, soubory jednotek a dílčí systémy. Vztah mezi nimi se uskuteční se společným prvkem dvou a více jednotek. Vztahy jsou **formální, formálně-významové a významové**. Hlavním úkolem lexikologie je popis fungování systémových jednotek.¹⁰⁶ Struktura slovníku se člení na paradigmatické a syntagmatické významové vztahy. S paradigmatickým významovým vztahem se propojují lexikální jednotky, které ve větě můžeme umístit na totéž místo, v daném kontextu se vzájemně vylučují (kromě synonym). Srov. pojmenování barev: Ve větě *Oblékla si bílý svetr* lze slovo *bílý* nahradit jinými označeními pro barvy, např. *černý, žlutý, zelený*, v kontextu jednoho určitého svetru může platit pouze jedna varianta. Syntagmatické významové vztahy jsou mezi jednotkami, které mají ve větě schopnost se spojovat. Např. slovo *pes* se váže se slovy *štěkat, vrčet, kousat, hlídat...*, na straně druhé se slovy *lovecký, chlupatý, zlý* apod.¹⁰⁷

¹⁰⁴ FILIPEC, Josef (ed.). K úkolům české lexikologie. *Slovo a slovesnost* [online]. 1968. roč. 29, č. 3. [cit. 2021-03-05] Dostupné z: <http://sas.ujc.cas.cz/archiv.php?art=1642>

¹⁰⁵ ČEJKA, Mirek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992, s. 11.

¹⁰⁶ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 128.

¹⁰⁷ KARLÍK, Petr, Marek NEKULA, Zdenka RUSÍNOVÁ a Miroslav GREPL. *Příruční mluvnice češtiny*. Vyd. 2., opr. [i.e. 4. vyd.]. Praha: NLN, Nakladatelství Lidové noviny, 2012, s. 79.

K paradigmatickým významovým vztahům se řadí antonymie, synonymie, polysémie, homonymie, hyperonymie, hyponymie a kohyponomie. Své specifické postavení má polysémie. Všechny tyto vztahy jsou zastoupeny ve všech přirozených jazycích.¹⁰⁸ **Synonymie** představuje relaci významové totožnosti dvou nebo více jazykových výrazů. Za synonymií se považuje sémantický vztah mezi významy lexikálních jednotek, který je oboustranně založen na implikaci. Synonymie se zakládá na relaci symetrické, tranzitivní a reflektivní. Kupříkladu můžeme uvést dvojice jazykových výrazů o různém stupni členění (nejen dvojice): *jazykověda – lingvistika, mluvnice – gramatika, kobyla – klisna*, ale i dvojice *sourozenci – bratři a sestry, vloni – minulého roku*.¹⁰⁹ **Antonymie** vedle synonymie představuje základní významový vztah lexikálních jednotek, který má základ v logickém protikladu, tzv. opozice dvou skutečností. Z hlediska kvantity je to omezený jev, ale sémanticky relativně jednoduchý.¹¹⁰ Antonymie, také zvaná polární, graduální (z hlediska logiky kontrární), je příznačné, že protikladné dvojice, např. *malý – velký, ledový – horký, hluboký – mělký*, nepokrývají významy celého rozsahu nadřazeného pojmu.¹¹¹

Hyponymie (podřazenost), **hyperonymie** (nadřazenost) a **kohyponymie** (souřadnost) jsou sémantické relace mezi lexikálními jednotkami jazyka. *Význam jedné lexikální jednotky může být složkou významu jiné lexikální jednotky*. Lexikální jednotka *kapr* je hyponymem lexikální jednotky *ryba*, lexikální jednotka *ryba* hyperonymem lexikální jednotky *kapr*. Tyto lexikální jednotky se společným hyperonymem (*ryba*) nazýváme kohyponyma (slova souřadná).¹¹² **Homonymie** vzniká z jedné jazykové formy (morfém, slovo, spojení slov, větná výpověď) psané nebo mluvené. Má dvě nebo více funkcí, významů, smyslů, které vzájemně spolu nesouvisí. Rozlišujeme homonymií vlastní (v tvarech jednoho lexému) a nevlastní (u různých lexémů), gramatická a lexikální. Příkladem homonymie může být koncovka *-e* (5. sg. u *pane*, 2. sg. u *meče*, 4. pl. u *muže*, 1. sg. u *může* aj.), sufixy *-č, -ka* u *leštič, leštička*, obojí život. i neživot. (označují činitele, činitelku a lešticí prostředek). Polysémií se označují lexikální jednotky, které představují základní významovou hierarchickou strukturu. Lexikální polysémie má dva a více různých významů, které spolu souvisí. Například adj. *lehký* je uvedeno v pěti

¹⁰⁸ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 128.

¹⁰⁹ ČEJKA, Mírek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992, s. 24.

¹¹⁰ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 129.

¹¹¹ KARLÍK, Petr, Marek NEKULA, Zdenka RUSÍNOVÁ a Miroslav GREPL. *Příruční mluvnice češtiny*. Vyd. 2., opr. [i.e. 4. vyd.]. Praha: NLN, Nakladatelství Lidové noviny, 2012, s. 85.

¹¹² ČEJKA, Mírek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992, s. 26.

významech: *lehký* 1. „mající malou váhu“, 2. „mající něco v malé míře“, 3. „málo namáhavý“, 4. *klidný*, 5. *svižný*, *pružný*.¹¹³

4.2 Slovo, lexém

Základní jednotkou slovní zásoby je slovo, které souhrnně tvoří v každém jazyce jeho slovní zásobu. Zkoumání slova je předmětem nejen v lexikologii, ale i v tvoření slov, tvarosloví a skladbě. Jeho podoba má svou grafickou i zvukovou stránku. Jedná se tedy o komplexní povahu útvaru.¹¹⁴ Slovo je obvykle chápáno jako **diskrétní forma**, tj. celek psaný dohromady (grafické slovo). V lingvistice patří k nejstarším pojmům s proměnlivou historií.¹¹⁵ Slovo může být buď jednočlenný lexém monosémický (jednočlenná lexie), nebo polysémický (polysémické slovo). Rozeznává se slovo potenciaální v systému (lexém) a aktuální v řeči a textu (alolex).¹¹⁶ Forma slova tvoří jeho hlásková stavba, ve většině případů je slovo vytvořeno ze skupiny hlásek (např. *úl*, *oko*, *šest*). Obsahem slova máme na mysli jeho význam, který se vztahuje k určitému úseku skutečnosti prostřednictvím našeho vědomí. Význam slova se váže na daný jazyk, proto slovo v jednom jazyce, nemusí jím být v jiném jazyce (např. *pes* – *čeština*, *Hund* – *němčina*, *sobaka* – *ruština*, *chien* – *francouzština*).¹¹⁷

Lexém označuje systémovou formálně samostatnou, ale funkčně a abstraktně chápanou znakovou jednotku. Formálně může být jednoduchá nebo složená (víceslovná). Tvořena jediným tvarem nebo všemi příslušnými tvary, které k sobě funkčně patří.¹¹⁸

4.3 Rozšiřování slovní zásoby

Slova po vstupu do jazyka se vývojem proměňují, dochází ke změně jejich významu a frekvence. Některá slova se přestávají užívat z důvodu, že zaniká označovaná skutečnost, až postupně vymizí ze slovní zásoby. Pokud se taková slova objeví, nazýváme je tzv. historismy (označování zaniklé reálie), např. názvy historických zbraní (*palcát*, *řemdih*, *halapartna*), názvy už neexistujících společenských hodností a vztahů (*purkmistr*, *purkrabí*). Slovní zásobu zachycují slovníky, ale nelze zachytit veškerou

¹¹³ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 137-141.

¹¹⁴ HAUSER, Přemysl. *Nauka o slovní zásobě*. 2. vyd. Praha: Státní pedagogické nakladatelství, 1986. Učebnice pro vysoké školy (Státní pedagogické nakladatelství), s. 9.

¹¹⁵ ČERMÁK, František. *Lexikon a sémantika*. Praha: Lidové noviny, 2010, s. 17.

¹¹⁶ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 33.

¹¹⁷ HAUSER, Přemysl. *Nauka o slovní zásobě*. 2. vyd. Praha: Státní pedagogické nakladatelství, 1986. Učebnice pro vysoké školy (Státní pedagogické nakladatelství), s. 9.

¹¹⁸ ČERMÁK, František. *Lexikon a sémantika*. Praha: Lidové noviny, 2010, s. 17.

slovní zásobu, jelikož se slovní zásoba neustále mění. Nejrozsáhlejší slovník češtiny, **Příruční slovník jazyka českého**, má velikost 250 000 slov. Individuální slovní zásoba každého jedince je mnohem menší. Aktivní slovní zásoba obsahuje slova, která jednotlivec aktivně používá při mluvení a psaní. V pasivní slovní zásobě se nachází slova, kterým rozumí a je mnohem rozsáhlejší než aktivní slovní zásoba. Rozsah pasivní slovní zásoby bývá individuální, záleží na vzdělání, prostředí, povolání a dalších aktivitách, zejm. četbě. Všeobecně se uvádí, že k běžnému dorozumění postačí 1 500 slov, člověk užívá 3 000 až 10 000 slov.¹¹⁹

Centrum slovní zásoby jazyka se skládá z jádra. V ní jsou slova, která pojmenovávají skutečnosti z běžného denního života a v komunikaci se užívají (např. *syn, voda, slunce, pes*). Jádrová slova náleží k vývojově nejstarším a nejstabilnějším jednotkám. Ze synchronního hlediska patří převážně k nemotivovaným slovům. Počet těchto slov není příliš velký, ale mají vysokou frekvenci výskytu. Málo frekventované lexikální jednotky představují periferii lexikálního systému (např. slova zastarávající – *batalion, kaprál* apod.).¹²⁰

Slovní zásoba celonárodní se pojí s určitým národním a národnostním jazykovým společenstvím. Rozlišuje se v ní **lexikon spisovaného jazyka** a **lexikon nespisovných jazykových útvarů**. Slovní zásoba se vymezuje teritoriálně, a to slova užívaná místně – dialektismy (např. *peřina : duchma, svrchnice, poduška, devětnice*), případně regionalismy. Zvláštní skupinu slovní zásoby tvoří obecná čeština, interdialektický útvar, užívaná především v Čechách. Z hlediska sociálního vymezení je slovní zásoba spjatá s profesní mluvou, slangem a argotem. V časovém hledisku se slovní zásoba rozděluje do slov zastarávajících (archaismy, historismy) a slov nových (neologismy).¹²¹

Slovní zásoba češtiny se obohacuje těmito způsoby:

1. Tvořením nových slov – nová slova se tvoří ze slov, která vznikla odvozováním a skládáním. Odvozování se tvoří přidáním předpon, přípon nebo koncovkou a jejich kombinací. Skládání slov je tvorba dvou základových slov. Další možností se stává také zkracování, vytváření slov z několikaslovných názvů;
2. Přejímáním slov z jiných útvarů národního jazyka a ze slangu – nejprve tato slova pronikají do hovorové vrstvy jazyka;

¹¹⁹ ČECHOVÁ, Marie. *Čeština - řeč a jazyk*. 3., rozš. a upr. vyd. Praha: SPN - pedagogické nakladatelství, 2011, s. 81.

¹²⁰ KARLÍK, Petr, Marek NEKULA, Zdenka RUSÍNOVÁ a Miroslav GREPL. *Příruční mluvnice češtiny*. Vyd. 2., opr. [i.e. 4. vyd.]. Praha: NLN, Nakladatelství Lidové noviny, 2012, s. 92-93.

¹²¹ Tamtéž, s. 93-95.

3. Přejímání slov z jiných jazyků a slov mezinárodních – nově vznikají základní pojmenovací jednotky, slova. Ojedinele se vrací i zastaralá slova (např. historismy – *samopal, tarasnice*);
4. Vznik nových významů slov (sémantické tvoření) – takto se rozšiřuje skupina slov mnohovýznamových a vznikají další významy slov, která byla mnohovýznamová již předtím. Do této oblasti spadá i rozpad mnohovýznamnosti v homonymii, slovo se osamostatňuje (např. *pokoj – klid, pokoj – světnice*);
5. Tvoření víceslovných pojmenovávání – tento způsob pojmenovávání nových věcí je nejčastější pro víceslovné neboli sdružené pojmenování.¹²²

4.4 Frazeologie

4.4.1 Obecné vlastnosti

Obohacovat slovní zásobu lze za pomoci frazeologie. Idiomatika a frazeologie je lingvistická disciplína, která postupem času ukazuje, kde všude její jednotky, idiomy a frazémy můžeme nacházet. Tyto jednotky můžeme vyhledat ve všech jazykových disciplínách. Idiomatika a frazeologie má i vazby na další nejazykové disciplíny. Méně se studuje souvislost s psychologií a logikou. Idiomatika a frazeologie zkoumá popis frazémů a idiomů ve všech úrovních. Vzhledem k početnosti se disciplína zaměřuje na lexémy víceslovné. *Pravidelnými kombinacemi v obecně syntagmatickém smyslu se míní všechny kombinace řízené analogickými pravidly, tj. pravidly opírajícími se o analogii, a tedy kombinace morfémů, lexémů, kolokací i vět.* V oblasti kombinatorických výsledků se idiomatika a frazeologie posouvá v dalších možnostech, tj. do oblasti kombinací, které podle pravidel v jazyce už nemůžou a nemají vzniknout (kombinace anomální). Např. frazém *ležet polem* je formálně anomální a má jinde v jazyce analogicky nenalezitelnou další instrumentálovou konstrukci, ale sémanticky se oba v idiomatice a frazeologii kombinují pouze jediným odlišným způsobem. V kombinatorickém principu můžeme mluvit o vytváření nových pojmenování včetně anomálních, a frazémů a idiomů.¹²³

Termín kolokace zavedl anglický lingvista J. R. Firth, který tímto termínem označoval charakteristická spojení slov. Taková slova mají vysokou pravděpodobnost souvškytu, což je důsledkem jejich vzájemné očekávatelnosti založené na jejich sémantických, nikoliv gramatických vlastnostech. Srov. *pes : štěkat, temný : noc*,

¹²² HAUSER, Přemysl. *Nauka o slovní zásobě*. 2. vyd. Praha: Státní pedagogické nakladatelství, 1986. Učebnice pro vysoké školy (Státní pedagogické nakladatelství), s. 97.

¹²³ ČERMÁK, František. *Frazeologie a idiomatika česká a obecná: Czech and general phraseology*. V Praze: Karolinum, 2007, s. 24-26.

blond : vlasy, vraný : kůň, žluklý : máslo. Monokolokabilní výrazy (např. bacha, pré, hin, duhu, lubem, holičkách, krážem, lehára, bandurskou, mrtě, šlak) nemohou mít samostatný význam, jelikož se nedají v kontextu ničím nahradit a v jiném kontextu neexistují. Od kolokalibity (systémová schopnost jazykové jednotky kombinovat se s jednotkami jinými) se liší kompatibilita jako sémantická spojitelnost jednotek v syntagmatu.¹²⁴

Zájem lexikografie o idiomatiku a frazeologii se nestal nikterak novou záležitostí. V minulosti se oba pojmy chápaly různě, vždy se však ve slovnících vyskytovaly. Byly uváděny jako zvláštní přívažek jednotlivých významů určitých slov. Mělo to řadu nevýhod, ať ve volbě takového slova, pod nímž měl být výraz uveden, tak i s rozlišením daného významu i výrazu celého. V 19. století vznikají první samostatné slovníkové sbírky idiomatiky a frazeologie, některé se používají i v dnešní době. U nás byly vydány první nedokonalé sbírky J. M. Sychry (1821-2), J. Šacha (1862), F. Šebka (1864). Lidová rčení J. Zaorálka (1947) představují cenný lexikografický počín. Tradice však sahá mnohem dál, až do období antiky. Mravní poučení, výstižné zásady, postřehy a myšlenky o životě (autorské či lidové) byly vždy velice žádané. První nálezy pochází se starého Egypta ve sbírce Šalamounových příslovích (Kniha Přísloví) ve Starém zákonu.¹²⁵

Samotná frazeologie je disciplínou, která zkoumá **nepravidelná, ustálená slovní spojení** (např. *hodit něco za hlavu, Achillova pata*). Frazémy v oblasti slovo tvorby, valence a reflektivity, společně s frazémy flektivními, označujeme jako lexikální frazémy. Kolokační anomálie lexikálních frazémů je kombinací formálně i sémanticky nekolokabilních morfémů. Rozlišují se čtyři základní typy lexikálních frazémů: 1. LF v oblasti slovo tvorby, které dále dělíme na frazémy kompoziční (*strašpytel, zčistajasna*) a derivační prefixální (*najít, úskok*), resp. sufixální (*psina, otčím*); 2. LF v oblasti flexe (*panečku, chrám Páně*); 3. LF v oblasti valence (*zahýbat někomu, kašlat na něcolněkoho*); 4. LF v oblasti reflexivity (*ulít se, zahazovat se*). Kompoziční frazémy mají výraznější významovou strukturu než frazémy derivované (jejich komponenty jsou významově relativně samostatné), a sémantická odchylka je proto u nich nápadnější (srov. *rozumbrada* a *podraz*). Lexikální frazémy nejsou ničím ohraničeny od pravidelného lexikonu ani od frazeologie. Míra frazeologičnosti lexémů se odvíjí od formálního stupně a sémantické anomálie. U lexikálních frazémů se vyskytuje

¹²⁴ ČEJKA, Mirek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992, s. 33.

¹²⁵ ČERMÁK, František. *Idiomatika a frazeologie češtiny*. Praha: Univerzita Karlova, 1982, s. 11.

nesamostatnost a častá polysemie jejich komponentů. Frazémy s komponenty vázanými, sémanticky vágními a s nejasnou strukturou řadí lingvisté do oblasti periferie.¹²⁶

Zcela není shodný názor při vymezování podstatných vlastností frazeologických jednotek. *Uvádí se řada těchto rysů, a to zejména ustálenost, obraznost, víceslovnost, nerozložitelnost (celost), expresivnost, jejich „hotovost“ před řečovým aktem, ekvivalentnost se slovem aj.* Avšak ne všechny rysy jsou stejně významné pro vymezení frazeologických jednotek. Obraznost se může skrýt při ztrátě motivovanosti (*mít něco za lubem*), dále se setkáváme s jednotkami, jejichž jádrem je spojení předložky se jménem (*být v tom*). Dochází i k absenci u dalších rysů, např. k stírání expresivnosti (*žlutý anděl*, označení pro pojízdnou službu motoristům na dálnici). V oblasti rozsahu frazeologie, co vše zahrnuje a co už nikoliv, se názory různí. Při širším chápání frazeologie do ní řadíme nejen jednodušší výrazy s funkcí nominace, ale i složená, rozsáhlejší spojení s funkcí komunikativní. Mimo ponecháváme útvary jako aforismy, mota, epigramy, hádanky aj. vyšší jednotky. Výzkum týkající se vztahu dětí a mládeže k frazeologii prokázal pokles ve tradiční části, lidové frazeologii. Zároveň prokázal znalost a rozšíření frazeologie tradiční, lidové, tak i novější, včetně slangové. Změna frazémů u mládeže vzniká z potřeby aktualizovat (*Mluvíti stříbro, mlčeti zlato – nedostatečná, Komu se zelení, od toho se opisuje*).¹²⁷

4.4.2 Idiom a frazém

Frazém je chápán jako sémanticky nečleněný a nečlenitelný celek. Frazeologické jevy lze definovat na většině jazykových rovin, chápeme je jako určité anomální kombinace jazykových jednotek. Frazémy a idiomy rozlišujeme jako dvojí (formální a sémantický) aspekt té samé lexikální frazeologické jednotky. Obecně frazém a idiom definujeme jako nemodelové a ustálené syntagma jazykových jednotek, kdy alespoň jedna je z hlediska členem omezeného a formálně, i sémanticky uzavřeného paradigmatu. Idiomy odpovídají jazykovým prvkům, které jsou u neidiomatických slov kompatibilní. Srov. idiomatický i neidiomatický význam spojení *vypustit pokusný balónek, chytat lelky, chcipl tam pes*.¹²⁸ Frazém a idiom nelze definovat jedinou (anomální) vlastností, ať formální, sémantickou či kolokační. Čím více má frazém anomálií, tím je idiomatičtější

¹²⁶ KLÖTZEROVÁ, Petra (ed.). Hranice frazeologie se posouvají. Lexikální frazémy v češtině. *Slovo a slovesnost* [online]. 1998. roč. 59, č. 4. [cit. 2021-02-21] Dostupné z:

<http://sas.ujc.cas.cz/archiv.php?art=3805>

¹²⁷ ČECHOVÁ, Marie (ed.). Dynamika frazeologie. *Naše řeč – Základní informace* [online]. 1986, roč. 69, č. 4 [cit. 2021-02-21] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=6639>

¹²⁸ ČEJKA, Mírek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992, s. 33-34.

a také naopak. Prvky do frazeologických kombinací v podobě jako komponenty formálně vstupují. Tyto prvky jsou různé podle své členské roviny, kam se zařazují, zvl. morfémy (kořeny nebo afixy), lexémy, kolokace, klauze nebo propozice (věty). Jejich podoba má určitý tvar, nikoli podobu lemmatu. Výsledné frazeologické celky, resp. frazémy jako útvary spadají na vyšší úroveň. *Srov. např. frazém lexikální přestřelit a jeho komponenty pře-střelit (proti pravidelnému pře-hodit apod.), frazém kolokační a jeho komponentní lexikální tvary mít strejčka (proti pravidelnému mít tetičku apod.) nebo frazém propoziční a jeho komponenty To je slovo do pranice (proti pravidelnému To je slovo do diskuse apod.).*¹²⁹

Do formální stránky frazémů patří pevná vnitřní struktura. Každý komponent má ve své struktuře pevnou strukturní pozici. Konstitutivní vlastnosti struktury (pořadí, druh obsazení a minimální počet komponentů) jsou tvořeny základní podobou. Strukturní pozice má danou příslušnost a povahu paradigmatických prvků, které ji realizují a obsazují.¹³⁰

Ve frekvenci idiomů a frazémů se objevují frekventovaná slova, tak i slova nefrekventovaná (řidká). *Srov. hlava dřavá, mít z něčeho těžkou hlavu, druhá míza, vymřít po přeslici.* Méně frekventované slovo může být současně monokolokabilní (např. *nechat na holičkách*). Doménu zde představuje skupina nejfrekventovanějších sloves typu mít, nechat a slov gramatických. *Jednou z mnoha cest, jak útvary idiomů a frazémů vznikají, je i přesun prvku, slova původně běžného na periferii úzu doprovázený rostoucím oslabením, ztrátou či anomálií různých vlastností.*¹³¹

4.4.3 Frazémy nevětné

U frazémů nevětných záleží na konkrétním kontextu věty a jsou v ní gramaticky formovány (např. lev salónů – Stal se lvem salónů). Uspořádání nevětných frazémů není omezeno proti volným slovním spojením (k pomezím případům patří např. některá spojení sloves s abstraktními substantivy: *získat motiv, mít motiv, ztratit motiv*).¹³²

¹²⁹ ČERMÁK, František. *Frazeologie a idiomatika česká a obecná: Czech and general phraseology*. V Praze: Karolinum, 2007, s. 31-33.

¹³⁰ FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické, s. 185.

¹³¹ ČERMÁK, František. *Idiomatika a frazeologie češtiny*. Praha: Univerzita Karlova, 1982, s. 54.

¹³² KARLÍK, Petr, Marek NEKULA, Zdenka RUSÍNOVÁ a Miroslav GREPL. *Příruční mluvnice češtiny*. Vyd. 2., opr. [i.e. 4. vyd.]. Praha: NLN, Nakladatelství Lidové noviny, 2012, s. 71.

Nejčastější strukturní typy nevětných frazémů:

1. frazémy s funkcí slovesnou obsahující sloveso: *měnit barvu, dělat kličky, ukápnout jedu, praštit se přes kapsu, dopadnout bledě*;
2. frazémy s funkcí jména v nominativu, které vznikly kombinací adjektiva a substantiva: *jidášský groš, šedá eminence, růžové brýle*;
3. frazémy s funkcí jména v nominativu vzniklé kombinací substantiv: *hodina duchů, lev salónů, zkouška ohněm*;
4. frazémy s funkcí adverbiální, jež vznikly kombinací substantiva a adjektiva, případně několika substantiv v jiných pádech než v nominativu: *levou rukou, živou mocí, s klidem Angličana, na dosah ruky*;
5. frazémy z ne autosémantických komponentů: *podle mého, pro nic za nic, jako takový, bodejť by ne/jo, s takovou*.¹³³

Ustálená přirovnání patří k frazémům nevětným, mají větnou strukturu, ale některé své pozice uvnitř se umisťují až v konkrétních kontextech (např. *spát jako na vodě – (někdo) spí jako vodě – Dědeček spí jako na vodě*). Za nevětné frazémy jsou považovány i v případě, kdy jejich součást tvoří věta, která se poté neodděluje čárkou (např. *rozprchli se(,) jako když do vrabců střelí*). Ve své základní podobě přirovnání je stabilní. Hlavní funkcí ustáleného přirovnání bývá intenzifikace, zdůraznění a kladné či záporné zhodnocení znaku.¹³⁴

4.4.4 Frazémy větné

Frazémy mohou mít podobu věty ve variantě slovesné (*Ranní ptáče dál doskáče*) nebo neslovesné (*Všude dobře, doma nejlíp*), případně souvětí (*Čiň čertu dobře, peklem se ti odmění*). V kontextu se chovají jako celek.¹³⁵

Vymezení frazémů podle publikace – Čeština, řeč a jazyk:

1. rčení (úsloví) – *domácí drak (o ženě), zaječí úmysly, má za ušima, má pod čepicí, poradí se i s medvědem, pálí mu to, je to hlava otevřená, ...*
2. přirovnání – se srovnávacím výrazem jako: *je chytrý jako liška, je jako dubnové počasí (nestálý), je jako slon v porcelánu (nešikovný), točí se jako korouhvička; ustálená přirovnání se srovnávacím výrazem než: udělá to, než bys řekl švec (rychle, hned)*;

¹³³ KARLÍK, Petr, Marek NEKULA, Zdenka RUSÍNOVÁ a Miroslav GREPL. *Příruční mluvnice češtiny*. Vyd. 2., opr. [i.e. 4. vyd.]. Praha: NLN, Nakladatelství Lidové noviny, 2012, s. 71.

¹³⁴ Tamtéž, s. 72.

¹³⁵ Tamtéž, s. 72.

3. pořekadla – charakter větný, shrnují často zkušenost lidí, hodnotící a imperativní: *Kam vítr, tam plášť, Panská láska po zajících skáče (je nestálá), shrnují i zkušenost protikladnou: Spěchej pomalu, Ráno moudřejší večera x Co můžeš udělat dnes, neodkládej na zítřek;*
4. pranostiky – výroky zabývající se roční dobou, počasím a závislostí prací v zemědělství: *Svatá Lucie noci upije a dne nepřidá; Únor bílý pole sílí; Na svatého Řehoře šelma sedlák, který neoře;*
5. přísloví – ustálené obrazné výpovědi dvousložkové (první část přísloví něco konstatuje, druhá část z toho vyvozuje poučení): *Kdo jinému jámu kopá, sám do ní padá; Tak dlouho se chodí se džbánem pro vodu, až se ucho utrhne; některá jsou významově protikladná: Mluvíti stříbro-mlčeti zlato x Líná huba, holé neštěstí.¹³⁶*

Výše uvedené příklady se řadí mezi tradiční (tzv. lidové) frazémy. Ve frazeologii existují i kulturní frazémy, kde jejich původ sahá až do antiky (*Kostky jsou vrženy – Alea iacta est* – výrok Gaia Julia Caesara). Můžeme se setkat i s dvojslovnými frazémy – rčení cizojazyčná (*in memoriam* – na paměť, k uctění památky. V současnosti vznikají nové frazémy nebo se znovu obměňují staré frazémy (*dát si šlofika – dát si lehára – dát si chrupku – odpočinout si*).¹³⁷

¹³⁶ ČECHOVÁ, Marie. *Čeština - řeč a jazyk*. 3., rozš. a upr. vyd. Praha: SPN - pedagogické nakladatelství, 2011, s. 66-67.

¹³⁷ Tamtéž, s. 67.

5 Výzkumná část

5.1 Cíle a organizace práce

Cílem a smyslem výzkumu bylo ověřit využití Českého národního korpusu ve výuce českého jazyka. Jeho funkčnost a smysluplnost byla ověřována pomocí pracovního listu, kteří žáci vyplňovali v hodině českého jazyka. Korpus může být vhodnou didaktickou pomůckou v hodinách českého jazyka u rodilých mluvčích. Cvičení vycházejí z autentických psaných textů a názorně tak vystihují podobu českého jazyka v komunikaci.

Výzkumu se zúčastnilo celkem 59 žáků sekundy a tercie z víceletého gymnázia. Gymnázium se nachází v jihočeském městě, které má přes 30 tisíc obyvatel. Školu navštěvuje přes 500 žáků a řadí se spíše k větším školám v regionu. Žákům je poskytováno osmileté, šestileté či čtyřleté studium s možností bilingvního studia ve francouzštině. Vedle široké nabídky studia jazyků nabízí škola i sportovní vyžití. Při realizaci výzkumu jsem žáky seznámila s tématem mé diplomové práce, vysvětlila jsem jim zaměření pracovního listu, jaká cvičení pracovní list obsahuje. Dále jsme si společně přečetli zadání jednotlivých cvičení, aby měli žáci případně možnost se doptat. Poté žáci pracovali individuálně. V každé třídě probíhal výzkum zvlášť, vyplňování pracovního listu se uskutečnilo anonymně při jedné vyučovací hodině. Na konci hodiny byla provedena reflexe a následná diskuse nad problematikou některých jevů.

Výzkumný vzorek představuje účastníky smíšeného kolektivu s různou čtenářskou gramotností. Na základě dostupných informací se v tercii nacházel žák se specifickou poruchou učení, která byla mírného charakteru.

5.2 Metodika práce

Pro diplomovou práci jsem zvolila kvalitativní metodu výzkumu, tato metoda se zakládá na jednotlivci či skupině a jejich interpretaci.

Kvalitativní výzkum interpretuje pohledy subjektů na zkoumaný předmět, výzkumník přebírá jejich perspektivu. Metodika kvalitativního výzkumu je založena na indukci.¹³⁸

¹³⁸ ŠVARŤÍČEK, Roman a Klára ŠEĎOVÁ. *Kvalitativní výzkum v pedagogických vědách*. Vyd. 2. Praha: Portál, 2014, s. 13-14.

Pracovní list se skládal z 8 cvičení, která se zaměřovala na slovní zásobu u žáků 2. stupně. Z toho 7 cvičení bylo otevřených a 1 uzavřené, kde žáci z vybraných možností doplňovali slova do textu. Součástí byla i cvičení na znalosti z frazeologie.

První cvičení bylo zaměřeno na přechylování. Žáci měli vytvořit z rodu mužského rod ženský v 1. pádě jednotného čísla. Jednalo se o tato slova: *soudce, tchán, neurovědec, tanečník, knihovnik, manžel, kluk*. Ve druhém cvičení žáci vymýšleli synonyma k vybraným podstatným jménům, přídavným jménům a slovesům v infinitivu: *naštvaný, pyšný, milý, choroba, obloha, debata, přemýšlet, stavět, opustit*. Smyslem třetího cvičení bylo utvořit slovní spojení, která spolu souvisí. Vybraná slova byla přídavná jména v množném čísle a slovesa v infinitivu: *bezpečnostní, městský, rodinný, sportovní, hrát, čekat, věnovat a plánovat*.

Ve čtvrtém cvičení měli žáci napsat, co znamenají jednotlivé zkratky, a to: *OSN, USA, Ing., ČNB, ODS, ZŠ*. Podstatou pátého cvičení bylo doplnit slova do správného tvaru, aby byl kontext srozumitelný. Jednalo se o přejatá slova, která byla uvedena pod zadáním a žáci je psali do vyznačeného místa ve větě. Cvičení obsahovalo tato slova: *virus, studentka, sezona, influencer, byznys*.

Následující cvičení se zaměřovala na procvičení frazeologie. V šestém cvičení měli žáci vysvětlit význam vybraných slovesných i neslovesných frazémů: *Tichá voda břehy mele; Boží mlýny melou pomalu, ale jistě.; Ranní ptáče dál do skáče.; Na hrubý pytel hrubá záplata; Jiný kraj jiný mrav.* Úkolem sedmého cvičení bylo napsat druhou část přirovnání: *Chytrý jako...; Mlsný jako...; Ostrý jako...; Pomalý jako...; Sladký jako....* V posledním cvičení měli žáci doplnit číslovku do frazému. Vybrané frazémy byly následující: *Kdo šetří má za tři. Každá mince má dvě strany. Jedna vlaštovka jaro nedělá. Když se dva perou, třetí se směje. Devatero řemesel desátá bída.*

Celý výzkum se skládal z jedné vyučovací hodiny v každé třídě. V hodině byly rozdány žákům pracovní listy, při jejichž vyplňování žáci pracovali samostatně. Poté následovala reflexe a diskuse nad jednotlivými cvičeními.

Výzkumu se účastnilo celkem 59 žáků víceletého gymnázia, přesněji 28 žáků ze sekundy a 31 žáků z tercie. Návratnost pracovních listů byla 100 %. Domnívám se, že s vybranými frazémy se žáci již setkali a budou znát jejich význam. Předpokládám, že i jejich slovní zásoba jakožto studentů gymnázia bude pestrá.

5.3 Forma a metoda

Pro vyplňování pracovního listu jsem zvolila individuální formu výuky, kdy žáci pracují na zadaných úkolech samostatně. Lze uplatnit i jinou formu výuky, a to práci ve dvojici či rozdělení do skupin. Učitel také může v hodině názorně ukázat žákům výsledky korpusu na data projektoru/interaktivní tabuli a společně je mohou porovnat s jejich odpověďmi.

Zvolená pomůcka využívá metodu slovní – práce s textem. Tato metoda u žáků zvyšuje jejich učební aktivitu. Specifické cíle této metody jsou popsány v kapitole 1 této práce.

5.4 Klíčové kompetence

Práce s pracovním listem rozvíjí u žáků tyto klíčové kompetence (RVP ZV, 2017):

Kompetence k učení

- Žák vyhledává a třídí informace na základě jeho pochopení. Propojuje širší poznatky v různých vzdělávacích oblastech. Získané informace efektivně využívá v procesu učení, tvůrčích činnostech a praktickém životě.

Kompetence komunikativní

- Žák se výstižně a kultivovaně vyjadřuje v písemném projevu.
- Žák rozumí různým typům textů, přemýšlí o nich a tvořivě je využívá ke svému rozvoji a začlenění se do společenského dění.

Kompetence sociální a personální

- Žák podporuje svoji sebedůvěru a samostatný rozvoj.

Kompetence pracovní

- Žák účinně užívá materiály, dodržuje vymezená pravidla a adaptuje se na nové pracovní podmínky.

6 Pracovní list

6.1 Tvorba pracovního listu

Ve výzkumné části sestavuji z korpusu cvičení zaměřená na slovní zásobu v českém jazyce, která jsou určena žákům 2. stupně. Pro uložení dat jsem si vytvořila subkorpus DP2021. Do dotazovacího řádku zadávám buď základní dotaz na konkrétní hledaná slova, nebo využiji pokročilejší dotaz, tzv. CQL. Tento složitější typ dotazu tvořím pomocí značek, tj. poziční atributy, regulární výrazy, morfologické značky a specifická značení v korpusu. Při tvorbě cvičení pracuji s korpusovým manažerem KonText a WaG. Pro práci s KonTextem je nutné přihlášení, aby byly využity všechny nástroje, které korpus poskytuje. KonText i WaG vyhodnocuje zadaný dotaz a zobrazí výsledky, jež jsem dále třídila a vybírala vhodná slova do pracovního listu. Data z WaGu pochází z korpusu syn2015. V KonTextu analyzuji data z nejnovějšího korpusu syn2020, dále syn v8 a online_now, který mapuje obsah českého internetu (internetová žurnalistika, diskuse a sociální sítě). Každé cvičení obsahuje konkordanční seznam s výsledky dotazu. Dále uvádím data o skupině textových typů.

Všechna data jsou získána z psaných projevů, které korpus shromáždil.

6.1.1 Přechylování pomocí přípon

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [word=".*yně" & tag="N.F.*"]

Přípona: -yně, slovní druh: podstatné jméno v rodě ženském

Výchozí atribut: word

Počet výskytů: 27 985; i.p.m.: 229,71

Nastavená minimální frekvence: 100

Zadání cvičení: Utvořte z rodu mužského rod ženský (přechylování) v 1. pádě jednotného čísla.

Výběr: soudce, tchán, neurovědec, tanečník, knihovník, manžel, kluk

Pořadí	Filtr	Word	Frekvence
1	p/n	kuchyně	3514
2	p/n	ministřyně	1595
3	p/n	kolegyně	1020
4	p/n	příteřkyně	921
5	p/n	předsedkyně	854
6	p/n	jeskyně	850
7	p/n	zástupkyně	628
8	p/n	náměstkyně	488

9	p/n	soudkyně	479
10	p/n	poslankyně	419
11	p/n	poradkyně	356
12	p/n	hospodyně	316
13	p/n	umělkyně	315
14	p/n	tchyně	300
15	p/n	místopředsedkyně	300
16	p/n	ministřyně	288
17	p/n	bohyně	275
18	p/n	svatyně	268
19	p/n	průvodkyně	230
20	p/n	vévodkyně	228

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Koček, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Nastavená minimální frekvence: 6

Pořadí	Filtr	Word	Frekvence
115	p/n	vojákyňě	7
116	p/n	neurovědkyňě	7
117	p/n	příjemkyňě	7
118	p/n	jmenovkyňě	7
119	p/n	leskyňě	6
120	p/n	nadšenyňě	6
121	p/n	čekyňě	6
122	p/n	spojenyňě	6
123	p/n	svůdkyňě	6
124	p/n	chotyňě	6
125	p/n	stoupenkyňě	6
126	p/n	rádkyňě	6

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Koček, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.txttype_group	Frekvence	i.p.m.
1	p/n	Publicistika	12215	305,63
2	p/n	Beletrie	9699	233,2
3	p/n	Oborová literatura	6071	150,76

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [word=".*ice" & tag="N.F.*"]

Přípona: -ice, slovní druh: podstatné jméno v rodě ženském

Výchozí atribut: word

Počet výskytů: 212 656; i.p.m. 1745,56

Nastavená minimální frekvence:100

Pořadí	Filtr	Word	Frekvence
58	p/n	čarodějnice	718
59	p/n	expedice	707
60	p/n	justice	697
61	p/n	borovice	665
62	p/n	Teplice	663
63	p/n	tanečnice	594
64	p/n	skořice	593
65	p/n	náušnice	585
66	p/n	konvice	570
67	p/n	porodnice	562
68	p/n	sliznice	545
69	p/n	pracovnice	538
70	p/n	intuice	521
71	p/n	hořčice	517
72	p/n	knihovnice	496
73	p/n	sestřenice	493
74	p/n	čtveřice	483
75	p/n	munice	482
76	p/n	kondice	482
77	p/n	popelnice	482

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.t xtype_group	Frekvence	i.p.m.
1	p/n	Publicistika	92915	2324,82
2	p/n	Oborová literatura	72582	1802,42
3	p/n	Beletrie	47159	1133,87

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [word=".*ka" & tag="N.F.*"]

Přípona: -ka, podstatné jméno v rodě ženském

Výchozí atribut: word

Počet výskytů: 369 897; i.p.m. 3036,25

Nastavená minimální frekvence: 2 000

Pořadí	Filtr	Word	Frekvence
1	p/n	matka	16160
2	p/n	otázka	9626
3	p/n	dívka	6356
4	p/n	maminka	5927
5	p/n	láska	5858
6	p/n	válka	4783
7	p/n	babička	4219
8	p/n	manželka	4064
9	p/n	ruka	3964
10	p/n	holka	3862
11	p/n	myšlenka	3723
12	p/n	nabídka	3250
13	p/n	republika	3175
14	p/n	herečka	3067
15	p/n	banka	2856
16	p/n	politika	2821
17	p/n	délka	2674
18	p/n	ředitelka	2642
19	p/n	značka	2614
20	p/n	výška	2459

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.t xtype_group	Frekvence	i.p.m.
1	p/n	Publicistika	146535	3666,45
2	p/n	Beletrie	121007	2909,4
3	p/n	Oborová literatura	102355	2541,77

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

6.1.2 Synonyma

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [tag="A.M.*"]

Tag: přídavné jméno v rodě mužském životném

Výchozí atribut: word

Počet výskytů: 1 174 626; i.p.m. 9 641,77

Zadání cvičení: Napiš k vybraným slovům synonyma (slova podobného významu).

Výběr: naštvaný, pyšný, milý, choroba, obloha, debata, přemýšlet, stavět, opustit

Pořadí	Filtr	Word	Frekvence
34	p/n	úspěšný	3581
35	p/n	celý	3451
36	p/n	milý	3445
37	p/n	poslední	3399
38	p/n	přítomný	3388
39	p/n	francouzský	3327
40	p/n	významný	3304
41	p/n	státní	3219
42	p/n	ruský	3168
43	p/n	vlastní	3165
44	p/n	nemocný	3036
45	p/n	silný	3027
146	p/n	pyšný	1116
192	p/n	naštvaný	897

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.t xtype_group	Frekvence	i.p.m.
1	p/n	Publicistika	440421	11019,76
2	p/n	Beletrie	381108	9163,21
3	p/n	Oborová literatura	353097	8768,42

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [tag="N.F.*"]

Tag: podstatné jméno v rodě ženském

Počet výskytů: 11 289 265; i.p.m. 92 666,52

Výchozí atribut: word

Pořadí	Filtr	Word	Frekvence
411	p/n	obloha	5287
412	p/n	herečka	5276
413	p/n	kavárna	5276
414	p/n	skladba	5273
415	p/n	obrazovka	5249
416	p/n	galerie	5244
417	p/n	pistole	5241
418	p/n	výhra	5238
419	p/n	pusa	5150
420	p/n	asociace	5137
421	p/n	pohoda	5128
422	p/n	demokracie	5122
431	p/n	debata	4979
459	p/n	choroba	4620

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.t xtype_group	Frekvence	i.p.m.
1	p/n	Oborová literatura	4082827	101388,35
2	p/n	Publicistika	4079282	102067,55
3	p/n	Beletrie	3127156	75188,08

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [tag="Vf.*"]

Tag: sloveso, infinitiv

Výchozí atribut: word

Počet výskytů: 2 289 025; i.p.m. 18 789,17

Pořadí	Filtr	Word	Frekvence
104	p/n	trvat	3241
105	p/n	stavět	3154
106	p/n	znamenat	3135
107	p/n	opustit	3132
108	p/n	zabránit	3121
109	p/n	napsat	3115
110	p/n	bránit	3112
111	p/n	očekávat	3107
112	p/n	těšit	3100
113	p/n	zaplatit	3084
114	p/n	navštívit	3075
115	p/n	vyřešit	3046
116	p/n	vydat	3025
117	p/n	přemýšlet	3023

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.txttype_group	Frekvence	i.p.m.
1	p/n	Beletrie	801082	19260,89
2	p/n	Publicistika	759629	19006,65
3	p/n	Oborová literatura	728314	18086,13

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

6.1.3 Slovní spojení

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [tag="AA.P.*"]

Tag: přídavné jméno, obyčejné, v množném čísle

Výchozí atribut: lemma

Počet výskytů: 3 121 555; i.p.m. 25 622,89

Zadání cvičení: Utvoř vhodná slovní spojení, která spolu souvisí.

Pořadí	Filtr	Lemma	Frekvence
68	p/n	bezpečnostní	4951
69	p/n	pouhý	4899
70	p/n	uvedený	4793
71	p/n	obchodní	4764
72	p/n	historický	4724
73	p/n	potřebný	4691
74	p/n	úvodní	4631
81	p/n	rodinný	4425
92	p/n	městský	4030
109	p/n	sportovní	3559

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

V korpusu WaG jsem položila základní dotaz, kde jsem vyhledala výše vybraná slova z korpusového manažeru KonText, verze syn2020. WaG představuje základní přehled o slovech, data pochází z dosud největšího českého korpusu syn2015. Tabulka kolokace uvádí nejčastější slovní spojení.

Bezpečnostní

Pořadí	Filtr	Lemma	Frekvence
1	p/n	opatření	461
2	p/n	pás	301
3	p/n	složka	398
4	p/n	agentura	315
5	p/n	kamera	225
6	p/n	archiv	152
7	p/n	riziko	211
8	p/n	prvek	191
9	p/n	sklo	137
10	p/n	služba	328

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Městský

Pořadí	Filtr	Lemma	Frekvence
1	p/n	policie	1057
2	p/n	úřad	1011
3	p/n	hromadný	362
4	p/n	doprava	588
5	p/n	část	1457
6	p/n	strážník	313
7	p/n	knihovna	403
8	p/n	rada	539
9	p/n	soud	475
10	p/n	divadlo	432

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Rodinný

Pořadí	Filtr	Lemma	Frekvence
1	p/n	dům	2678
2	p/n	příslušník	567
3	p/n	doměk	511
4	p/n	rozpočet	181
5	p/n	tradice	182
6	p/n	farma	129
7	p/n	kruh	156
8	p/n	hrobka	105
9	p/n	zázemí	114
10	p/n	výstavba	136

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Sportovní

Pořadí	Filtr	Lemma	Frekvence
1	p/n	hala	476
2	p/n	areál	322
3	p/n	klub	429
4	p/n	aktivita	330
5	p/n	manažer	232
6	p/n	akce	336
7	p/n	svaz	164
8	p/n	ředitel	277
9	p/n	vyžití	92
10	p/n	kulturní	162

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [tag="VB.*"]

Tag: sloveso, tvar přítomného nebo budoucího času

Výchozí atribut: lemma

Počet výskytů: 7 780 739; i.p.m. 63 867,22

Pořadí	Filtr	Lemma	Frekvence
25	p/n	čekat	20871
26	p/n	začít	20797
27	p/n	dostat	20124
28	p/n	nabízet	19922
29	p/n	žít	19503
30	p/n	řící	19188
31	p/n	umět	18334
32	p/n	hrát	18155
111	p/n	věnovat	7553
176	p/n	plánovat	4712

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Koček, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Stejně tak jako u přídavných jmen jsem zadávala do WaGu i slovesa. Pořadí nejčastějších slovních spojení udává tzv. logDice, který frekventuje slova napříč všemi korpusy, nehledě na jejich velikost.

Hrát

Pořadí	Filtr	Lemma	Frekvence
1	p/n	role	4794
2	p/n	fotbal	956
3	p/n	hra	1114
4	p/n	divadlo	623
5	p/n	zápas	687
6	p/n	liga	497
7	p/n	dobře	1071
8	p/n	hokej	397
9	p/n	důležitý	667
10	p/n	karta	419

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Čekat

Pořadí	Filtr	Lemma	Frekvence
1	p/n	až	3723
2	p/n	dlouho	1012
3	p/n	trpělivě	595
4	p/n	my	2487
5	p/n	vy	1430
6	p/n	na	14568
7	p/n	odpověď	452
8	p/n	co	2403
9	p/n	další	1101
10	p/n	oni	1962

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Věnovat

Pořadí	Filtr	Lemma	Frekvence
1	p/n	pozornost	3095
2	p/n	problematika	233
3	p/n	kapitola	276
4	p/n	čas	793
5	p/n	péče	234
6	p/n	naplno	145
7	p/n	plně	170
8	p/n	téma	233
9	p/n	studium	209
10	p/n	podrobně	137

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Plánovat

Pořadí	Filtr	Lemma	Frekvence
1	p/n	původně	130
2	p/n	budoucnost	86
3	p/n	příští	142
4	p/n	budoucnost	102
5	p/n	výstavba	75
6	p/n	dopředu	57
7	p/n	výlet	52
8	p/n	letos	112
9	p/n	svatba	46
10	p/n	otevření	34

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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

6.1.4 Zkratky

Korpusový manažer: KonText, syn2020

Zadání dotazu: CQL, [tag="BN.*"]

Tag: zkratka, podstatné jméno obyčejné

Výchozí atribut: word

Počet výskytů: 813 668; i.p.m.: 6 678,89

Zadání cvičení: Napiš, co znamenají tyto zkratky.

Výběr: OSN, USA, Ing, ČNB, ODS, ZŠ

Pořadí	Filtr	Word	Frekvence
29	p/n	ČSSD	6431
30	p/n	obr	6145
31	p/n	kg	5922
32	p/n	f	5598
33	p/n	Ing	5125
34	p/n	ml	5100
35	p/n	In	4777
36	p/n	ODS	4252
37	p/n	str	4178
38	p/n	kw	4087

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	Word	Frekvence
8	p/n	USA	10467
31	p/n	kg	5922
36	p/n	ODS	4252
74	p/n	OSN	1981
90	p/n	ČNB	1548
95	p/n	ZŠ	1469

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

Pořadí	Filtr	doc.txtype_group	Frekvence	i.p.m.
1	p/n	Oborová literatura	456479	11335,69
2	p/n	Publicistika	312458	7818
3	p/n	Beletrie	44731	1075,49

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelinek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-31] Dostupný z WWW: <http://www.korpus.cz>

6.1.5 Práce s textem, přejatá slova

Korpusový manažer: KonText, online_now

Zadání dotazu: základní dotaz

Výchozí atribut: word

Zadání cvičení: Doplň slova do textu ve správném tvaru, aby byl kontext srozumitelný.

Byznys

Počet výskytů: 28 506; i.p.m. 29,59

Z mého pohledu profesionála působícího v internetovém reklamním **byznysu** více než 20 let jsou EuroZpravy.cz seriózním zdrojem mainstreamového zpravodajství, které garantuje provozovatel serveru mediální skupina INCORP a. s.

Zdroj: sdeleni.idnes, [2020-09-01]. Dostupné na:

https://sdeleni.idnes.cz/zpravy/eurozpravy-cz-dusledne-bojuji-proti-sireni-lzi-na-internetu.A200901_105549_zpr_sdeleni_hradr

Pořadí	Filtr	doc.media_type	Frekvence	i.p.m.
1	p/n	Twitter	8887	28,43
2	p/n	Facebook	8079	42,41
3	p/n	discussions	5740	36,57
4	p/n	Hlavní proud	2998	25,97
5	p/n	Ostatní	1000	14,63
6	p/n	Antisystémové weby	476	24,78
7	p/n	forums	427	8,18
8	p/n	Bulvární media	253	13,4
9	p/n	Politický bulvár	231	31,38
10	p/n	Názorové deníky	163	32,14

Cvrček, V. – Procházka, P.: ONLINE_NOW: monitorovací korpus internetové češtiny. Ústav Českého národního korpusu FF UK, Praha 2020 [cit. 2021-05-31]. Dostupný z WWW: <http://www.korpus.cz>

Influencer

Počet výskytů: 1 364; i.p.m. 1,42

Influencer byl velmi populární vzhledem ke svému mladému věku. Na Instagramu měl přes půl milionu sledujících a 139 tisíc odběratelů na videoportálu YouTube.com.

Zdroj: blesk.cz, [2020-09-07]. Dostupné na: <https://www.blesk.cz/clanek/zpravy-krimi/654691/tragicka-smrt-mlade-hvezdy-internetu-17-predavkoval-se-boji-se-pratele.html>

Pořadí	Filtr	doc.media_type	Frekvence	i.p.m.
1	p/n	Twitter	788	2,52
2	p/n	Facebook	167	0,88
3	p/n	Hlavní proud	130	1,13
4	p/n	Ostatní	102	1,49
5	p/n	discussions	65	0,41
6	p/n	Bulvární media	39	2,07
7	p/n	forums	31	0,59
8	p/n	Market-driven media	16	1,75
9	p/n	Antisystémové weby	13	0,68
10	p/n	Politický bulvár	12	1,65

Cvrček, V. – Procházka, P.: ONLINE NOW: monitorovací korpus internetové češtiny. Ústav Českého národního korpusu FF UK, Praha 2020 [cit. 2021-05-31]. Dostupný z WWW: <http://www.korpus.cz>

Virus

Počet výskytů: 94 450; i.p.m. 98,05

Než přišla koronavirová pandemie, prodávali Jakub Sapák a Lukáš Cafourek zhruba tisíc látkových sáčků na ovoce, zeleninu nebo pečivo měsíčně. Sami je navrhli a nechali ušít. V březnu ale Evropu ochromil **virus**, lidé v obchodech raději sáhli po tradičních plastových pytlících a majitelům českého start-upu pojmenovaného Saaczech to srazilo prodeje na desetinu loňského stavu.

Zdroj: byznys.ihned.cz, [2020-09-01]. Dostupné na: <https://archiv.ihned.cz/c1-66810180-podnikatele-z-brandysa-siji-pytliky-z-armadnich-padaku-v-dobe-nakazy-ale-vedou-ty-plastove>

Pořadí	Filtr	doc.media_type	Frekvence	i.p.m.
1	p/n	discussions	35043	223,28
2	p/n	Twitter	23173	74,13
3	p/n	Facebook	16341	85,78
4	p/n	Hlavní proud	8635	74,81
5	p/n	Antisystémové weby	4024	209,51

6	p/n	Ostatní	2382	34,84
7	p/n	forums	1311	25,12
8	p/n	Politický bulvár	1051	144,84
9	p/n	Bulvární media	941	49,84
10	p/n	Náborové deníky	806	158,95

Cvrček, V. – Procházka, P.: *ONLINE_NOW: monitorovací korpus internetové češtiny. Ústav Českého národního korpusu FF UK, Praha 2020 [cit. 2021-05-31]. Dostupný z WWW: <http://www.korpus.cz>*

Sezona

Počet výskytů: 90 808; i.p.m. 94,27

Jsem strašně hrdý na Česko, pořád jsme parádní hokejový národ. Švýcarský šampionát o rok později, kde měl Řiha znovu vést reprezentaci, byl kvůli koronaviru zrušen. Zkušenému trenérovi po **sezóně** skončila smlouva a pro nový ročník jej u reprezentace nahradil Filip Pešán.

Zdroj: praha.idnes.cz, [2020-09-01]. Dostupné na:
https://www.idnes.cz/hokej/reprezentace/milos-riha-umrti-trener-hokej.A200901_133744_reprezentace_zele

Pořadí	Filtr	doc.media_type	Frekvence	i.p.m.
1	p/n	Hlavní proud	36320	314,66
2	p/n	Ostatní	22171	324,27
3	p/n	Facebook	9493	49,83
4	p/n	Bulvární media	9260	490,49
5	p/n	Twitter	8091	25,88
6	p/n	discussions	3083	19,64
7	p/n	forums	1341	25,69
8	p/n	Market-driven media	567	61,86
9	p/n	Web instituce	198	52,43
10	p/n	Stranické weby	162	44,76

Cvrček, V. – Procházka, P.: *ONLINE_NOW: monitorovací korpus internetové češtiny. Ústav Českého národního korpusu FF UK, Praha 2020 [cit. 2021-05-31]. Dostupný z WWW: <http://www.korpus.cz>*

Studentka

Počet výskytů: 7 496; i.p.m. 7,78

Patnáctiletá **studentka** z Havířova ostravské 1st International School of Ostrava zaujala porotu mj. zpěvem a také znalostí tří světových jazyků. Na finálovém večeru zpívala v italštině svou vlastní píseň Caruso, k níž letos v březnu natočila na Bali také videoklip.

Zdroj: karvinsky.denik.cz, [2020-09-01]. Dostupné na:

https://karvinsky.denik.cz/zpravy_region/podivejte-se-divkou-roku-2020-se-stala-vali-angel-z-havirova-20200901.html

Pořadí	Filtr	doc.media_type	Frekvence	i.p.m.
1	p/n	Twitter	2080	6,65
2	p/n	Hlavní proud	1837	15,91
3	p/n	Facebook	1333	7
4	p/n	discussions	622	3,96
5	p/n	Ostatní	590	8,63
6	p/n	forums	326	6,25
7	p/n	Bulvární média	272	14,41
8	p/n	Antisystémové weby	121	6,3
9	p/n	Web instituce	77	20,39
10	p/n	Market-driven média	75	8,18

Cvrček, V. – Procházka, P.: *ONLINE_NOW: monitorovací korpus internetové češtiny. Ústav Českého národního korpusu FF UK, Praha 2020 [cit. 2021-05-31]. Dostupný z WWW: <http://www.korpus.cz>*

6.1.6 Význam frazémů

Korpusový manažer: KonText, syn v8

Zadání dotazu: CQL [tag="AA.*" & col_type="M.*"] [tag="NN.*" & col_type="M.*"]

Tag: přídavné jméno, obyčejné a podstatné jméno, obyčejné

Výchozí atribut: col_lemma

Col_type: M: přísloví, okřídlená rčení, citace apod.

Počet výskytů: 29 130; i.p.m. 5,4

Nastavená minimální frekvence: 100

Zadání cvičení: Vysvětli význam přísloví a pořekadla.

Výběr: Tichá voda břehy mele; Boží mlýny melou pomalu, ale jistě.; Ranní ptáče dál do skáče.; Na hrubý pytel hrubá záplata; Jiný kraj jiný mrav.

Vybrané frazémy jsou slovesné i neslovesné.

Pořadí	Filtr	Word	Frekvence
1	p/n	jiný mrav	1388
2	p/n	zlaté dno	1307
3	p/n	jiné kraj	1220
4	p/n	stejně řeky	1127
5	p/n	zdravý duch	969
6	p/n	stará láska	964
7	p/n	marná lásky	964
8	p/n	zdravém těle	957
9	p/n	nový rok	952
10	p/n	celý rok	847
16	p/n	boží mlýny	577
21	p/n	hrubý pytel	442
37	p/n	tichá voda	239
48	p/n	ranní ptáče	145

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.: *Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>*

6.1.7 Přirovnání

Korpusový manažer: KonText, syn v8

Zadání dotazu: základní dotaz

Výchozí atribut: lemma

Zadání cvičení: Napiš druhou část přirovnání.

Po zadání základního dotazu mi KonText vyhodnotil výsledky v konkordančním řádku, v horní liště jsem zadala kolokaci s vlastním nastavením. Atribut jsem zvolila lemma v kontextu 0 do 1, tedy první slovo, které následuje za zadaným výrazem. Minimální frekvenci kolokátu v korpusu jsem ponechala na stupni 3.

Chytrý jako

Počet výskytů: 719; i.p.m. 0,13

Pořadí	Filtr	Lemma	Frekvence
1	p/n	opice	72
2	p/n	liška	56
3	p/n	rádio	38
4	p/n	já	22
9	p/n	četník	8
13	p/n	Einstein	7

16	p/n	čert	5
17	p/n	Edison	5

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

Mlsný jako

Počet výskytů: 36; i.p.m. 0,01

Pořadí	Filtr	Lemma	Frekvence
1	p/n	koza	13
2	p/n	kočka	3
3	p/n	kocour	3

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

Ostrý jako

Počet výskytů: 2 093; i.p.m. 0,39

Pořadí	Filtr	Lemma	Frekvence
1	p/n	břitva	950
2	p/n	žiletka	124
3	p/n	nůž	119
4	p/n	jehla	57
6	p/n	meč	42
7	p/n	dýka	33
10	p/n	břit	19

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

Pomalý jako

Počet výskytů: 128; i.p.m. 0,02

Pořadí	Filtr	Lemma	Frekvence
1	p/n	želva	13
2	p/n	šnek	11
5	p/n	hlemýžď	4
10	p/n	řidič	3

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

Sladký jako

Počet výskytů: 669; i.p.m. 0,12

Pořadí	Filtr	Lemma	Frekvence
1	p/n	med	126
2	p/n	láska	42
3	p/n	cukr	20
6	p/n	cumel	12
11	p/n	bonbón	6
14	p/n	limonáda	5

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

6.1.8 Doplnění číslovky do frazému

Korpusový manažer: KonText, syn v8

Zadání dotazu: CQL, [tag="Cl.*" & col_type="M.*"]

Tag: číslovka, základní 1-4, „půl“,...; sto a tisíc v nesubstantivním skloňování

Výchozí atribut: col_lemma

Col_type: M: přísloví, okřídlená rčení, citace apod.

Počet výskytů: 9 825; i.p.m. 1,82

Zadání cvičení: Doplně číslovku do přísloví.

Výběr: Kdo šetří, má za tři. Každá mince má dvě strany. Jedna vlaštovka jaro nedělá.

Pořadí	Filtr	Lemma	Frekvence
1	p/n	jeden	5974
2	p/n	dva	3529
3	p/n	tři	318
4	p/n	sto	4

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

Korpusový manažer: KonText, syn v8

Zadání dotazu: CQL, [tag="Cr.*" & col_type="M.*"]

Tag: číslovka, řadová

Výchozí atribut: col_lemma

Col_type: M: přísloví, okřídlená rčení, citace apod.

Počet výskytů: 2 612; i.p.m. 0,48

Výběr: Když se dva perou, třetí se směje. Devatero řemesel desátá bída.

Pořadí	Filtr	Lemma	Frekvence
1	p/n	první	1484
2	p/n	třetí	605
3	p/n	druhý	477
4	p/n	desátý	40
5	p/n	čtvrtý	3
6	p/n	pátý	3

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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

6.2 Vyhodnocení pracovních listů

Díky získaným datům z pracovních listů jsem vyhodnotila výsledky, jaká je slovní zásoba u žáků ze sekundy a tercie na víceletém gymnáziu. U cvičení, kde mohli žáci napsat více variant, srovnávám s frekvencí v korpusu na základě dostupných dat.

6.2.1 První cvičení

V prvním cvičení měli žáci vytvořit z rodu mužského rod ženský, tzv. přechylování, v 1. pádě jednotného čísla. Cvičení mělo u každého slova pouze jednu správnou variantu kromě slova *kluk*, kde mohla být varianta *dívka* nebo hovorový výraz *holka*: *soudkyně, tchyně, neurovědkyně, tanečnice, knihovnice, manželka, holka/dívka*. Žáci správně vytvořili tvar ženského rodu, někteří uváděli i obě varianty *holka/dívka*. Více odpovídali variantou *holka*, která bývá užívána častěji. Ve slově *tchyně* většina žáků psala chybně dlouhý tvar, který je tvořen kořenem *tch-* a příponou *-yně*. Korpusová databáze KonTextu syn v8 uvádí i tento dlouhý tvar, který se v textech často vyskytuje. K psaní dlouhého ypsilonu dochází vzhledem k dlouhé výslovnosti.

6.2.2 Druhé cvičení

Následující cvičení se zaměřovalo na synonyma, kde žáci psali mnoho variant slov podobného významu. U přídavných jmen *naštvaný, pyšný a milý* odpovídali slovy: *rozzlobený, rozdurděný, rozzuřený, rozhořčený, rozhněvaný, nasraný, rozčilený, zlý, nespokojený, vytočený, nazlobený, uražený, nepříjemný, durdivý, vzteklý; nafrněný, napýřený, rozmazlený, hrdý, namyšlený, egoistický, arogantní, nafoukaný, nadutý, nafrněný, nadmutý, sebestředný, povrchní; hodný, příjemný, zdvořilý, laskavý, přívětivý, přátelský, vřelý*.

K podstatným jménům *choroba, obloha a debata* doplnili žáci tato slova: *nemoc, postižení, onemocnění, nákaza, virus; nebe, nebesa, příloha, mráčky; mluvení, diskuse/diskuze, rozhovor, konverzace, debatní kroužek, výměna názorů, dialog, rozmluva, komunikace*.

V případě sloves *přemýšlet, stavět a opustit* napsali žáci varianty synonym: *myslet, přemítat, uvažovat, zamyslet se, dumat, bádát, zapojit mozkové závity, polemizovat; dělat, budovat, skládat, vytvářet, tvořit, složit, konstruovat; odejít, zanechat, nechat, odloučit se, nechat být*.

V tomto cvičení se u žáků prokázala široká znalost slovní zásoby, zvláště u přídavných jmen, kde se vyskytovalo několik různých variant, čímž dané slovo žáci nahradili. U dvou žáků se vyskytla ve slově *přemítat* pravopisná chyba, kdy žáci napsali slovo s ypsilonem. Mohli si to chybně odvodit od slova *myslet*. V některých pracovních listech se objevovala prázdná místa bez odpovědi, takových bylo minimálně.

6.2.3 Třetí cvičení

Třetí cvičení se zaměřovalo na vytvoření vhodných slovních spojení ke 4 přídavným jménům a 4 slovesům v infinitivu: *bezpečnostní, městský, rodinný, sportovní, hrát, čekat, věnovat, plánovat*. Z odpovědí žáků plyne několik různých variant, jaká smysluplná slovní spojení mohou existovat. Utvořili taková slovní spojení, která zasahují do různých odvětví, např. sportu, hudby.

- Bezpečnostní: *pás, opatření, varování, zámek, vesta, služba, systém, helma, zábradlí, východ, pokyn, riziko, prověrka, kamery, alarm, prvek, složka, páska, síly.*
- Městský: *úřad, park, rozhlas, radnice, dům, znak, byt, vůz, autobus.*
- Rodinný: *kruh, dům, výlet, podnik, přítel, typ, obyčej, sněm, příslušník, svaz, rozpočet, zájezd, vztah, erb, poklad, poradce, obraz, zvyk, člen, majetek, automobil, oběd.*
- Sportovní: *oblečení, stadion, soutěž, obuv, hala, akce, členka, boty, zápas, vybavení, utkání, centrum, cvičení, náčini, klub, areál, hřiště, auto, úbor, úraz, podprsenka, manažer.*
- Hrát: *fotbal, hokej, hru/hry, florbal, šachy, karty, na flétnu, tenis, na housle, o peníze, na hudební nástroj, na kytaru, na nervy, volejbal.*
- Čekat: *na hřišti, na prohlídku, dítě, na autobus, na někoho, miminko, na výsledch, návštěvu, na vlak, na výsledky, na smrt, na sestru, na jídlo, na přítele, v čekárně, na ticho, na kamaráda, déšť, balíček, doma, změnu, venku, na oběd, na zázrak, na signál, na zastávce, návštěvu, na znamení, dlouho, pohromu.*
- Věnovat: *se sportu, dar/dárek, se učení, památku, pozornost, se práci, kytici/květinu, svůj čas, album, vyznamenání, peníze, polibek, přání, srdce, se kreslení, majetek, podíl, se škole, knihu, se hudbě, se sportu, se koničkám, pozornost, věno, ledvinu, pozdrav.*
- Plánovat: *dovolenou, výlet, dopředu, svatbu, cestu, akci, den, plán, projekt, atentát, život, trasu, odjezd, prázdniny, loupež, do budoucna, pohřeb, pomstu.*

Porovnáním dat s korpusem syn2015 jsou ve výsledcích žákovských odpovědí zastoupena nejčastěji frekventovaná slova: *opatření, pás, složka, služba, úřad, dům, příslušník, rozpočet, hala, areál, klub, fotbal, hokej, karty, dlouho* a spojení s předložkou *na, pozornost, čas, výlet, dopředu, budoucno*.

6.2.4 Čtvrté cvičení

V dalším cvičení měli žáci napsat, co znamenají názvy zkratk: *OSN, USA, Ing., ČNB, ODS a ZŠ*. Správně byly vysvětleny celým názvem: *Organizace spojených národů, United States of America/Spojené státy americké, inženýr, Česká národní banka, Občanská demokratická strana a základní škola*. U zkratky *Ing.* uváděli někteří žáci, že se jedná o titul a doslovně ho napsali. V případě *USA* byla zkratka napsána buď v angličtině, nebo v českém překladu. Žáci převážně zkratky znali, u pár jedinců bylo místo pro odpověď prázdné. Při reflexi jsem se žáků dotazovala, zda v uvedených zkratkách byla nějaká, kterou neznají. Odpověděli mi, že si přesně nepamatují doslovný překlad *OSN* a *ODS*. Věděli, že se zkratky týkají mezinárodní organizace a politické strany, ale v komunikaci se spíše užívá právě zkrácený tvar. Ve zkratkách se objevovala chyba v psaní velkých písmen, např. *Občanská Demokratická Strana*.

6.2.5 Páté cvičení

Smyslem pátého cvičení bylo doplnit přejatá slova do kontextu ve správném tvaru. Žáci měli slova k dispozici pod zadáním: *virus, studentka, sezona, influencer, byznys*. Kromě jednoho žáka, který zaměnil 2 slova ve větách, všichni správně doplnili a vyskoňovali slova do vět tak, aby byl zachován smysl sdělení. Z reflexe mohu shledat, že všem daným slovům rozuměli a setkali se s nimi v komunikaci, publicistice nebo na internetu.

Správné řešení: *influencer, virus, studentka* – 1. pád, jednotného čísla; *byznysu/byznyse* – 6. pád, jednotného čísla; *sezoně/sezóně* – 6. pád, jednotného čísla. Jednoznačně u slova *byznys* se vyskytoval tvar *byznysu*. V korpusu *online_now* je výskyt tvaru *byznysu* několikanásobně vyšší oproti tvaru *byznyse*.

Z mého pohledu profesionála působícího v internetovém reklamním byznysu více než 20 let jsou EuroZpravy.cz seriózním zdrojem mainstreamového zpravodajství, které garantuje provozovatel serveru mediální skupina INCORP a. s.

Influencer byl velmi populární vzhledem ke svému mladému věku. Na Instagramu měl přes půl milionu sledujících a 139 tisíc odběratelů na videoportálu YouTube.com.

Než přišla koronavirová pandemie, prodávali Jakub Sapák a Lukáš Cafourek zhruba tisíc látkových sáčků na ovoce, zeleninu nebo pečivo měsíčně. Sami je navrhli a nechali ušít. V březnu ale Evropu ochromil virus, lidé v obchodech raději sáhli po

tradičních plastových pytlících a majitelům českého start-upu pojmenovaného Saaczech to srazilo prodeje na desetinu loňského stavu.

*Jsem strašně hrdý na Česko, pořád jsme parádní hokejový národ. Švýcarský šampionát o rok později, kde měl Říha znovu vést reprezentaci, byl kvůli koronaviru zrušen. Zkušenému trenérovi po **sezóně** skončila smlouva a pro nový ročník jej u reprezentace nahradil Filip Pešán.*

*Patnáctiletá **studentka** z Havířova ostravské 1st International School of Ostrava zaujala porotu mj. zpěvem a také znalostí tři světových jazyků. Na finálovém večeru zpívala v italštině svou vlastní píseň Caruso, k níž letos v březnu natočila na Bali také videoklip.*

6.2.6 Šesté cvičení

Šesté cvičení se věnovalo vysvětlení následujících frazémů: *Tichá voda břehy mele.*; *Boží mlýny melou pomalu, ale jistě.*; *Ranní ptáče dál do skáče.*; *Na hrubý pytel hrubá záplata.*; *Jiný kraj jiný mrav.* Toto cvičení bylo pro žáky nejnáročnější, aby přesně a srozumitelně formulovali význam daného frazému. O problematice frazémů jsem s žáky diskutovala v závěrečné reflexi, kde mi sami potvrdili, že bylo obtížnější najít přesná slova k vyjádření frazémů, když se běžně tyto frazémy používají a všichni jim v komunikaci rozumí. Ve cvičení nebyly některé frazémy vysvětleny.

Tichá voda břehy mele žáci vystihli formulací, že nevýrazní lidé mohou něčím překvapit, něco v životě dokázat. Chybné nebo nepřesné interpretace byly např.: *Nevšimne si něčeho, co se stalo.* „*Neviditelné*“ věci toho spoustu dělají. *Nejhlasitější není nejlepší.*

Boží mlýny melou pomalu, ale jistě vyjádřili respondenti větou: *Na každého jednou dojde. Spravedlnost každého dostihne. Pokud někdo udělá něco špatného, trest ho dřív nebo později nemine.*

S tímto frazémem se pojilo spíše více špatných odpovědí, např.: *Když děláme něco pomalu, tak se nám to povede. Měli bychom jít na jistotu. Není třeba rychlosti, ale kvality. Není třeba spěchat. Nedělat věci rychle a zbrkle.*

Ranní ptáče dál do skáče bylo charakterizováno, že když člověk brzy vstane, víc toho přes den stihne udělat. Odpovědi se týkaly množství práce, kterou je člověk schopen za den vykonat.

Na hrubý pytel hrubá záplata formulovali respondenti: *Ke každému se chovat tak, jak si zaslouží. Jak on tobě, tak ty jemu. Když se někdo chová zle, budou se tak k němu chovat i ostatní.* V některých odpovědích byla i chyba např.: *Když praskne pytel musí se zalepit.*

Jiný kraj jiný mrav, tento frazém žáci vysvětlovali, že jsou v různých zemích nebo krajích různé zvyklosti a tradice. Pochopili, že se jedná o odlišnosti v jednotlivých zemích. Vyskytovala se zde u 2 žáků pravopisná chyba ve slově: *mravy*, *zvyky*.

6.2.7 Sedmé cvičení

V následujícím cvičení žáci vymýšleli druhou část přirovnání. Existuje několik typů přirovnání, proto i v tomto cvičení bylo několik různých odpovědí. Také se vyskytovalo pár nevyplněných přirovnání. Nejčastěji žáci doplnili zvíře, věc, jídlo nebo osobu. Kolokační seznam korpusu syn v8 vyhodnotil nejčastější výskyt vybraných přirovnání, kde se nachází *chytrý jako opice* s frekvencí 72, u žáků toto spojení nebylo zaznamenáno. V žakovských pracích bylo nejvíce zastoupeno *chytrý jako liška*, které je v korpusu umístěno na 2. místě. U přirovnání *mlsný jako* a *pomalý jako* se výsledky žáků a korpusu shodovaly. V dalších přirovnáních *ostrý jako* a *sladký jako* se vyskytovalo několik variant, které vyhledal i korpus ve své databázi s různou frekvencí.

- Chytrý jako: *liška, Elon Musk, učebnice, rádio, sova, Einstein*
- Mlsný jako: *kozel, koza, medvěd, kočka, vlk*
- Ostrý jako: *břitva, nůž, šíp, pila, meč, rys, chilli, zub, ježek, žiletka, dyka, žihadlo*
- Pomalý jako: *želva, hlemýžď, šnek, lenochod*
- Sladký jako: *med, pasta, cumel, jahoda, cukr, šťáva, cukrkandl, bonbón, cukrová vata, čokoláda*

K chybným možnostem patří např. *chytrý jako sova*, *chytrý jako učebnice*, *ostrý jako zub*.

6.2.8 Osmé cvičení

Úkolem posledního cvičení bylo doplnit číslovku do přísloví. Žáci shodně odpověděli na první čtyři přísloví, kde vyplnili číslovky: *tři*, *dvě*, *jedna/první*, *třetí*. V posledním přísloví byly číslovky různé nebo bylo místo nevyplněné. Respondenti uváděli číslovku *jedna* nebo zájmeno *žádná*. Při reflexi mi žáci sdělili, že toto přísloví znají méně, někteří se s ním setkali poprvé. Frekvence tohoto přísloví v korpusu syn v8 je nízká, není tedy příliš frazém užíván.

Správné řešení: *Kdo šetří, má za tři. Každá mince má dvě strany. Jedna/první vlaštovka jaro nedělá. Když se dva perou, třetí se směje. Devatero řemesel desátá bída.*

7 Závěr

Diplomová práce se zabývala využitím Českého národního korpusu ve výuce českého jazyka u žáků ze sekundy a tercie víceletého gymnázia. Práce se dělí na dvě části – teoretickou a výzkumnou.

Teoretická část vymezuje na základě odborné literatury a zdrojů didaktiku českého jazyka; korpusovou lingvistiku, její historii, metodologii a typy korpusů. Představuje akademický projekt, Český národní korpus, jako databázi autentických textů, který má několik korpusových aplikací. Dále podává přehled o lexikologii jako vědní lingvistické disciplíně a obohacování slovní zásoby.

Výzkumná část popisuje tvorbu pracovního listu zaměřeného na slovní zásobu a hodnotí výsledky těchto pracovních listů, které žáci vyplňovali při hodině českého jazyka. Bylo zkoumáno, zda může být korpusová databáze vhodnou didaktickou pomůckou pro učitele, kteří by ji uplatnili ve výuce. Dále práce zjišťuje úroveň slovní zásoby u žáků nižšího stupně gymnázia a porovnává výsledky s korpusem. Při tvorbě pracovního listu jsem pracovala s korpusovým manažerem KonText a WaG, kde jsem zadávala základní nebo pokročilý dotaz. V KonTextu jsem získala data z nejnovějších korpusů syn2020 a online_now, dále z verze syn v8 a syn2015. Po vyhodnocení jednotlivých dotazů jsem výsledky zaznamenala do tabulek, následně třídila a vybírala vhodná slova do pracovního listu.

Pracovní list se skládá z 8 cvičení na slovní zásobu, poslední 3 cvičení se specializují na frazeologii. Všechna cvičení rozvíjí klíčové kompetence dle RVP ZV. Výzkum probíhal na víceletém gymnáziu při jedné vyučovací hodině. V prvním úkolu žáci tvořili z rodu mužského rod ženský, tzv. přechylování. Druhé cvičení se zaměřovalo na synonyma k vybraným slovům. Třetí cvičení podporovalo kreativitu při vymýšlení vhodných slovních spojení. Ve čtvrtém cvičení žáci vysvětlovali, co znamenají jednotlivé zkratky. Následující cvičení vychází z práce s textem, kde žáci doplňovali slova do vět, aby byl zachován kontext. Šesté cvičení ověřuje znalost z frazeologie, žáci měli krátce vysvětlit význam 5 frazémů. V sedmém cvičení doplňovali druhou část přirovnání. Poslední cvičení bylo na doplnění číslovky do frazému.

Na základě výsledků zúčastněných tříd se ukázalo, že žáci disponují velmi dobrou slovní zásobou, i v porovnání s korpusem. Díky závěrečné reflexi jsem zjišťovala, která cvičení byla pro žáky obtížnější a zda všem slovům ve cvičeních rozuměli. Nejvíce se zamýšleli nad šestým cvičením, kde měli formulovat význam frazémů. Z dostupných dat

korpusu byly některé frazémy méně frekventované a žáci je příliš neznali. V ostatních cvičeních všem slovům rozuměli a setkali se s nimi např. v publicistice, beletrii. Některá cvičení nebyla zcela vyplněna, pravopisných chyb bylo minimum.

Výsledky výzkumné části považuji za velmi zdařilé. Potvrdily se mé domněnky, že jejich slovní zásoba je na kvalitní úrovni. U žáků se projevila i vysoká znalost frazémů, kterou získali z hodin českého jazyka a běžné komunikace. Využití Českého národního korpusu ve výuce českého jazyka shledávám jako přínosné. Při tvorbě pracovního listu se mi s ním dobře pracovalo. Uplatnila jsem zde znalosti nejen z bakalářské práce, kde jsem již v praktické části s korpusem pracovala, ale i z korpusových workshopů, které pořádal Ústav Českého národního korpusu. Pro žáky může korpus představovat zpestření do výuky a obohacují si s ním svou slovní zásobu.

8 Seznam literatury a zdrojů

Knížní publikace

- CARTER, Ronald, O' KEEFFEE, Anne, McCARTHY, Michael. *From corpus to classroom: Language use and language teaching*. UK: Cambridge university press, 2007. ISBN-13 978-0-521-85146-6
- ČAPEK, Robert. *Moderní didaktika: lexikon výukových a hodnoticích metod*. Praha: Grada, 2015. Pedagogika (Grada). ISBN 978-80-247-3450-7.
- ČECHOVÁ, Marie a Vlastimil STYBLÍK. *Čeština a její vyučování: didaktika českého jazyka pro učitele základních a středních škol a studenty učitelství*. Praha: SPN - pedagogické nakladatelství, 1998. ISBN 80-85937-47-6.
- ČECHOVÁ, Marie. *Čeština - řeč a jazyk*. 3., rozš. a upr. vyd. Praha: SPN - pedagogické nakladatelství, 2011. ISBN 978-80-7235-413-9.
- ČEJKA, Mírek. *Česká lexikologie a lexikografie*. Brno: Masarykova univerzita, 1992. ISBN 80-210-0393-6.
- ČERMÁK, František a Renata BLATNÁ, ed. *Korpusová lingvistika - stav a modelové přístupy*. Praha: NLN, Nakladatelství Lidové noviny, 2006. Studie z korpusové lingvistiky. ISBN 80-7106-861-6.
- ČERMÁK, František a Renata BLATNÁ. *Jak využívat český národní korpus*. Praha: Lidové noviny, 2005. ISBN 80-7106-736-9.
- ČERMÁK, František. *Frazeologie a idiomatika česká a obecná: Czech and general phraseology*. V Praze: Karolinum, 2007. ISBN 978-80-246-1371-0.
- ČERMÁK, František. *Idiomatika a frazeologie češtiny*. Praha: Univerzita Karlova, 1982.
- ČERMÁK, František. *Korpus a korpusová lingvistika*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. ISBN 978-80-246-3710-5.
- ČERMÁK, František. *Lexikon a sémantika*. Praha: Lidové noviny, 2010. ISBN 978-80-7422-020-3.
- FILIPEC, Josef a František ČERMÁK. *Česká lexikologie*. Praha: Academia, 1985. Studie a práce lingvistické.
- HAUSER, Přemysl. *Didaktika českého jazyka pro 2. stupeň ZŠ*. Brno: Masarykova univerzita, 2007. ISBN 978-80-210-4244-5.
- HAUSER, Přemysl. *Nauka o slovní zásobě*. 2. vyd. Praha: Státní pedagogické nakladatelství, 1986. Učebnice pro vysoké školy (Státní pedagogické nakladatelství).
- HRABAL, Vladimír, František MAN a Isabella PAVELKOVÁ. *Psychologické otázky motivace ve škole*. 2., upr. vyd. Praha: Státní pedagogické nakladatelství, 1989. Knihnice psychologické literatury. ISBN 80-04-23487-9.
- JANIŠ, Kamil. *Obecná didaktika - vybraná témata*. Vyd. 4. Hradec Králové: Gaudeamus, 2010. ISBN 978-80-7435-047-4.

KARLÍK, Petr, Marek NEKULA, Zdenka RUSÍNOVÁ a Miroslav GREPL. *Příruční mluvnice češtiny*. Vyd. 2., opr. [i.e. 4. vyd.]. Praha: NLN, Nakladatelství Lidové noviny, 2012. ISBN 978-80-7106-624-8.

KOSTEČKA, Jiří. *Do světa češtiny jinak: (didaktika českého jazyka a slohu pro vyučovací praxi)*. Jinočany: H & H, 1993. ISBN 80-85787-24-5.

KUČERA, Karel, ed. *Český národní korpus: úvod a příručka uživatele*. Praha: Filozofická fakulta UK, 2000. ISBN 80-85899-94-9.

NELEŠOVSKÁ, Alena. *Pedagogická komunikace v teorii a praxi*. Praha: Grada, 2005. Pedagogika (Grada). ISBN 80-247-0738-1.

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

PETTY, Geoffrey. *Moderní vyučování*. 6., rozš. a přeprac. vyd. Přeložil Jiří FOLTÝN. Praha: Portál, 2013. ISBN 978-80-262-0367-4.

PRŮCHA, Jan. *Přehled pedagogiky: úvod do studia oboru*. 4., aktualizované vydání. Praha: Portál, 2015. ISBN 978-80-262-0872-3.

SIEGLOVÁ, Dagmar. *Konec školní nudy: didaktické metody pro 21. století*. Praha: Grada, 2019. ISBN 978-80-271-2254-7.

SKALKOVÁ, Jarmila. *Obecná didaktika: vyučovací proces, učivo a jeho výběr, metody, organizační formy vyučování*. Praha: Grada, 2007. Pedagogika (Grada). ISBN 978-80-247-1821-7.

ŠEBESTA, Karel, ŠORMOVÁ, Kateřina a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická univerzita UK, 2019. ISBN 987-80-7308-897-2.

ŠEBESTA, Karel. *Od jazyka ke komunikaci: didaktika českého jazyka a komunikační výchova*. Vyd. 2. Praha: Karolinum, 2005. Acta Universitatis Carolinae. ISBN 80-246-0948-7.

ŠTĚPÁNÍK, Stanislav, Eva HÁJKOVÁ, Klára ELIÁŠKOVÁ, Ludmila LIPTÁKOVÁ a Marta SZYMAŃSKA. *Školní výpravy do krajiny češtiny: didaktika českého jazyka pro základní školy*. Plzeň: Fraus, 2020. ISBN 978-80-7489-595-1.

ŠULC, Michal. *Korpusová lingvistika: první vstup*. Praha: Karolinum, 1999. ISBN 80-7184-847-6.

ŠVAŘÍČEK, Roman a Klára ŠEĐOVÁ. *Kvalitativní výzkum v pedagogických vědách*. Vyd. 2. Praha: Portál, 2014. ISBN 978-80-262-0644-6.

TŘEŠTÍK, Michael, ed. *Kdo je kdo: osobnosti české současnosti: 5000 životopisů*. Praha: Agentura Kdo je kdo, 2002. ISBN 80-902586-7-0.

ZORMANOVÁ, Lucie. *Výukové metody v pedagogice: tradiční a inovativní metody, transmisivní a konstruktivistické pojetí výuky, klasifikace výukových metod*. Praha: Grada, 2012. Pedagogika (Grada). ISBN 978-80-247-4100-0.

Časopisy

CVRČEK, Václav, KOVÁŘÍKOVÁ, Dominika (eds.). Možnosti a meze korpusové lingvistiky. *Naše řeč – Základní informace* [online]. 2011. roč. 94, č. 3. [cit. 2021-02-14] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8191>

ČECHOVÁ, Marie (ed.). Dynamika frazeologie. *Naše řeč – Základní informace* [online]. 1986, roč. 69, č. 4 [cit. 2021-02-21] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=6639>

ČERMÁK, František (ed.). Jazykový korpus: Prostředek a zdroj poznání. *Slovo a slovesnost* [online]. 1995, roč. 56, č. 2 [cit. 2021-02-14] Dostupné z: <http://sas.ujc.cas.cz/archiv.php?art=3627>

FILIPEC, Josef (ed.). K úkolům české lexikologie. *Slovo a slovesnost* [online]. 1968. roč. 29, č. 3. [cit. 2021-03-05] Dostupné z: <http://sas.ujc.cas.cz/archiv.php?art=1642>

CHLUMSKÁ, Lucie, RICHTEROVÁ, Olga (eds.). Překladová čeština v korpusech. *Naše řeč – Základní informace* [online]. 2014. roč. 97, č. 4-5. [cit. 2021-02-15] Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8343>

CHROMÝ, Jan (ed.). Korpus a reprezentativnost. *Naše řeč – Základní informace* [online]. 2014. roč. 97, č. 4-5 [cit. 2021-01-24]. Dostupné z: <http://nase-rec.ujc.cas.cz/archiv.php?art=8337>

KLÖTZEROVÁ, Petra (ed.). Hranice frazeologie se posouvají. Lexikální frazémy v češtině. *Slovo a slovesnost* [online]. 1998. roč. 59, č. 4. [cit. 2021-02-21] Dostupné z: <http://sas.ujc.cas.cz/archiv.php?art=3805>

Internetové zdroje

Cvrček, V. – Procházka, P.: ONLINE_NOW: monitorovací korpus internetové češtiny. Ústav Českého národního korpusu FF UK, Praha 2020 [cit. 2021-05-31]. Dostupný z WWW: <http://www.korpus.cz>

Cvrček, Václav - Richterová, Olga (eds). manualy:keywords [online]. Příručka ČNK; 2017. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:keywords&rev=1506346748>.

Cvrček, Václav - Richterová, Olga (eds). manualy:morfio [online]. Příručka ČNK; 2017. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:morfio&rev=1506346481>.

Cvrček, Václav - Richterová, Olga (eds). manualy:proskoly [online]. Příručka ČNK; 2017. [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:proskoly&rev=1487855476>.

Cvrček, Václav - Richterová, Olga (eds). manualy:syd [online]. Příručka ČNK; 2017. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:syd&rev=1506346256>.

- Cvrček, Václav - Richterová, Olga (eds). manualy:treq [online]. Příručka ČNK; 2020. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:treq&rev=1597072085>.
- Cvrček, Václav - Richterová, Olga (eds). manualy:wag [online]. Příručka ČNK; 2020. [cit. 2021-01-25]. Dostupné na: <http://wiki.korpus.cz/doku.php?id>manualy:wag&rev=1600851035>.
- Cvrček, Václav - Richterová, Olga (eds). pojmy:korpus [online]. Příručka ČNK; 2017. [cit. 2021-02-11]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:korpus&rev=1496929366>.
- Cvrček, Václav - Richterová, Olga (eds). pojmy:korpusovy_manazer [online]. Příručka ČNK; 2016. [cit. 2021-01-24]. Dostupné na: http://wiki.korpus.cz/doku.php?id=pojmy:korpusovy_manazer&rev=1470210588.
- Cvrček, Václav - Richterová, Olga (eds). pojmy:paralelni [online]. Příručka ČNK; [cit. 2021-02-18]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:paralelni&rev=1566383701>.
- Cvrček, Václav - Richterová, Olga (eds). pojmy:reprezentativnost [online]. Příručka ČNK; 2019. [cit. 2021-01-24]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=pojmy:reprezentativnost&rev=1571233796>.
- Cvrček, Václav - Richterová, Olga (eds). seznamy:kontext_verze [online]. Příručka ČNK; 2021. [cit. 2021-01-25]. Dostupné na: http://wiki.korpus.cz/doku.php?id=seznamy:kontext_verze&rev=1610643179.
- Cvrček, Václav - Richterová, Olga (eds). start [online]. Příručka ČNK; [cit. 2021-01-23]. Dostupné na: <http://wiki.korpus.cz/doku.php?id=start&rev=1602059248>.
- ČERMÁKOVÁ, Anna, CHLUMSKÁ, Lucie. What are the suitable corpus data to support L1 teaching (not only) in Czech? [online]. In: Teaching and language corpora conference 20. - 23. July, Univerzita Giessen, Německo, 2016, s. 20 [cit. 2021-01-24]. Dostupné z: <https://www.unigiessen.de/faculties/f05/engl/ling/talc/home/programme/abstracts>
- 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.: SYN2015: reprezentativní korpus psané češtiny[online]. Ústav Českého národního korpusu FF UK, Praha 2015. [cit. 2021-05-31] 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.: Korpus SYN, verze 8 z 12. 12. 2019 [online]. Ústav Českého národního korpusu FF UK, Praha 2019. [cit. 2021-06-01] Dostupný z WWW: <https://www.korpus.cz>

Křen, M. – Cvrček, V. – Henyš, J. – Hnátková, M. – Jelínek, T. – Kocek, J. – Kovářiková, D. – Křivan, J. – Milička, J. – Petkevič, V. – Procházka, P. – Skoumalová, H. – Šindlerová, J. – Škrabal, M.: SYN2020: reprezentativní korpus psané češtiny [online]. Ústav Českého národního korpusu FF UK, Praha 2020. [cit. 2021-05-30] Dostupný z WWW: <http://www.korpus.cz>

Machálek, Tomáš (2019): Slovo v kostce – agregátor slovních profilů. FF UK, Praha. [cit. 2021-05-30] Dostupný z WWW: <http://korpus.cz/slovo-v-kostce>

VEČERKA, Radoslav. Biografickobibliografické medailonky českých lingvistů: bohemistů a slavistů (2. vydání) [online]. *Linguistica online*. Brno: electronic journal of the Department of Linguistics and Baltic Languages at Masaryk University, Czech Republic, 2008, s. 22 [cit. 2021-01-23]. ISSN 1801-5336. Dostupné z: <http://www.phil.muni.cz/linguistica/art/vecerka/vec-medailonky2.pdf>

9 Přílohy

9.1 Příloha č. 1

PRACOVNÍ LIST

Pracovní list se zaměřuje na slovní zásobu českého jazyka u žáků 2. stupně. Obsahuje data z Českého národního korpusu (www.korpus.cz). Výsledky jsou shromážděny za účelem diplomové práce s názvem: Využití Českého národního korpusu ve výuce českého jazyka.

Autorka: Bc. Žaneta Zemanová

Pedagogická fakulta, Jihočeská univerzita v Českých Budějovicích

Zakroužkuj: sekunda tercie

1. Vytvoř z rodu mužského rod ženský (přechylování) v 1. pádě jednotného čísla.

soudce	knihovník
tchán	manžel
neurovědec	kluk
tanečník	

2. Napiš k vybraným slovům synonyma (slova podobného významu).

naštvaný	debata
pyšný	přemýšlet
milý	stavět
choroba	opustit
obloha	

3. Utvoř vhodná slovní spojení, která spolu souvisí.

bezpečnostní	
městský	
rodinný	
sportovní	
hrát	
čekat	
věnovat	
plánovat	

4. Napiš, co znamenají tyto zkratky.

OSN	ČNB
USA	ODS
Ing.	ZŠ

5. Doplně slova do textu ve správném tvaru, aby byl kontext srozumitelný.

Virus, studentka, sezona, influencer, byznys

Z mého pohledu profesionála působícího v internetovém reklamním _____více než 20 let jsou EuroZpravy.cz seriózním zdrojem mainstreamového zpravodajství, které garantuje provozovatel serveru mediální skupina INCORP a. s.

_____byl velmi populární vzhledem ke svému mladému věku. Na Instagramu měl přes půl milionu sledujících a 139 tisíc odběratelů na video portálu YouTube.com.

Než přišla koronavirová pandemie, prodávali Jakub Sapák a Lukáš Cafourek zhruba tisíc látkových sáčků na ovoce, zeleninu nebo pečivo měsíčně. Sami je navrhli a nechali ušít. V březnu ale Evropu ochromil_____, lidé v obchodech raději sáhli po tradičních plastových pytlících a majitelům českého start-upu pojmenovaného Saaczech to srazilo prodeje na desetinu loňského stavu.

Jsem strašně hrdý na Česko, pořád jsme parádní hokejový národ. Na švýcarském šampionátu o rok později, kde měl Říha znovu vést reprezentaci, byl kvůli koronaviru zrušen. Zkušenému trenérovi po _____ skončila smlouva a pro nový ročník jej u reprezentace nahradil Filip Pešán.

Patnáctiletá _____ z Havířova ostravské 1st International School of Ostrava zaujala porotu mj. zpěvem a také znalostí tří světových jazyků. Na finálovém

večeru zpívala v italštině svou vlastní píseň Caruso, k níž letos v březnu natočila na Bali také videoklip.

6. Vysvětli význam přísloví a pořekadla.

Tichá voda břehy mele

Boží mlýny melou pomalu, ale jistě.

Ranní ptáče dál do skáče.

Na hrubý pytel hrubá záplata

Jiný kraj jiný mrav

7. Napiš druhou část přirovnání.

Chytrý jako _____

Mlsný jako _____

Ostrý jako _____

Pomalý jako _____

Sladký jako _____

8. Dopln číslou do přísloví.

Kdo šetří má za _____.

Každá mince má _____ strany.

_____ vlaštovka jaro nedělá.

Když se dva perou, _____ se směje.

Devatero řemesel _____ bída.

9.2 Příloha č. 2

Ukázka vyplněného pracovního listu žáka/žákyně ze sekundy

PRACOVNÍ LIST

Pracovní list se zaměřuje na slovní zásobu českého jazyka u žáků 2. stupně. Obsahuje data z Českého národního korpusu (www.korpus.cz). Výsledky jsou shromážděny za účelem diplomové práce s názvem: Využití Českého národního korpusu ve výuce českého jazyka.

Autorka: Bc. Žaneta Zemanová

Pedagogická fakulta, Jihočeská univerzita v Českých Budějovicích

Zakroužkuj: sekunda tercie

1. Vytvoř z rodu mužského rod ženský (přechylování) v 1. pádě jednotného čísla.

soudce	soudkyně	knihovnik	knihovnice
tchán	tchyně	manžel	manželka
neurovědec	neurovědkyně	kluk	chůvka
tanečník	tanečnice		

2. Napiš k vybraným slovům synonyma (slova podobného významu).

naštvaný	rozsvěšený	debata	diskuse
pyšný	namyšlený	přemýšlet	přemýšlet
milý	hulavý	stavět	budovat
choroba	nemoc	opustit	opustit
obloha	nebe		

3. Utvoř vhodná slovní spojení, která spolu souvisí.

bezpečnostní	opatření
městský	rodnice
rodinný	dlům
sportovní	cvičení
hrát	tenis
čekat	na smrt

věnovat	dar
plánovat	dovoleno

4. Napiš, co znamenají tyto zkratky.

OSN <i>Organizace spojených národů</i>	ČNB <i>Česká národní banka</i>
USA <i>(Spojené Státy Americké)</i> <i>United States of America</i>	ODS <i>Občanská demokratická strana</i>
Ing. <i>Inženýr</i>	ZŠ <i>Základní škola</i>

5. Doplň slova do textu ve správném tvaru, aby byl kontext srozumitelný.

Virus, studentka, sezona, influencer, byznys

Z mého pohledu profesionála působícího v internetovém reklamním byznysu více než 20 let jsou EuroZprávy.cz seriózním zdrojem mainstreamového zpravodajství, které garantuje provozovatel serveru mediální skupina INCORP a. s.

Influencer byl velmi populární vzhledem ke svému mladému věku. Na Instagramu měl přes půl milionu sledujících a 139 tisíc odběratelů na video portálu YouTube.com.

Než přišla koronavirová pandemie, prodávali Jakub Sapák a Lukáš Cafourek zhruba tisíc látkových sáčků na ovoce, zeleninu nebo pečivo měsíčně. Sami je navrhli a nechali ušít. V březnu ale Evropu ochromil virus, lidé v obchodech raději sáhli po tradičních plastových pytlících a majitelům českého start-upu pojmenovaného Saaczech to srazilo prodeje na desetinu loňského stavu.

Jsem strašně hrdý na Česko, pořád jsme parádní hokejový národ. Na švýcarském šampionátu o rok později, kde měl Říha znovu vést reprezentaci, byl kvůli koronaviru zrušen. Zkušenému trenérovi po sezone skončila smlouva a pro nový ročník jej u reprezentace nahradil Filip Pešán.

Patnáctiletá studentka z Havířova ostravské 1st International School of Ostrava zaujala porotu mj. zpěvem a také znalostí tří světových jazyků. Na finálovém večeru zpívala v italštině svou vlastní píseň Caruso, k níž letos v březnu natočila na Bali také videoklip.

6. Vysvětli význam přísloví a pořekadla.

Tichá voda břehy mele

Tvůri a bylosti mohou být silnější než se zdají

Boží mlýny melou pomalu, ale jistě.

Každému se časem oplácí co udělal

Ranní ptáče dál do skáče.

Kdo dříve vstává více stihne

Na hrubý pytel hrubá záplata

Jak^{ne} sobě tak by jemu

Jiný kraj jiný mrav

V každém místě jsou jiné tradice, zvyky, kulura a oběci

7. Napiš druhou část přirovnání.

Chytrý jako *liška*

Mlsný jako *hlava*

Ostrý jako *bílva*

Pomalý jako *šnek*

Sladký jako *umel*

8. Dopln číslou do přísloví.

Kdo šetří má za *šití*.

Každá mince má *dvě* strany.

Jedna vlaštovka jaro nedělá.

Když se dva perou, *tráčí* se směje.

Devatero řemesel *desátí* bída.

9.3 Příloha č. 3

Ukázka vyplněného pracovního listu žáka/žákyně z tercie

PRACOVNÍ LIST

Pracovní list se zaměřuje na slovní zásobu českého jazyka u žáků 2. stupně. Obsahuje data z Českého národního korpusu (www.korpus.cz). Výsledky jsou shromážděny za účelem diplomové práce s názvem: Využití Českého národního korpusu ve výuce českého jazyka.

Autorka: Bc. Žaneta Zemanová

Pedagogická fakulta, Jihočeská univerzita v Českých Budějovicích

Zakroužkuj: sekunda tercie

1. Vytvoř z rodu mužského rod ženský (přechylování) v 1. pádě jednotného čísla.

soudce	soudkyně	knihovnik	knihovnice
tchán	tchyně	manžel	manželka
neurovědec	neurovědkyně	kluk	holka
tanečník	tanečnice		

2. Napiš k vybraným slovům synonyma (slova podobného významu).

naštvaný	rozlobený	debata	konverzace
pyšný	hrdý	přemýšlet	uvážovat
milý	hodný	stavět	budovat
choroba	nákaza	opustit	odejít
obloha	nebe		

3. Utvoř vhodná slovní spojení, která spolu souvisí.

bezpečnostní	pas
městský	úřad
rodinný	clán
sportovní	vytavení
hrát	hru
čekat	počkání

věnovat	knihku
plánovat	výhled

4. Napiš, co znamenají tyto zkratky.

OSN	Organizace spojených národů	ČNB	Česká národní banka
USA	Spojené státy americké	ODS	Ověřená demokratická strana
Ing.	inženýr	ZŠ	Základní škola

5. Doplně slova do textu ve správném tvaru, aby byl kontext srozumitelný.

Virus, studentka, sezona, influencer, byznys

Z mého pohledu profesionála působícího v internetovém reklamním byznysu více než 20 let jsou EuroZpravy.cz seriózním zdrojem mainstreamového zpravodajství, které garantuje provozovatel serveru mediální skupina INCORP a. s.

Influencer byl velmi populární vzhledem ke svému mladému věku. Na Instagramu měl přes půl milionu sledujících a 139 tisíc odběratelů na video portálu YouTube.com.

Než přišla koronavirová pandemie, prodávali Jakub Sapák a Lukáš Cafourek zhruba tisíc látkových sáčků na ovoce, zeleninu nebo pečivo měsíčně. Sami je navrhli a nechali ušít. V březnu ale Evropu ochromil COVID, lidé v obchodech raději sáhli po tradičních plastových pytlících a majitelům českého start-upu pojmenovaného Saaczeh to srazilo prodeje na desetinu loňského stavu.

Jsem strašně hrdý na Česko, pořád jsme parádní hokejový národ. Na švýcarském šampionátu o rok později, kde měl Říha znovu vést reprezentaci, byl kvůli koronaviru zrušen. Zkušenému trenérovi po sezóně skončila smlouva a pro nový ročník jej u reprezentace nahradil Filip Pešán.

Patnáctiletá studentka z Havířova ostravské 1st International School of Ostrava zaujala porotu mj. zpěvem a také znalostí tří světových jazyků. Na finálovém večeru zpívala v italštině svou vlastní píseň Caruso, k níž letos v březnu natočila na Bali také videoklip.

6. Vysvětli význam přísloví a pořekadla.

Tichá voda břehy mele,

Nejméně vytrvalá voda nejvíce opouští

Boží mlýny melou pomalu, ale jistě.

Na každém kámen dopadne

Ranní ptáče dál do skáče.

První dráček nosíme, více toho získáme

Na hrubý pytel hrubá záplata

Na špatné chování odpovídá stejně špatným

Jiný kraj jiný mrav

Jiné město, jiné mravy

7. Napiš druhou část přirovnání.

Chytrý jako liška

Mlsný jako medvěd

Ostrý jako nůž

Pomalý jako šnek

Sladký jako med

8. Doplň číslovku do přísloví.

Kdo šetří má za šest

Každá mince má dvě strany.

První vlaštovka jaro nedělá.

Když se dva perou, první se směje.

Devatero řemesel desáta bída.