

Úvod do výuky češtiny jako druhého/cizího jazyka

Barbora Hrabalová

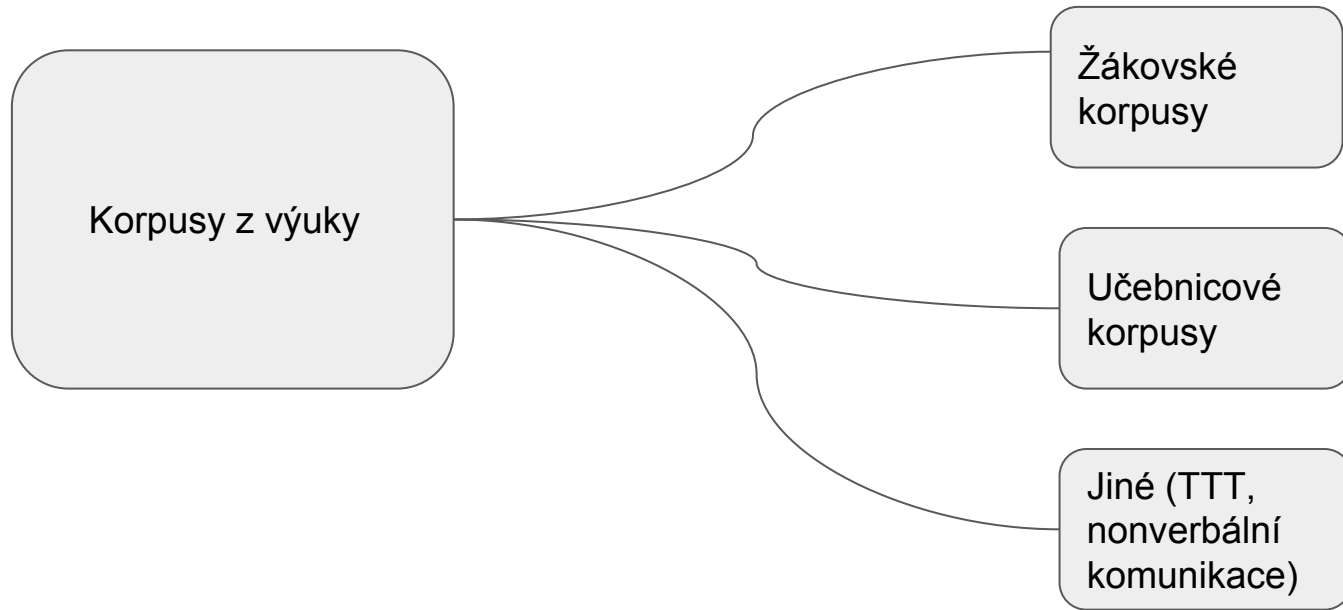
Přednáška 4

Termíny a klíčová slova

- korpus (obecný X speciální)
- L1 / L2
- SERR
- jazykový input/output
- jazykový transfer
- kognáty a falešní přátelé
- mezijazyk
- DDL

Korpusy pro výuku
(obecné korpusy)

Korpusy z výuky
(speciální korpusy)



Žákovské korpusy

- žák jako ekvivalent k anglickému “learner” = learner corpus
- 90. léta 20. století
- podtyp akvizičních korpusů
- klasifikace:
 - čas sběru = synchronní X sběrné (longitudinální)
 - data = L1, tasks, úroveň
 - přístup = komerční X nekomerční
- specifické parametry budování korpusu (data, velikost, autentičnost, metadata, anotace)
- výzkum: kontrastivní analýza, chybová analýza, teorie mezijazyka

Jak probíhá osvojování druhého/cizího jazyka?

- L. Selinker (70. léta 20. století) = autor teorie mezijazyka (v originále interlanguage)
- mezijazyk je svébytný jazykový útvar, proměnlivý a dynamický systém, který obsahuje řadu nestandardních prvků
 - míšení kódů
 - chyby
 - jazykový transfer, pozitivní (kognáty) a negativní (false friends)
 - naučené fráze versus improvizace

Výzkum je zaměřen na popis procesů a utváření mezijazyka či mimojazykové vlivy (např. věk).

Žákovské korpusy u nás


- 167 žákovských korpusů ve světě (květen 2018, Learner corpus association)
- projekt AKCES (2005)
 - Ústav českého jazyka a teorie komunikace FF UK + další pracoviště
 - prof. PhDr. Karel Šebesta, CSc.
 - osvojování češtiny jako L1 i L2
 - žákovský korpus češtiny jako cizího jazyka
- databáze CHRUP (2013)
 - prof. Ph.Dr. Hana Gladkova, CSc., FF UK
 - didaktický cíl: specifika osvojování L2, který je blízký L1
- mezinárodní žákovský korpus MERLIN (2012)

Žákovský korpus CzeSL - SGT

- Czech as a Second Language with Spelling, Grammar and Tags
- Alexandr Rosen, Barbora Štindlová, Svatava Škodová, Jirka Hana, Tomáš Jelínek, Vladimír Petkevič a další
- Ústav českého národního korpusu
- 2014
- 958 tisíc slov
- A1-C1
- rozsáhlá metadata
- částečná automatická chybová analýza

Korpus: czesl-sgt

Hledat v korpusu

Korpus: / Typ dotazu: [Klávesnice](#) | [Předchozí dotazy](#)

Dotaz:

V dotazu CQL můžete vložit další řádek stisknutím klávesy Shift+ENTER  **Specifikovat kontext**  **Omezit hledání**

Using MERLIN

[Simple search](#)

[Advanced search](#)

[Define a subcorpus](#)

[Statistics](#)

MERLIN for research

MERLIN corpus

MERLIN annotations


Documentation


Download whole corpus

Search for words in the learner texts and display them in context (e.g., full text):

examples: *welchen*, *gli*, *rád*

[Search for word](#)  

Search in: 

Corpus: 

No subcorpus yet? Compile your own text collection » [here](#).

The MERLIN corpus

MERLIN provides access to 2.286 texts written by learners of Czech, Italian and German.

The learner texts stem from standardized language tests and they have been reliably related to the CEFR levels. » [read more](#)

Use MERLIN ...

... to better understand the levels of the Common European Framework of Reference (CEFR). » [read more](#)

Video tutorial ...



Pozitiva žákovských korpusů

- zpřístupnění dat pro analýzu žákovského jazyka
- nový nástroj pro výzkum procesů osvojování druhého/cizího jazyka
- napomáhají rozvoji korpusových nástrojů (např. pro automatickou chybovou analýzu)
- mají široké využití pro tvorbu výukových a studijních materiálů

Ukázka využití informací z ŽK ve studentském slovníku nakladatelství Longman

Grammar: information


Do not say "an information" or "informations." Say, for example: some information, a lot of information, or a piece of information:

- Do you have any information about tickets?

- I need a couple more pieces of information before I can finish my report.

Zdroj: The Longman Corpus Network, 2015, online.

Využití ŽK v učebnicích z nakladatelství Cambridge University Press

 Spelling mistakes are among the most frequent errors made by candidates in the exam. Find and correct the spelling mistakes made by candidates in the sentences below.

1 As you can see in the advertisement, the holiday is quite cheap.

Zdroj: Brook-Hart – Haines, 2014–2015, s. 21.

Negativa žákovských korpusů

- sběr dat je velmi složitý a časově náročný
- je velmi složité dosáhnout vyváženosti dat a vytvářet analýzy s obecnou platností (například v korpusu CzeSL převažují data od rusky mluvících studentů)
- strategie vyhýbání se (angl. avoidance)- není možné získat veškerá možná data, jelikož se uživatelé mohou cíleně vyhýbat jevům, u nichž si nejsou jisti jejich užitím (např. gramatikou, výslovností atd.)
- sledujeme pouze produktivní dovednosti, receptivní dovednosti zůstávají upozadněny

Učebnicové korpusy

- začínají vznikat po roce 2000
- Ute Römer
- doplňují klasické analýzy učebnic, která se zaměřuje na různé aspekty tvorby učebnic (strukturování, design, zastoupení jazykových dovedností, funkce didaktických ilustrací atd.)
- zájem o jazyk
 - Jaký je jazyk instrukcí?
 - Jaká jsou specifika učebnicového jazyka?
 - Jaká slovní zásoba je používána? Jaká je nadužívána/podužívána?
 - stereotypizace

Učko1, Učko2

- projekt korpusu učebnic češtiny pro cizince
- autorkou je Mgr. Pavlína Vališová (Meta, o.p.s)
- neveřejný
- kompilován a uložen ve Sketch Engine
- 18 učebnic pro úrovně A1-B1
- obsahuje modelové věty, přepisy nahrávek, delší texty ze cvičení
- aktuální velikost je 100 tisíc slov

Korpusy pro výuku

Korpusy z výuky

Korpusy pro výuku

- nepřímé využití
 - zdroj pro frekvenční analýzy, zdroj pro tvorbu výukových materiálů, zdroj pro jazykové plánování
 - projekty, které využívají ČNK - Česky krok za krokem (2016, L. Holá), aplikace Ten Ta To (2015), Česká čítanka (2012, I. Kořánová)
- přímé využití
 - studenti používají korpusy nebo data z korpusů přímo ve výuce (přístup DDL)
 - projekt Korpusy pro školy - repozitář cvičení, popularizace korpusu ze strany ÚČNK
 - nástroj vystavěný podle zásad DDL - Sketch Engine for Language Learning

Data Driven Learning

- přístup ve výuce cizích jazyků, který je založený na předpokladu, že student potřebuje autentická data = korpus
- podporuje induktivní přístup (student hledá v datech, vybraná data klasifikuje a tvoří vlastní generalizace)
- další podobné přístupy: corpus-based teaching, computer-assisted language learning ...
- Tim Johns (1991), Ute Römer, Alex Boulton, James Thomas, Pavlína Vališová, Adrian Zasina ...
- korpus může být využíván pro výuku L1 i L2
- dvě varianty
 - hands on (student pracuje přímo v korpusu)
 - hands off (student dostává materiály od učitele)

Odkazy

Learner Corpus Association <https://www.learnercorpusassociation.org/>

Projekt AKCES <http://akces.ff.cuni.cz/>

Databáze CHRUP <http://chrup.ff.cuni.cz/About.aspx>

Mezinárodní žákovský korpus MERLIN <https://merlin-platform.eu/>

Projekt Korpusy pro školy <https://www.korpus.cz/proskoly>

SKELL <https://skell.sketchengine.eu/#home?lang=en>

Literatura

ŠEBESTA, Karel. *Vyučování cizího jazyka: terminologický slovník*. Praha: Filozofická fakulta Univerzity Karlovy, 2017.

ŠEBESTA, Karel. *Druhý a cizí jazyk: osvojování a vyučování: terminologický slovník*. Praha: Filozofická fakulta Univerzity Karlovy, 2014.

ŠORMOVÁ, K., ŠEBESTA K. a kol. *Korpusy v jazykovém vyučování*. Praha: Filozofická fakulta Univerzity Karlovy, 2019.

ŠTINDLOVÁ, Barbora. *Žákovský korpus češtiny a evaluace jeho chybové anotace*. Vydání první. V Praze: Filozofická fakulta Univerzity Karlovy, 2013.

Computer learner corpora, second language acquisition, and foreign language teaching. Edited by Sylviane Granger - Joseph Hung - Stephanie Petch-Tyson. Amsterdam: John Benjamins, 2002.

JOHNS, T. Should you be persuaded. Two samples of data-driven learning materials. *Classroom Concordancing: ELR Journal*, 4, s. 1–16, 1991.