

Abstrakt

Lexikální rozmanitost je termín popisující rozsah užití slovní zásoby v textu. Jedná se o podkategorii jazykové komplexity, která je jedním ze tří základních komponentů teorie označované CAF, která zahrnuje komplexnost (complexity), přesnost (accuracy) a plynulost (fluency). Tyto termíny jsou v lingvistickém výzkumu používány k popisu jazykové úrovně rodilých (L1) mluvčích a jsou také součástí např. rámce CEFR, jenž klasifikuje nerodilé (L2) mluvčí do kategorií na základě souhrnných jazykových dovedností. Většina zkoumání v oblasti lexikální rozmanitosti se i dnes soustřeďuje na jazyk psaný a analýza mluveného projevu bývá opomíjena, jelikož je třeba promluvy transkribovat a získaná data vytržít. Východiskem může být použití již zkompilevaného korpusu a tato možnost byla také využita k získání dat pro tuto práci. Jedná se o subkorpus LINDSEI_CZ, jenž zaznamenává transkripci českých L2 mluvčích angličtiny, kteří byli zařazeni do jazykových úrovní B2 až C2 dle rámce CEFR. Cílem této práce bylo zjistit, zda lze v projevu mluvčích na úrovni B2 a C1 v daném korpusu pozorovat signifikantní rozdíl v lexikální rozmanitosti. Bylo zkoumáno všech dvanáct mluvčích na úrovni B2 a k nim bylo náhodným výběrem přiřazeno dvanáct mluvčích C1. Pro porovnání jejich jazykové rozmanitosti byly použity dva indexy a to TTR (Type-Token Ratio) a MTLN (Measure of Textual Lexical Diversity), které jsou dostupné v online nástroji Text Inspector. Hypotéza této práce předjímala rozdíl mezi dvěma zmíněnými skupinami mluvčích. Ani jeden z užitých indexů ovšem nepozoroval signifikantní rozdíl v lexikální rozmanitosti B2 a C1 mluvčích. Tento výsledek je možné přisoudit malému vzorku mluvčích nebo jejich velmi málo rozdílným kompetencím. Lze také spekulovat o přesnosti vyhodnocení úrovně jednotlivých mluvčích, nebo o tom, zda je lexikální rozmanitost zásadním faktorem při vnímání pokročilosti v mluveném projevu. V psaném projevu je užitá slovní zásoba klíčovým indikátorem jazykových dovedností, ale v mluveném projevu může hlavní roli přebírat plynulost.

Klíčová slova: CAF, lexikální rozmanitost, mluvený projev, L2 mluvčí, LINDSEI_CZ, korpus, TTR, MTLN, Text Inspector, CEFR