

Abstrakt

Diplomová práce analyzuje lexikální komplexnost písemného jazyka nerodilých mluvčích (L2) anglického jazyka na úrovni C2 s následnou komparací výsledné lexikální komplexnosti s komplexností u rodilých mluvčích (L1). Lexikální komplexnost jazyka je sledována ve dvou klíčových parametrech, v dimenzi lexikální rozmanitosti (lexical diversity) a lexikální propracovanosti (lexical sophistication). Práce obsahuje kvantitativní analýzu dat s pomocí jednoduchých indikátorů, která je doprovázena analýzou s kompozitními metrikami VOCD-D a MTLD v oblasti lexikální rozmanitosti. Lexikální sofistikovanost je analyzována prostřednictvím nástroje English Vocabulary Profile (EVP), který měří výskyt slov v textu na základě předdefinovaného frekvenčního seznamu slov podle kategorií jazykové úrovně A1-C2. Základní stanovená hypotéza předpokládala nižší lexikální komplexnost nerodilých mluvčích jazyka ve srovnání s rodilými mluvčími, a to i přes nejvyšší možnou jazykovou kompetenci nerodilých mluvčích na úrovni C2, která je přirovnávána k „native-speaker competence”. V rámci jednotlivých skupin uživatelů jazyka (L2 a L1) se očekávaly srovnatelné výstupy. Další pomocnou hypotézou byl nižší předpokládaný počet slov na úrovni C2 (low-frequency words) u L2 ve srovnání s L1.

Soubor dat, které byly podrobeny analýze, obsahuje 20 srovnatelných textů nerodilých mluvčích anglického jazyka na úrovni C2 co do jejich rozsahu (s průměrnou délkou 781 slov), žánru (ekonomie), cíle (napsání komentáře k publikovanému článku) a struktury (doporučená struktura komentáře), a 20 textů rodilých mluvčích anglického jazyka opět s podobnými parametrickými vlastnostmi. Data byla analyzována ve svých skupinách (L2 a L1) samostatně a také vzájemně mezi sebou. Pro analýzy byl využit automatizovaný webový nástroj Text Inspector, jehož bližší popis je součástí práce.

Lexikální rozmanitost byla srovnatelná v rámci jednotlivých skupin. Při komparaci byla prokazatelně vyšší u L1 uživatelů jazyka a to při použití obou nástrojů (VOCD-D a MTLD). Dalším poznatkem je, že L2 uživatelé mají tendenci více opakovat stejná slova bez většího využití synonymie. Výsledky srovnávání v dimenzi lexikální propracovanosti s využitím frekvenčních seznamů slov nepotvrdily hypotézu, neboť nebyly statisticky signifikantní. Rozdělení slov podle jednotlivých slovních kategorií se ukázalo jako přibližně stejné s klesajícím trendem od A1 do C2. Problémem je zde větší množství nekategorizovaných slov, což je důsledkem nástroje EVP, který mnohdy neobsahuje specifickou slovní zásobu z oblasti ekonomie, a také specifčnosti ekonomických textů, které obsahují více čísel, vlastních jmen a

zkratk (převážně akronymů) nežli texty obecné povahy. Pro větší spolehlivost výsledků by bylo vhodné výzkum dále rozšířit, ať už by se jednalo o větší délku textů nebo o jejich množství. Studie poukázala na citlivost žánru a proto by bylo žádoucí v dalším případném výzkumu stejný žánr dodržet. Dalším ovlivněním výzkumu je tzv. „priming effect“ (tendence L2 studentů opakovat stejná slova vyskytující se v článcích L1 autorů) , který automatizované nástroje nejsou schopny zachytit.

Klíčová slova

Lexikální komplexnost, jazyk nerodilých mluvčích, písemný jazyk, frekvenční seznam slov, English Vocabulary Profile (EVP), Text Inspector