
A Study of Lexical Diversity in Multilingual Student's Texts

Elke G. Montanari¹, Gertrud Faäß¹ & Aivars Glaznieks²

¹Universität Hildesheim, ²Eurac Research, Bozen/Bolzano

elke.montanari@uni-hildesheim.de, gertrud.faass@uni-hildesheim.de,

aivars.glaznieks@eurac.edu

In general, multilingual children have lower vocabulary competence than monolingual children when considering only one language (but more when considering all languages). If sufficient input is provided in all languages, however, the Complementarity Principle (Grosjean 2016) hypothesizes equivalent vocabulary for monolinguals and multilinguals speaking the same language. It is our objective to evaluate the bilingual vocabulary of monolingual and bilingual students who have been studying together for several years and should, according to the Complementarity Principle, have the same vocabulary. For the analysis of student texts, we will use three sets of texts taken from the LEONIDE corpus (Glaznieks et al., 2022). The sets were built to take into account the language background of the students (monolingual German-speakers, bilingual German-speakers, and multilingual non-L1-German-speakers). Due to the fact that the texts of the students are not equal, we will utilize the Lexical Diversity Measure to compare the vocabulary of each student. It has been shown that learner characteristics profoundly influence measures of lexical diversity, range, and sophistication (e.g., Meurers 2015). This poster will present an evaluation of different tools and measures for lexical diversity, including an evaluation of a state-of-the-art tagger. Considering school-related topics, we expect that the vocabulary of monolingual and bilingual students after a sufficient period of schooling is equally large. Moreover, we hope to gain knowledge about which measures of lexical diversity are more useful for (short) student texts. Consequently, lexical diversity measures would likely be useful tools for comparing the vocabulary of individual texts, and the expected results would have important implications for the judgement of the vocabulary of individuals.

References:

- Glaznieks, A., Frey, J.-C., Stopfner, M., Zanasi, L. & Nicolas, L. (2022). LEONIDE: A longitudinal trilingual corpus of young learners of Italian, German and English. *International Journal of Learner Corpus Research* 8(1), 97–120. <https://doi.org/10.1075/ijlcr.21004.gla>
- Grosjean, F. (2016). The Complementarity Principle and its impact on processing, acquisition, and dominance. In J. Treffers-Hammou, B. A., Larouz, M., & Fagroud, M. (2021). Word frequency, Range and Lexical diversity: Picking out Changes in Lexical Proficiency among University Learners in an EFL Context. *International Journal of Linguistics and Translation Studies*, 2(2), 22-38.
- Meurers, D. (2015). Learner corpora and natural language processing. *The Cambridge Handbook of Learner Corpus Research*, 537–566.