



Contribution ID: 127

Type: **Individual paper / Exposés individuels / Einzelvorträge**

Lexical Diversity in Multilingual Student Texts

Wednesday 23 July 2025 15:00 (20 minutes)

Mehrsprachige Kinder zeigen in Wortschatztests geringere Werte als einsprachige Kinder, sofern die Auswertung nur eine Sprache berücksichtigt, aber höhere Werte, wenn alle Sprachen einbezogen werden. Der Beitrag soll dieses in zahlreichen Studien replizierte Ergebnis kritisch auf der Grundlage des Komplementaritätsprinzips (CP, Grosjean 2016) diskutieren und eine innovative Methodik vorschlagen. Grob gesagt, stützen sich bisherige Testverfahren für Wortschatz in der Regel auf den Alltag als Handlungsfeld, der typischerweise jedoch von mehrsprachigen SchülerInnen zu einem großen Teil mehrsprachig erlebt wird; sie verteilen also ihre Sprachen komplementär auf einzelne Handlungsbereiche (CP). Zielführender für Voraussagen zum Wortschatz in Bildungskontexten wäre es daher, dies in das Testverfahren einzubeziehen.

Unser Ziel ist es, den Wortschatz von einsprachigen und zweisprachigen Schülern zu evaluieren, die seit mehreren Jahren zusammen lernen und für die das CP einen vergleichbaren Wortschatz prognostiziert. Dabei wendet sich die vorgeschlagene Methode in einem ersten Schritt an die wissenschaftliche Öffentlichkeit; bei positiver Evaluierung ist sie auch für die Anwendung in der Schule geeignet.

Im Grundsatz schlagen wir vor, nicht Individuen zu testen, sondern Texte aus dem schulischen Umfeld mit computerlinguistischen Methoden zu untersuchen. Für die Analyse der Schülertexte werden wir drei Textsätze aus dem LEONIDE-Korpus (Glaznieks et al., 2022) verwenden. Die Sets wurden so zusammengestellt, dass sie den sprachlichen Hintergrund der Schüler berücksichtigen (einsprachige Deutschsprachige, zweisprachige Deutschsprachige und mehrsprachige Nicht-L1-Deutschsprachige). Da die Texte der Schüler nicht gleich sind, werden wir Maße der lexikalischen Diversität einsetzen (z.B. Meurers 2015). Der Vortrag wird eine Evaluierung verschiedener Instrumente und Maße für lexikalische Vielfalt versuchen. In Bezug auf schulische Themen zeigen die Daten, dass der verwendete Wortschatz auch bei Betrachtung nur einer Sprache von monolingualen und bilingualen Schülern nach einer ausreichenden Schulzeit äquivalent ist.

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.
- Kyle, K., Crossley, S. A., & Jarvis, S. (2021). Assessing the Validity of Lexical Diversity Indices Using Direct Judgements. *Language Assessment Quarterly*, 18(2), 154-170. <https://10.1080/15434303.2020.1844205>
- Meurers, D. (2015). Learner corpora and natural language processing. *The Cambridge Handbook of Learner Corpus Research*, 537-566.
- Schmidt, H. (2019). Deep Learning-Based Morphological Taggers and Lemmatizers for Annotating Historical Texts, DATeCH, May 2019, Brussels, Belgium.

Content/contenu/Inhalt 2

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 vocabulary of monolingual and bilingual students who have been studying together for several years and should, according to the Complementarity Principle, have a comparable 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 try an evaluation of different tools and measures for lexical diversity, including an evaluation of the RNN-tagger developed by Schmidt (2019). 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 multilingual individuals.

Title/ titre/Titel 2

Lexical Diversity in Multilingual Students' Texts

Only for symposia / Seulement pour symposiums / Nur für Symposien

Authors: Dr GLAZNIEKS, Aivars (EURAC); MONTANARI, Elke G.

Presenter: MONTANARI, Elke G.

Session Classification: Slot 5

Track Classification: 3: Research and dissemination / Recherche et diffusion / Forschung und Dissemination