



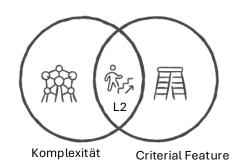


Automatisierte Werkzeuge zur Analyse sprachlicher Komplexität und Verständlichkeit

Denise Löfflad, Detmar Meurers

Wie können sprachliche Merkmale quantifiziert werden?

In der Forschung zum Zweitspracherwerb (L2) stehen traditionell Maße zu Komplexität, Genauigkeit und Sprachfluss (CAF-Triade) im Mittelpunkt (Weiss, 2024). Für die Praxis mit Lehrkräften ist jedoch der Bezug zum Common European Framework of Reference (CEFR) über sogenannte Criterial Features (Hawkins & Buttery, 2010) besonders geeignet. Wir stellen für beide automatisierte Werkzeuge bereit.



Komplexität

Aggregierender Ansatz:

- Teil der CAF-Triade zur Messung von L2-Kompetenz
- Charakterisiert Elaboriertheit und Variabilität sprachlicher Realisierungen (Bulté, Housen, Pallotti, 2025)
- Abstrahierend über sprachliche Realisierungen
- Immer bezogen auf messbare Eigenschaften von Texten

Criterial Features



Sprachentwicklungsbasierter Ansatz:

- Ableitung linguistischer Muster aus Entwicklungsverläufen im Zweitspracherwerb
- Fokus auf Merkmale, die Ewerbsstufen abbilden und differenzieren (Hawkins and Buttery, 2010)
- Enge Anbindung an den CEFR, aber detaillierter und datenbasiert (Gaillat et al., 2022)

Komplexität messen

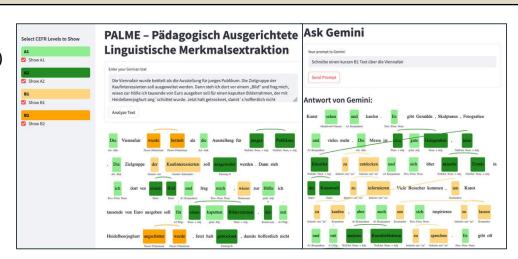


CTAP Plattform (Chen and Meurers, 2016)

- CTAP (Common Text Analysis Platform) (Chen and Meurers, 2016; Weiss, 2024): webbasierte Plattform für automatische Sprachanalyse
- >1.000 Maße in fünf Sprachen (Syntax, Lexikon, Morphologie, Diskurs, Sprachgebrauch)
- Erfolgreich eingesetzt für Lesbarkeitsmodelle (Weiss & Meurers, 2018) oder die Analyse der Sprache von Lehrkräften (Heppt et al., 2025)

Criterial Features messen

- Entwicklung des German Grammar Profile (GGP) mit 150 Criterial Features
- Top-Down-Ansatz, von Expert:innen geprüft für sprachliche Validität & pädagogische Relevanz
- PALME (Löfflad et al., 2025): Webplattform zur automatischen Extraktion mit direkt nutzbaren Einblicken für Lehrkräfte
- Work in progress: derzeit sind 70 Merkmale implementiert



PALME Plattform (Löfflad et al., 2025).

5.

Simpkin, A., Ballier, N., Steams, B., Sousa, A., Bouyé, M., & Zarrouk, M. (2022). Predicting CEFR levels in learners of English: The use of Meurers, D. (2016). CTAP: A web-based tod supporting automatic complexity analysis.

(2024). An Integrative Approach to Linguistic Complexity Analysis for German (Doctoral dissentation, Eberhard Karls Universität Tütingen).

Is Meurers, D. (2018, August). Modeling the readability of German targeting adults and children: An empirically broad analysis and its cross-Beuttler, B., & Meurers, D. (2025, September). German Grammar Profile for Learners: Pedagogical Feature Definition and Automated Extremence. Volume 2: Workshops (Vol. 9, p. 212).

Leibniz-Institut für Wissensmedien AG Sprache und KI in der Bildung Schleichstr.6, 72072/Tübingen d.loefflad@iwm-tuebingen.de