



Entwicklung von textuellen Oberflächenmerkmalen in Schüleraufsätzen in den Klassen 4, 8 und 10

Ermittlung von Indikatoren für die Kompetenzeinschätzung von Schülertexten mittels computerlinguistischer Methoden und dialogische Entwicklung eines Prototyps für die computergestützte Analyse von Aufsätzen

Jasmine Bennöhr

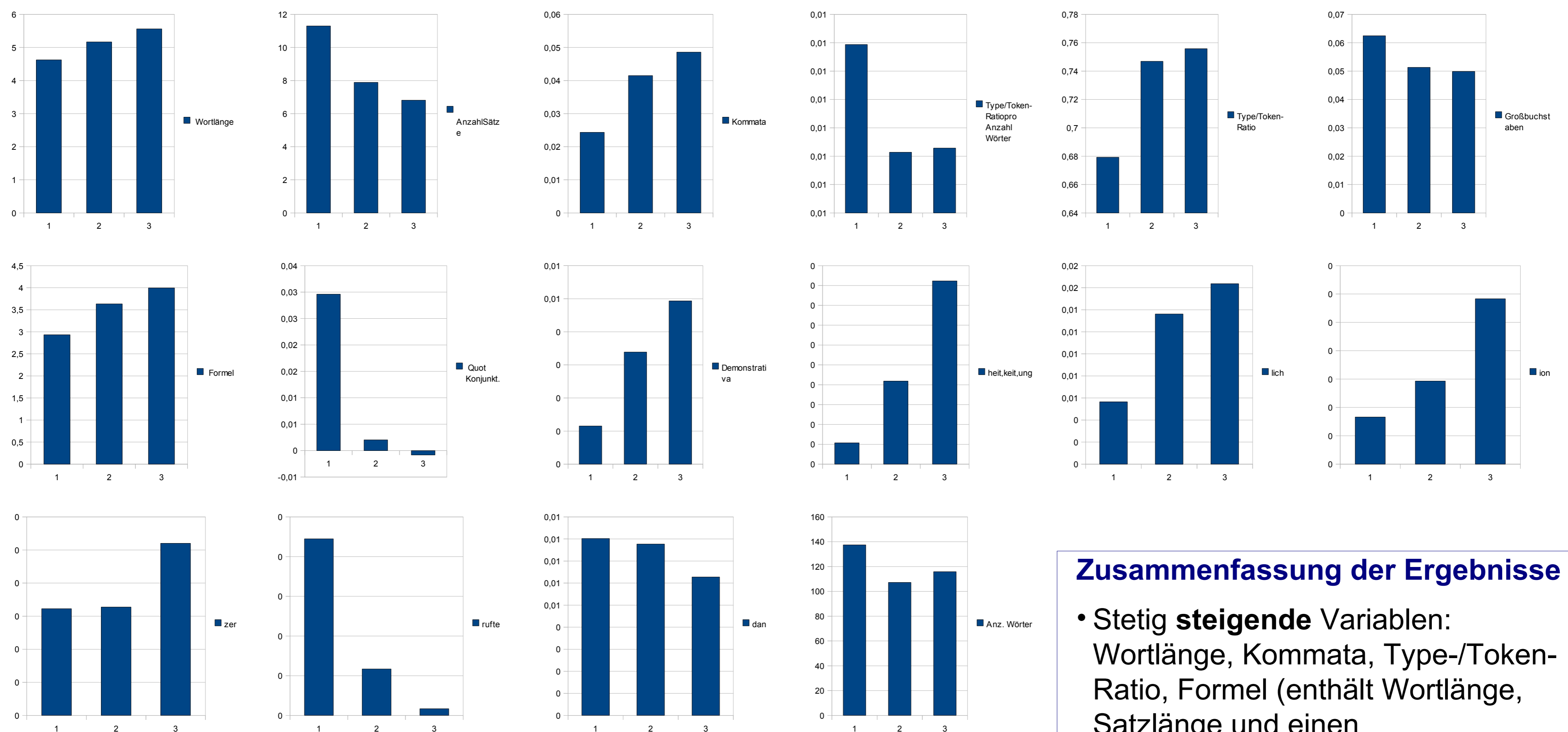
Humboldt-Universität zu Berlin & Landesinstitut für Lehrerbildung und Schulentwicklung, Hamburg

Daten

- **Schüleraufsätze in deutscher Sprache aus der Hamburger Längsschnittstudie KESS** (Kompetenzen und Einstellungen von Schülerinnen und Schülern) study, thematische Streuung
- Für diesen Vergleich wurden **706** Aufsätze aus **Klasse 4**, **1751** Texte aus **Klasse 8** und **1192** Aufsätze aus **Klasse 10** herangezogen
- Rohdaten verwendet, **Zielhypothesen** sollen in Zukunft berücksichtigt werden

Hinweise zur Interpretation der Diagramme

- **Normalisierung der Variablen:** meist Anzahl pro Wort
- 1, 2, 3 entspricht den Klassenstufen 4, 8, 10
- Analyse durchgeführt mit **Rohdaten**, die von Schülern gemachte **Fehler** enthalten z. B. Interpunktionsfehler (zu lange Sätze) oder Rechtschreibfehler (versuchen vs. fersuchen)



Ausblick

- Klasse 7 und Texte von Erwachsenen als Vergleichskorpus in Analyse mit einbeziehen
- Erhöhung der Qualität und Quantität des KESS-Korpus
- Differenziertere Betrachtungen, z. B. nach Themengebiet differenzieren
- Vergleich erneut mit **Zielhypothesenkorpus** durchführen (z. B. Auswirkungen auf Satzlänge)
- Weitere **Oberflächenmerkmale**
- Weiterführende statistische Analysen

References

- Steinig, W., Betzel, D., Geider, F. J., Herbold, A. (2009). **Schreiben von Kindern im diachronen Vergleich**. Texte von Viertklässlern aus den Jahren 1972 und 2002. Waxmann. Münster.
- Feilke, H. (1996). **From Syntactical to Textual Strategies of Argumentation**. Syntactical Development in Written Argumentative Texts by Students Aged 10 to 22. *Argumentation* 10, pp. 197-212.

Zusammenfassung der Ergebnisse

- **Stetig steigende** Variablen: Wortlänge, Kommata, Type-/Token-Ratio, Formel (enthält Wortlänge, Satzlänge und einen Konjunktionsquotienten, Demonstrativa, -heit/-keit/-ung, -lich, -ion und zer-
- Satzlänge sinkt anders als bei Feilke (1996), Vorzeichen kehrt sich vermutlich um, wenn Zielhypothesenkorpus als Analysegrundlage dient
- **Stetig fallende** Variablen: Anzahl Sätze, Großbuchstaben, Konjunktionsquotient aus leichten und schweren Konjunktionen, fehlerhafte Wörter rufte und dan
- Thematische Streuung hat Einfluss auf die Textlänge
- **Wortlänge** ist einer der stabilsten Indikatoren (vgl. Lesbarkeitsformeln) ist aber weder in Feilke (1996) noch in Steinig et. al (2009) erwähnt