

Gute Wörter für Delta: Verbesserung der Autorschaftsattribu- tion durch autorspezifische distinktive Wörter

Friedrich-Michael Dimpel · Thomas Proisl

Friedrich-Alexander-Universität Erlangen-Nürnberg

Einleitung

- Burrows Delta und Varianten: etablierte Verfahren zur Autorschaftsattribu-
tion
- Abstands-basiertes Verfahren zur Autorschaftsattribu-
tion
 - Relative Häufigkeiten der n häufigsten Wörter
 - Standardisierung zu z-Scores
 - Abstandsmaß (Burrows Delta: Manhattan-Distanz)
- Steigerung der Erkennungsquote durch autorspezifische distinktive
Wörter möglich?

F	Wh1	Pz2	Trist3	Mean	StDev	Z1	Wh	Z2	Pz	Z3	Gleiche Autoren		Versch. Autoren		L2Diff
											Abs Diff	Abs Diff	Abs Diff	Abs Diff	
1 der	3,19	2,62	2,13	2,65	0,53	1,02	-0,05	-0,98	1,07	2,00	1,07	2,00	0,93		
2 daz	2,02	1,93	2,77	2,24	0,46	-0,48	-0,67	1,15	0,19	1,63	1,15	0,19	1,63	1,44	
3 und	1,97	1,01	3,52	2,17	1,27	-0,15	-0,91	1,07	0,76	1,22	1,07	0,76	1,22	0,46	
4 ir	1,33	1,81	2,14	1,76	0,41	-1,06	0,12	0,93	1,18	1,99	1,18	1,99	0,81		
5 er	1,51	1,63	1,97	1,70	0,24	-0,80	-0,32	1,12	0,48	1,92	1,12	0,48	1,92	1,44	
6 in	1,30	1,25	1,66	1,40	0,23	-0,48	-0,67	1,15	0,19	1,63	1,15	0,19	1,63	1,44	
7 ich	1,29	1,43	1,30	1,34	0,08	-0,67	1,15	-0,48	1,82	0,18	1,82	0,18	-1,63		

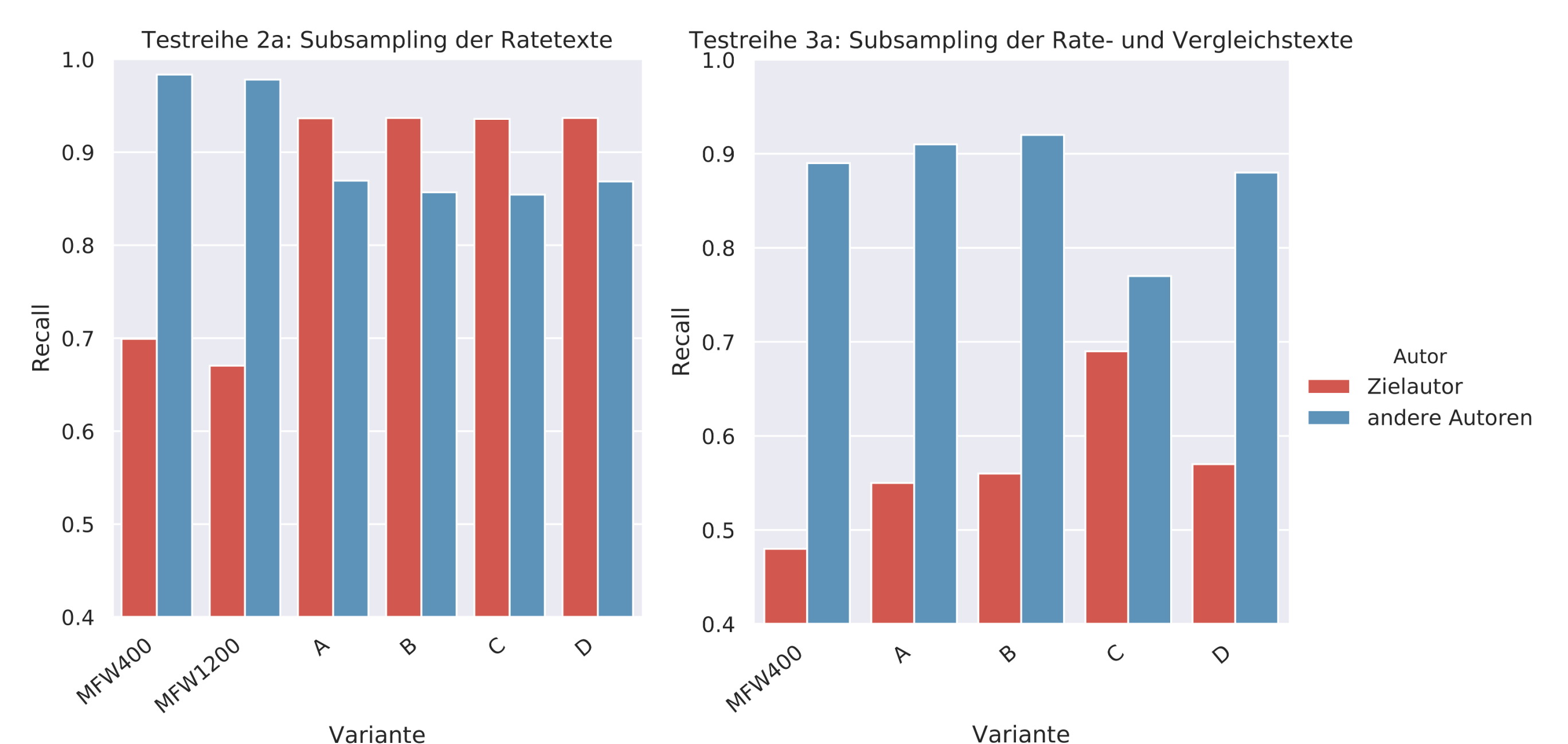
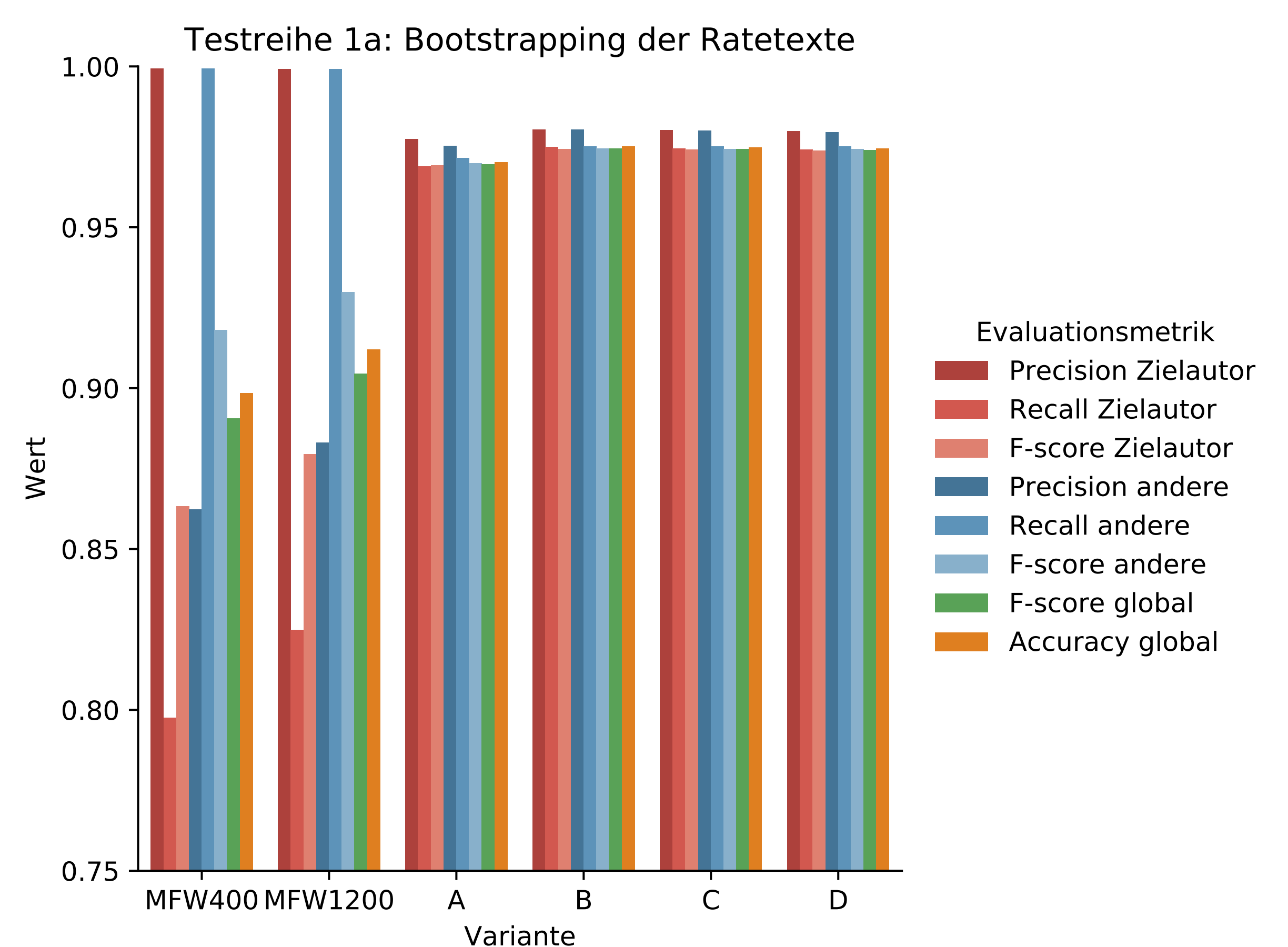
- Ermittlung der distinktiven Wörter über **Level-2-Differenzen**, d.h. Diffe-
renz zwischen
 1. z-Wert-Differenzen zwischen Autor- und Distraktortexten
 2. z-Wert-Differenzen zwischen den Autortexten
- Differenzen zwischen den Autortexten sollen kleiner sein als Differen-
zen zwischen Autortexten und Distraktortexten

$$I = \frac{\sum_{a_i \in A} \sum_{a_j \in A \setminus a_i} \sum_{d \in D} (\text{abs}(a_i - d) - \text{abs}(a_i - a_j))}{|A| |D| (|A| - 1)}$$

Experimente

- Evaluationskorpus: 280 deutsche Texte von 32 Autoren (ca. 19 Jhd.)
- Evaluationstexte nicht zum Ermitteln der guten Wörter verwendet
- Experimente auf ganzen Texten und auf 2.000-Wort-Stichproben
- Untersuchte Varianten:
 - A:** Gute Wörter: Mittelwerte aller Level-2-Differenzen $\geq 0,4$
 - B:** Gute Wörter: Mittelwerte aller Level-2-Differenzen $\geq 0,2$
 - C:** Wie A, zusätzlich Spitzenwert-Kriterium (Level-2-Differenz $> 1,64$)
 - D:** Wie C, Spitzenwert-Kriterium $> 1,2$; Mittelwerte $\geq 0,4$; schlechteste Level-2-Differenzen ignorieren
- 3×2 Testreihen:
 - Testreihe 1a:** Bootstrapping der Ratetexte. Vergleichskorpus enthält je einen Text von allen 32 Autoren. Ratekorpora für jeden Autor umfas-
sen drei Texte des Zielautors und drei Texte anderer Autoren.
 - Testreihe 1b:** Ratekorpus wie 1a. Vergleichskorpus: Ein Text des Ziel-
autors, 24 Texte anderer Autoren, jedoch nicht die richtigen Autoren
der drei anderen Autoren der Ratekorpora.
 - Testreihe 2a/b:** Rate- und Vergleichskorpus wie unter 1a/b, jedoch
Subsampling der Ratetexte (2.000 Wörter).
 - Testreihe 3a/b:** Alternatives Setting bei Rate- und Vergleichskorpus.
Beim Berechnen der False-Positives-Quote Sampling der Rate-
texte und der Vergleichstexte (2.000 Wörter).

Ergebnisse



- Testreihen 1b, 2b und 3b liefern ähnliches Bild mit niedrigeren Werten

Effekt der guten Wörter

- Deutlich bessere globale F-Scores (+ 6–8 Punkte)
- Deutlich höherer Recall beim Zielauteur, dafür etwas höhere False-
Positive-Rate (= niedrigerer Recall bei anderen Autoren)

Methode	Zielauteur	andere Autoren
klassisches Delta	Precision > Recall	Precision < Recall
gute Wörter	↓ Precision > ↑ Recall	↑ Precision < ↓ Recall

Fazit und Ausblick

- Bessere Erkennungsquoten durch gute Wörter
- Leicht höhere False-Positive-Rate, doch positive Effekte sind größer
- False-Positive-Rate stark autorabhängig
→ Vorab-Test der Anwendbarkeit auf spezifischen Autor denkbar
- Weitere Verbesserungen durch maschinell gelerntes autorspezifisches
Vokabular?