

Artikel erschienen in:

*J.M.M. Brown, Andreas Schmidt,
Marta Wierzba (Eds.)*

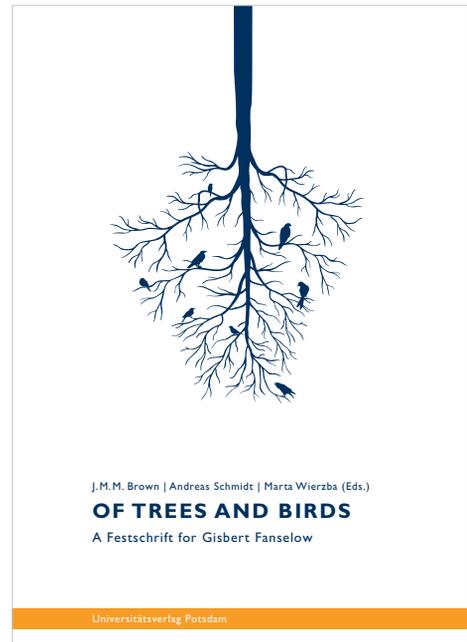
Of trees and birds

A Festschrift for Gisbert Fanselow

2019 – 435 S.

ISBN 978-3-86956-457-9

DOI <https://doi.org/10.25932/publishup-42654>



Empfohlene Zitation:

Stede, Manfred: Noch kindlich oder schon jugendlich? Oder gar erwachsen? Betrachtung von Komplexitätsmerkmalen altersspezifischer Texte, In: Brown, J.M.M. / Schmidt, Andreas / Wierzba, Marta (Eds.): *Of trees and birds*. A Festschrift for Gisbert Fanselow, Potsdam, Universitätsverlag Potsdam, 2019, S. 323–334.

DOI <https://doi.org/10.25932/publishup-43256>

Soweit nicht anders gekennzeichnet ist dieses Werk unter einem Creative Commons Lizenzvertrag lizenziert: Namensnennung 4.0. Dies gilt nicht für zitierte Inhalte anderer Autoren: <https://creativecommons.org/licenses/by/4.0/deed.de>

Noch kindlich oder schon jugendlich? Oder gar erwachsen?

Betrachtung von Komplexitätsmerkmalen altersspezifischer Texte¹

Manfred Stede, Universität Potsdam

1 Einführung

Dass Texte unterschiedlich schwierig sind, ist unter ganz unterschiedlichen Gesichtspunkten – vor allem in der Angewandten Linguistik – ausführlich untersucht worden. Dabei besteht Konsens darüber, dass “Schwierigkeit” jeweils in Bezug auf bestimmte Lese-Zielgruppen charakterisiert werden muss: Nicht-MuttersprachlerInnen empfinden andere Dinge als schwierig als es junge LeserInnen tun; der Bildungsstand ist eine andere Einflussgröße; ebenso können unterschiedliche kognitive Einschränkungen mit jeweils verschiedenen Bedürfnissen an die Einfachheit von Texten einhergehen. Schließlich sei auch auf die aktuellen

1. Vielen Dank (i) an Erik Haegert für die Unterstützung bei der Erstellung der Syntax-Parses und Berechnung der zugehörigen Werte in Abschnitt 3 sowie für die TF/IDF-Rangfolgen in Abschnitt 4, und (ii) an Peter Bourgonje für seine Bestimmung der desambiguierten Konnektoren.

Diskussionen zu *einfacher* und *leichter* Sprache verwiesen, die uns im Folgenden allerdings nicht weiter beschäftigen werden.

Stattdessen greifen wir in diesem Beitrag exemplarisch die Dimension *Altersstufe* heraus und führen eine kleine Fallstudie durch, die versuchen soll, Phänomene der Schwierigkeit dingfest zu machen. Dazu stellen wir ein kleines Arbeitskorpus aus drei verschiedenen Enzyklopädien zusammen: eine "Standard" Enzyklopädie; eine, die sich an Jugendliche wendet, und eine speziell für Kinder entwickelte. Um auch lexikalische Aspekte vergleichen zu können, ist uns daran gelegen, dass die Texte thematisch möglichst gut übereinstimmen. Dieses Ziel erreichen wir durch eine Fokussierung auf Einträge aus dem Reich der Vögel, die jeweils in allen drei Quellen zu finden sind. Die Datenmengen fallen relativ klein aus, und deshalb ist unser Ziel hier nicht eine ausgefeilte statistische Analyse, sondern nurmehr eine exemplarische Fallstudie, die Hinweise darauf gewinnen soll, ob vergleichsweise einfache Methoden Unterschiede in den Texten zeigen können.

Nachdem wir die Erstellung der Datengrundlage erläutert haben, stellen wir vergleichende Untersuchungen zur Satzkomplexität, zum Wortschatz, und zu ausgewählten Diskursphänomenen an. Danach unterwerfen wir unsere Texte noch einigen gängigen und recht einfach berechenbaren Indizes aus der Leseforschung und schließen dann mit einem Fazit.

2 Daten

Die drei Textquellen für unsere Untersuchung sind das *klexikon*², eine online Enzyklopädie für Kinder, die Brockhaus Jugendenzyklopädie und die "große" Brockhaus Enzyklopädie.³ Von diesen ist das *klexikon* die mit Abstand kleinste Quelle. Da wir an einer inhaltlich gut vergleichbaren Datenmenge interessiert sind, haben wir zunächst aus dem *klexikon* alle der Vogelwelt zuzurechnenden Artikel extrahiert; dies waren 23 verschiedene. Sie beschreiben entweder eine einzelne Art (z.B. *Pfau* oder *Kranich*) oder Gruppen (z.B. *Spechte* oder *Möwen*); hinzu kommt der generische Artikel *Vögel*. Zu allen Artikeln fanden sich Pendants in den

2. <https://klexikon.zum.de>

3. Beide sind unter <https://www.brockhaus.de/ecs/> erreichbar.

Tabelle 22.1: Syntaktische Komplexitätsindikatoren (Durchschnittswerte) für die drei Textmengen

| Textmenge | Satzlänge | Baumhöhe | NP/Satz | VP/Satz | S/Satz |
|-----------|-----------|----------|---------|---------|--------|
| K-Texte | 10.8 | 4.4 | 1.4 | 0.3 | 1.4 |
| J-Texte | 15.4 | 5.6 | 2.1 | 0.4 | 1.5 |
| E-Texte | 18.9 | 6.4 | 3.3 | 0.7 | 1.9 |

anderen beiden Enzyklopädien, diese wurden also gleichfalls extrahiert.

Zur Illustration zeigt Abb. 22.1 den Artikel *Papageien* aus dem klexikon, sowie die gekürzten Pendants aus den beiden anderen Enzyklopädien.

Als vorverarbeitende Schritte wurden Bildunterschriften entfernt, sofern sie nicht in den Text integriert waren, auch wurden einige offenkundige Tippfehler korrigiert. Im Folgenden sprechen wir der Einfachheit halber von K-Texten, J-Texten und E-Texten. Eine etwas unerfreuliche erste Beobachtung besteht darin, dass die J-Texte erheblich kürzer sind als die anderen, was die Vergleichbarkeit bestimmter Maße natürlich etwas einschränkt, hier aber nicht zu ändern ist. Insgesamt umfassen (nach Vorverarbeitung) alle K-Texte 7354 Wörter, die J-Texte 1983 Wörter und die E-Texte 7563 Wörter. Alle Texte einer Zielgruppe werden jeweils in einer einzelnen Datei zusammengefasst, und die resultierenden drei Dateien bilden die Grundlage für die nachfolgend beschriebenen Berechnungen.

3 Satzkomplexität

Syntaktische Komplexität gilt als offenkundiger Faktor für die Textverständlichkeit, doch wie man sie gut “operationalisiert”, ist weniger klar. In der Computerlinguistik sind, spätestens seit der vielzitierten Studie von Schwarm & Ostendorf (2005), diese Merkmale als Repräsentanten gebräuchlich: durchschnittl. Satzlänge, durchschnittl. Höhe des Parse-Baumes, durchschnittl. Anzahl von NPs, VPs, und von subordinierten Sätzen je Satz. Wir haben für unsere Texte mit *Stanford CoreNLP*⁴ Konstituentenstrukturen erstellt und anhand derer die genannten Maße errech-

4. <https://stanfordnlp.github.io/CoreNLP/>

Klexikon:

Papageien sind Vögel. Es gibt davon über 300 Arten, aber nicht alle von ihnen können das, wofür Papageien bekannt sind: Sie können die Stimmen von Menschen nachmachen. Papageien haben ein ziemlich großes Gehirn, darum können sie gut lernen. Zu den Papageien gehören auch die Sittiche und Kakadus.

Der Körper des Vogels steht aufrecht und ist eher schwer. Papageien mögen Körner, Nüsse und Früchte, darum ist der Schnabel kräftig und gebogen. Die Federn haben bei manchen Arten viele verschiedene Farben, während andere Arten fast einfarbig sind.

Einige Menschen halten Papageien als Haustiere. Allerdings hält man in vielen Ländern Papageien für Schädlinge, weil sie die Früchte in der Landwirtschaft wegfressen. Es werden auch Papageien gejagt und dann als Haustiere gehalten. Manche Papageienarten sind auch darum vom Aussterben bedroht.

Papageien leben normalerweise in den warmen Gegenden der Welt: in Südamerika, Afrika, Australien und im Süden von Asien. Einige Hauspapageien sind ihren Besitzern weggeflogen, so dass es heute auch Papageien in nördlichen Ländern gibt.

Brockhaus Jugendzyklopädie:

Papageien (französisch), in den Tropen und Subtropen verbreitete Vogelfamilie. Papageien sind 9-100cm lange, meist bunte Tiere, die mithilfe ihres stark gebogenen, hakigen Schnabels und ihrer Greiffüße geschickt auf Bäumen klettern. Sie leben meist gesellig und brüten vorwiegend in Baumhöhlen. In Gefangenschaft erlernen viele Papageienarten die Nachahmung von Geräuschen und können mitunter einige Wörter behalten (vor allem der afrikanische Graupapagei) und diese sogar im passenden Augenblick gebrauchen, ohne jedoch den Inhalt der Worte zu verstehen.

Am bekanntesten sind die von Australien bis zu den Philippinen verbreiteten Kakadus, mit einer aufrichtbaren Federhaube am Kopf, die bunten, mittelgroßen Loris aus Südostasien und Australien, die sehr farbenprächtigen Aras Südamerikas mit langem Schwanz, die in den Wäldern Mittel- und Südamerikas beheimateten Amazonen und die in der australischen Steppe heimischen Sittiche.

Brockhaus Enzyklopädie:

Papageien (französisch papagai, wohl von arabisch babagā'), Psittaciformes, Ordnung der Vögel mit etwa 365 Arten in 86 Gattungen.

Körperbau: Papageien sind knapp 9 bis 100 cm groß und haben einen charakteristischen, kräftigen Beißschnabel mit stark gebogenem Oberkiefer und Wachshaut an der Basis, in der auch die Nasenlöcher liegen. Die zweite und dritte Fußzehe sind nach vorn, die erste und vierte nach hinten gerichtet und damit hervorragend zum Klettern geeignet. Das Gefieder ist meist auffallend bunt.

Verbreitung: Papageien besiedeln hauptsächlich die Tropen und die Südhalbkugel; Fossilien von Papageien beziehungsweise papageienähnlichen Vögeln aus dem Miozän belegen, dass sie ursprünglich auch in Europa vorkamen. So wurde 2008 in einem dänischen Steinbruch ein 55 Mio. Jahre alter fossiler Papageienknochen gefunden, vermutlich das älteste Papageienfossil überhaupt. Dieser Fund deutet also sogar darauf, dass sich Papageien in der nördlichen Hemisphäre entwickelten, bevor sie sich im Süden und später in den Tropen verbreiteten. Heutzutage beherbergen Südamerika und die australische Region die meisten Arten.

Lebensweise: (...)

Systematik: (...)

Artenschutz: (...)

Kulturgeschichte: (...)

Abbildung 22.1: Beispieltex te aus den drei Textquellen

net – mit dem Unterschied, dass wir nicht speziell subordinierte Sätze ermitteln, sondern lediglich die durchschnittliche Zahl der ‘S’ Knoten pro Satz. Dann ergeben sich die in Tabelle 22.1 gezeigten Ergebnisse. Die erkennbaren Reihungen Kind–Jugend–Erwachsen korrespondieren weitgehend mit einschlägigen Studien zum Englischen, wonach vor allem Satzlänge und Baumhöhe zuverlässige Indikatoren des Komplexitätsgrades sind. Pitler & Nenkova (2008) und Feng u. a. (2010) haben in ihren Daten (in denen die Zielgruppen allerdings nicht per Altersstufe definiert waren) gefunden, dass die Zahl der VPs pro Satz eine prädiktive Variable ist, die der eingebetteten Sätze jedoch nicht; in unserem Fall weisen die einfachen Zählungen hingegen für sämtliche Maße dieselbe Tendenz auf (was freilich noch nicht viel über ihren statistischen Wert aussagt).

4 Wortschatz

Für einfache Wortschatz-Untersuchungen gibt es verschiedene Online-Werkzeuge, recht verbreitet sind die *Voyant Tools*⁵. U.a. zählen sie Wortfrequenzen aus und übersetzen diese auch in illustrative *word clouds*. Abb. 22.2 zeigt diejenige für die K-Texte. Insgesamt geben die Voyant Tools für die drei Textmengen die folgenden “Top-10” Inhaltswörter und Frequenzen:

- K-Texte: eier (45); leben (35); weibchen (34); arten (29); vögel (29); fliegen (26); menschen (25); nester (24); fressen (23); europa (20)
- J-Texte: vögel (13); arten (12); schnabel (9); singvögel (8); familie (7); greifvögel (7); lange (7); leben (7); insekten (6); schwanz (6)
- E-Texte: arten (68); vögel (29); schnabel (22); adler (19); lebt (18); papageien (18); eier (17); schwarz (17); afrika (16); gefieder (15)

Diese Zahlen sind für einen Vergleich allerdings nur wenig aussagekräftig (nicht nur wegen der viel kürzeren J-Texte). Es empfiehlt sich vielmehr, mit Hilfe des *tf/idf* Maßes zu bestimmen, welche Wörter besonders

5. <https://voyant-tools.org/>

typisch für ihre Textmenge (K, J, E) in Relation zur Gesamtmenge aller Texte ($K \cup J \cup E$) sind. Dann ergeben sich andere Top-10:⁶

- K-Texte: dann / woche / nennen / jungtier / weil / feind / sechs / verbringen / schweiz / mögen
- J-Texte: nord- / heimisch / klettern / mitunter / mittelgroß / ver-
sehen / rabe / jährlich / suche / ergreifen
- E-Texte: schwarz / farbschlag / aquila / symbol / regel / überwie-
gend / früh / ägypten / rom / kulturgeschichte

Ein Effekt dieses Schrittes ist beispielsweise, dass ‘Vögel’, ‘Arten’ sowie ‘leben/lebt’ in der einfachen Zählung in den Top-10 aller drei Textmengen auftauchen, in den tf/idf Top-10 folgerichtig aber nicht mehr. Wir sehen weiterhin, dass die frequenten Wörter der K-Texte, wie sie in Abb. 22.2 aufbereitet sind, gar nicht typisch für die K-Texte sind (da sie in den J- und E-Texten ebenfalls häufig sind). Stellen wir einmal die tf/idf-gewichteten K- und E-Top-10 einander gegenüber, fällt auf, dass die tf/idf Top-10 der K-Texte kürzer sind als die der E-Texte. Gilt diese Beobachtung generell für die Texte? In der Tat, die durchschnittlichen Wortlängen sind: K-Texte 5.15 (1.6 Silben) / J-Texte 5.57 (1.7 Silben) / E-Texte 5.94 (1.8 Silben). Der letzte Punkt, auf den wir hier aufmerksam machen wollen: Die K- und J-Texte haben drei bzw. vier Verben unter den tf/idf Top-10, die E-Texte keines.

Als letztes Maß betrachten wir noch die *type-token ratio*, die gemeinhin einen Eindruck von der Komplexität des Wortschatzes vermitteln soll. Sie beträgt für die K-Texte 0.26, für die J-Texte 0.48 und für die E-Texte 0.37. Bekanntlich reagiert TTR sensibel auf die Textlänge, so dass wir den J-Wert hier nicht interpretieren sollten; der beobachtete Unterschied für die fast gleich langen K- und E-Texte wäre jedoch festzuhalten.

6. Die Berechnung erfolgte mit der python Bibliothek *textblob*.



Abbildung 22.2: *Word cloud* für die Kinder-Texte (Quelle: voyant-tools.org)

5 Diskurs

In computerlinguistischen Studien wurden gelegentlich zwei Diskursphänomene in die Lesbarkeitsbestimmung einbezogen: Koreferenz und Konnektoren, letztere auch mit Kohärenzrelationen. Ein sehr einfacher Weg zur Behandlung von Koreferenz ist die Bestimmung des Anteils von Pronomen an den Wortarten; wenn wir das für unsere Texte tun, finden wir wieder eine Reihung: K-Texte 6.7 % / J-Texte 3,3 % / E-Texte 2,1 %. Allein, es fällt schwer, eine Erklärung zu formulieren, warum Pronomenverwendung in direktem Zusammenhang mit Textkomplexität stehen würde. (Im übrigen haben Barzilay & Lapata (2008) die gegenläufige Reihung gefunden, also weniger Pronomen in leicht verständlichen Texten.) Interessanter sind Versuche, komplexere Merkmale über Koreferenzketten heranzuziehen: Die Zahl der Ketten pro Text, ihre Länge, ihre Dichte. Da Koreferenzresolution aber weit mehr fehlerbehaftet ist als etwa Syntax-Analyse, ist bei automatischer Auswertung Vorsicht geboten. Ein großer Nutzen von Koreferenzmerkmalen (in Kombination mit grammatischen Rollen) für die Lesbarkeitsbestimmung war von Barzilay & Lapata (2008) auf manuell annotierten Daten gefunden worden;

demgegenüber haben Todirascu u. a. (2016) 65 verschiedene Koreferenzmerkmale mit automatischer Analyse getestet und dabei kaum messbare Effekte gefunden.

Für Konnektoren gibt es eine ermutigende Studie von Davoodi & Koseim (2016): Sie haben ein breites Spektrum von Syntax- und Diskursmerkmalen auf zwei Datensätzen zur Textlesbarkeitsbestimmung getestet, und Kombinationen aus Konnektor und Kohärenzrelation zählten zu den zuverlässigsten Merkmalen überhaupt. Für ein eigenes Experiment können wir hier auf den Konnektor-Desambiguator von Bourgonje & Stede (2018) zurückgreifen⁷ und den Anteil der Konnektoren an den Wörtern, sowie die Streuung der Konnektorenmenge (TTR) ermitteln. Dabei sollten wir aber die J-Texte wegen der geringen Textmenge außer acht lassen. Das Ergebnis ist für die K-Texte 4.2% mit TTR 0.23 und für die E-Texte 2.2% mit TTR 0.31. In den K-Texten finden sich also mehr Konnektoren, aber weniger verschiedene. Auffällig ist, dass in den K-Texten unter den Top-20 Konnektoren insgesamt 31 Instanzen von Kausalkonnektoren sind (*weil, deshalb, also, darum*), in den K-Texten hingegen Kausalität nur durch zwei Instanzen von *daher* vertreten ist. Dies dürfte aber ebenfalls kein Faktor der Textkomplexität *per se* sein, sondern ein Faktor der Abfassung von explikativen Texten bzw. der Textsorte *Lexikon*: Für Kinder wird erheblich mehr erklärt und begründet als für Erwachsene.

6 Lesbarkeitsindizes

Zum Schluss der Betrachtung von (potenziellen) Komplexitätsmerkmalen wollen wir uns mit den weit verbreiteten Lesbarkeitsindizes beschäftigen, wobei wir uns hier auf die Maße Flesch Reading Ease, Flesch/Kincaid grade level und LIX beschränken. Sie basieren auf der durchschnittlichen Satzlänge (DSL), der durchschnittlichen Wortlänge in Silben (DSW) und der Anzahl langer Wörter (ALW; mehr als 6 Buchstaben) bezogen auf die Anzahl der Wörter (AW). Hier sind die Definitionen:

7. So wie die Koreferenzresolution ist allerdings auch diese Aufgabe nicht fehlerfrei lösbar, daher ist ein wenig Vorsicht geboten.

- **Flesch Reading Ease (FRE):**

$$\text{FRE}_{\text{deutsch}} = 180 - \text{DSL} - (58,5 \times \text{DSW})$$

Je höher der Wert, desto verständlicher ist der Text; der mittlere Wert für die Zielgruppe 13-15-jähriger Schüler liegt im Bereich 60-70.⁸

- **Flesch/Kincaid grade level (FKGL):**

$$(0,39 \times \text{DSL}) + (11,8 \times \text{DSW}) - 15,59$$

Die resultierende Zahl soll die Anzahl der Schuljahre angeben, die junge Lesende absolviert haben sollen, damit sie den Text gut verstehen.⁹

- **LIX:**

$$\text{DSL} + ((\text{ALW} * 100) / \text{AW})$$

Diese heute sehr populäre Formel wurde 1968 von Björnsson vorgeschlagen. Ein Wert von 20 gilt als 'Jugendliteratur', 40 als 'Bellettristik' und 60 als 'Fachliteratur'.¹⁰

Die Berechnung erfolgt für FRE mit einem Online-Service¹¹, der die o.g. für Deutsch angepasste Formel verwendet. FKGL und LIX berechnen wir mit der Python Bibliothek *readability*, die für die drei Sprachen Englisch, Niederländisch und Deutsch nutzbar ist.¹² Eingabetexte müssen in Sätze aufgeteilt und tokenisiert sein; dafür verwenden wir den 'German Tokenizer'¹³ von Stefanie Dipper. Die Ergebnisse sind wie folgt:

- K-Texte: FRE 66 / Kincaid 7.02 / LIX 34.29
- J-Texte: FRE 56 / Kincaid 10.14 / LIX 46.54
- E-Texte: FRE 46 / Kincaid: 11.51 / LIX 50.54

8. <https://de.wikipedia.org/wiki/Lesbarkeitsindex>

9. <https://de.wikipedia.org/wiki/Lesbarkeitsindex>

10. <https://www.supertext.ch/tools/lix>

11. <https://fleschindex.de>

12. <https://pypi.org/project/readability/> Die FKGL Formel gilt eigentlich für englische Texte; es ist leider nicht klar, ob die Bibliothek eine für Deutsch angepasste Variante nutzt.

13. <https://www.linguistics.ruhr-uni-bochum.de/~dipper/resources/tokenizer.html>

Die absoluten Werte sind mit ein wenig Vorsicht zu nehmen, weil sie stark vom Vorgehen bei der Tokenisierung und der Güte der Berechnung von Silbenzahlen abhängen. Die ersichtlichen Unterschiede zwischen den drei Textmengen hingegen sollten zweifelsfrei sein.

7 Fazit

Die Vogelwelt wird Kindern, Jugendlichen und Erwachsenen unterschiedlich nahegebracht. Während für inhaltliche Auswertungen hier kein Platz blieb, haben wir für unser kleines Beispielkorpus eine Reihe quantitativer Indikatoren bestimmt, die auf unterschiedlichen sprachlichen Ebenen operieren: Lexikon, Syntax, Diskurs; wobei letzterer einstweilen für schlüssige Vorhersagen weniger ergiebig ist. In einer Bestandsaufnahme zur Textlesbarkeitsforschung aus primär computerlinguistischer Perspektive hatten Pitler & Nenkova (2008) festgestellt, dass trotz vieler Ergebnisse zu einzelnen Indikatoren wenig darüber bekannt ist, wie wichtig die einzelnen Faktoren sind und wie ihr Zusammenspiel modelliert werden sollte. Dies gilt bis heute, und ein wesentlicher Grund ist, dass (wie eingangs bemerkt) ‘Lesbarkeit’ keine textimmanente Eigenschaft ist, sondern zumindest zu einem Teil zielgruppenspezifisch beurteilt werden muss (Schriver 1989). Hinzu kommt, dass für das Deutsche – auf jeden Fall aus der Computerlinguistik – weniger Erkenntnisse vorliegen als zum Englischen. Unsere kleinen Experimente haben Hinweise darauf geliefert, dass explikative Texte, die für unterschiedliche Altersstufen geschrieben wurden, anhand der hier untersuchten Merkmale unterschieden werden können: Die ausprobierten Indikatoren von Syntax und Wortschatz haben dieselbe Reihung von K- über J- zu E-Texten angezeigt. Selbstverständlich müssten im nächsten Schritt diese initialen Hinweise nun statistisch sorgfältig untersucht werden, und ComputerlinguistInnen würden (auch für praktische Zwecke) einen automatischen Klassifikator trainieren und den *information gain* per Merkmal ausrechnen. Während all dies der “future work” vorbehalten bleibt, schließen wir mit der nicht ganz unerwarteten Frage: Wird es sich lohnen? Sind nicht die ganz einfachen, oberflächenbasierten Lesbarkeitsindizes völlig ausreichend? Oder, schlimmstenfalls, gar das Maß aller Dinge?

Literatur

- Barzilay, Regina & Mirella Lapata. 2008. Modeling local coherence: An entity-based approach. *Computational Linguistics* 34(1). 1–34.
- Bourgonje, Peter & Manfred Stede. 2018. Identifying explicit discourse connectives in German. In Kazunori Komatani, Diane Litman, Kai Yu, Alex Papangelis, Lawrence Cavendon & Mikio Nakano (Hrsg.), *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, 327–331. Melbourne, Australia: Association for Computational Linguistics.
- Davoodi, Elnaz & Leila Kosseim. 2016. On the contribution of discourse structure on text complexity assessment. In Raquel Fernandez, Wolfgang Minker, Giuseppe Carenini, Ryuichiro Higashinaka, Ron Artstein & Alesia Gainer (Hrsg.), *Proceedings of the 17th Annual SIGdial Meeting on Discourse and Dialogue*, 166–174. Los Angeles, CA: Association for Computational Linguistics.
- Feng, Lijun, Martin Jansche, Matt Huenerfauth & Noémie Elhadad. 2010. A comparison of features for automatic readability assessment. In Chu-Ren Huang & Dan Jurafsky (Hrsg.), *Proceedings of the 23rd International Conference on Computational Linguistics (COLING): Posters*, 276–284. Beijing, China: COLING 2010 Organizing Committee.
- Pitler, Emily & Ani Nenkova. 2008. Revisiting readability: a unified framework for predicting text quality. In Mirella Lapata & Hwee Tou Ng (Hrsg.), *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 186–195. Honolulu, Hawaii: Association for Computational Linguistics.
- Schriver, Karen E. 1989. Evaluating text quality: The continuum from text-focused to reader-focused methods. *IEEE Transactions on Professional Communication* 32(4). 238–255.
- Schwarm, Sarah & Mari Ostendorf. 2005. Reading level assessment using support vector machines and statistical language models. In Kevin Knight, Hwee Tou Ng & Kemal Oflazer (Hrsg.), *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, 523–530. Ann Arbor, MI: Association for Computational Linguistics.
- Todirascu, Amalia, Thomas Francois, Delphine Bernhard, Núria Gala & Anne-Laure Ligozat. 2016. Are cohesive features relevant for text

readability evaluation? In Marcello Federico & Akiko Aizawa (Hrsg.), *Proceedings of the 26th International Conference on Computational Linguistics (COLING)*, 987–997. Osaka, Japan: COLING 2016 Organizing Committee.