Analyzing Readability

Abstract:

In this seminar, we discuss the conceptual nature, linguistic modeling, and empirical evidence related to readability and explore where the formalization and automatic analysis offered by computational linguistics can support research and applications.

Instructor: Detmar Meurers
- *Office:* Room 1.28, SfS/Blochbau (Wilhelmstr. 19)
- *Email:* detmar.meurers@uni-tuebingen.de
- *Office hours:* Tuesdays 14–16 (please arrange slot by email beforehand)

Course meets: 4 SWS
- Tue 8:30 – 10:00 (SfS room 0.01)
  (first session 26.4.22)
- Thu 8:30 – 10:00 (SfS room 0.02)
- in case of online sessions: https://zoom.us/j/96719107835

Credit Points:
- Core Computatonal Linguistics Hauptseminar 6 CP, with term paper 9 CP.

Syllabus (this file):

- html-Version (http://purl.org/dm/22/ss/hs)
- pdf-Version (http://purl.org/dm/22/ss/hs/syllabus.pdf)

Moodle page: https://moodle.zdv.uni-tuebingen.de/course/view.php?id=2409

Please enroll in this course by logging into this moodle course.

Nature of course and our expectations: This is a research-oriented, hands-on Hauptseminar, in which we jointly explore the topic and gain practical experience in conducting analyses using CTAP. Everyone is expected to

1. regularly and actively participate in class, read the assigned papers and post a meaningful question on Moodle to the "Discussion Forum" on each reading *at the latest on the day before the topic is discussed* in class.

2. explore and present a topic (individually or as part of a group)

   - thoroughly research the topic, mainly based on the mentioned reference
   - prepare the presentation with slides, send them to the instructor by email to discusss them by email *a week before the presentation*
   - start a new Moodle thread on the "Discussion Forum" specifying what every course participant should read to prepare for your presentation *a week before your presentation*

- present and discuss the topic in class

3. successfully complete small projects assigned during the semester and present the results to the seminar,

4. if you pursue the 9 CP option, work out a project term paper

  - in the week before the end of the semester, select a topic and submit a one-page abstract (e.g., spelling out the analysis goal, data set, features and approach to be used).
    - Note for Computational Linguistics students: The term paper must be produced in LaTeX using the ACL conference format or the Computational Linguistics journal format; BibTeX must be used for the bibliography.

Academic conduct and misconduct: Research is driven by discussion and free exchange of ideas, motivations, and perspectives. So you are encouraged to work in groups, discuss, and exchange ideas. At the same time, the foundation of the free exchange of ideas is that everyone is open about where they obtained which information. Concretely, this means you are expected to always make explicit when you've worked on something as a team – and keep in mind that being part of a team always means sharing the work.

For text you write, you always have to provide explicit references for any ideas or passages you reuse from somewhere else. Note that this includes text "found" on the web, where you should cite the url of the web site in case no more official publication is available.

Class etiquette: Please do not read or work on materials for other classes in our seminar. All portable electronic devices such as cell phones and laptops should be switched off for the entire length of the flight, oops, class.

Topics (first sketch: this will develop as the semester proceeds)

- Current overview:
  - (Vajjala 2021)

- Traditional readability measures
  - (DuBay 2004, 2006; François & Miltsakaki 2012)

- Neural approaches:
  - (Martinc et al. 2021)
  - (Azpiazu & Pera 2019; Madrazo Azpiazu & Pera 2020a,b)
  - (Tseng et al. 2016; Liu et al. 2017)
  - BERT (Imperial 2021)
  - combination with linguistic features (Deutsch et al. 2020; Lee et al. 2021)
  - combination of features for French as L2 (Yancey et al. 2021)

2

- ranking (Lee & Vajjala 2022)
- comparison with rule-based (Chan et al. 2021)

- Various CL approaches:

  - first statistical approach (Si & Callan 2001)
  - (Heilman et al. 2007, 2008)
  - (Petersen & Ostendorf 2006a, 2007, 2006b, 2009; Schwarm & Ostendorf 2005)
  - (Pitler & Nenkova 2008; Pitler et al. 2010)
  - CohMetrix (Crossley et al. 2007)

- different target audiences

  - adults with intellectural disabilities (Feng et al. 2009)

- Psychological and Psycholinguistic Measures

  - Dependency Locality Theory (Gibson 2000; Shain et al. 2016)
  - Surprisal (Levy 2008; Boston et al. 2011; Levy & Gibson 2013)
  - Propositional Idea Density (Brown et al. 2008)
  - Cohesion/Coherence (McNamara et al. 2010)
  - Reading corpora as evidence (Demberg & Keller 2008; Demberg & Sayeed 2011; van Schijndel & Schuler 2016; Boston et al. 2008)

- Languages other than English:

  - German
    * DeLite (Vor der Brück et al. 2008a,b)
    * readability (Hancke et al. 2012; Weiss & Meurers 2018)
  - Bulgarian (Nikolova 2015)
  - French (François & Fairon 2012)
  - Greek (Georgatou 2016; Chatzipanagiotidis et al. 2021)
  - Italian (Dell'Orletta et al. 2011)
  - Russian (Reynolds 2016)
  - Swedish (Pilán et al. 2015)

- Evaluation (Huenerfauth et al. 2009; van Oosten et al. 2010; Van Oosten et al. 2011)

References

Azpiazu, I. M. & M. S. Pera (2019). Multiattentive recurrent neural network architecture for multilingual readability assessment. *Transactions of the Association for Computational Linguistics* 7, 421–436.

Boston, M. F., J. T. Hale, U. Patil, R. Kliegl & S. Vasishth (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research* 2(1), 1–12. URL http://www.jemr.org/online/2/1/1.

Boston, M. F., J. T. Hale, S. Vasishth & R. Kliegl (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes* 26(3), 301–349.

Brown, C., T. Snodgrass, S. J. Kemper, R. Herman & M. A. Covington (2008). Automatic measurement of propositional idea density from part-of-speech tagging. *Behavior Research Methods* 40(2), 540–545.

Chan, C. R., C. Pethe & S. Skiena (2021). Natural language processing versus rule-based text analysis: Comparing BERT score and readability indices to predict crowdfunding outcomes. *Journal of Business Venturing Insights* 16, e00276.

Chatzipanagiotidis, S., M. Giagkou & D. Meurers (2021). Broad linguistic complexity analysis for Greek readability classification. In *Proceedings of the 16th Workshop on Innovative Use of NLP for Building Educational Applications*. pp. 48–58.

Crossley, S., D. F. Dufty, P. M. McCarthy & D. S. Mcnamara (2007). Toward a New Readability: A Mixed Model Approach. In *Proceedings of the 29th annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, pp. 197–202.

Dell'Orletta, F., S. Montemagni & G. Venturi (2011). READ-IT: Assessing Readability of Italian Texts with a View to Text Simplification. In *Proceedings of the 2nd Workshop on Speech and Language Processing for Assistive Technologies*. pp. 73–83.

Demberg, V. & F. Keller (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition* 109(2), 193 – 210.

Demberg, V. & A. Sayeed (2011). Linguistic cognitive load: implications for automotive UIs. In *Adjunct Proceedings of AutomotiveUI'11*.

Deutsch, T., M. Jasbi & S. M. Shieber (2020). Linguistic Features for Readability Assessment. In *Proceedings of the Fifteenth Workshop on Innovative Use of NLP for Building Educational Applications*. pp. 1–17. URL https://aclanthology.org/2020.bea-1.1.pdf.

DuBay, W. H. (2004). *The Principles of Readability*. Costa Mesa, California: Impact Information. URL http://www.impact-information.com/impactinfo/readability02.pdf.

DuBay, W. H. (2006). *The Classic Readability Studies*. Costa Mesa, California: Impact Information.

Feng, L., N. Elhadad & M. Huenerfauth (2009). Cognitively Motivated Features for Readability Assessment. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*. Athens, Greece: Association for Computational Linguistics, pp. 229–237. http://aclweb.org/anthology/E09-1027.

François, T. & C. Fairon (2012). An "AI readability" formula for French as a foreign language. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. https://www.aclweb.org/anthology/D12-1043.

François, T. & E. Miltsakaki (2012). Do NLP and machine learning improve traditional readability formulas? In *Proceedings of the First Workshop on Predicting and Improving Text Readability for target reader populations*. Association for Computational Linguistics, pp. 49–57.

Georgatou, S. (2016). Approaching readability features in Greek school books. Master thesis in computational linguistics, Department of Linguistics, University of Tübingen.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita & W. O'Neil (eds.), *Image, language, brain: papers from the First Mind Articulation Project Symposium*, MIT, pp. 95–126.

Hancke, J., S. Vajjala & D. Meurers (2012). Readability Classification for German using lexical, syntactic, and morphological features. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING)*. Mumbai, India, pp. 1063–1080. http://aclweb.org/anthology-new/C/C12/C12-1065.pdf.

Heilman, M., K. Collins-Thompson, J. Callan & M. Eskenazi (2007). Combining Lexical and Grammatical Features to Improve Readability Measures for First and Second Language Texts. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL-07)*. Rochester, New York, pp. 460–467.

Heilman, M., K. Collins-Thompson & M. Eskenazi (2008). An Analysis of Statistical Models and Features for Reading Difficulty Prediction. In *Proceedings of the 3rd Workshop on Innovative Use of NLP for Building Educational Applications at ACL-08*. Columbus, Ohio.

Huenerfauth, M., L. Feng & N. Elhadad (2009). Comparing evaluation techniques for text readability software for adults with intellectual disabilities. In *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. New York, NY, USA: ACM, Assets '09, pp. 3–10. http://doi.acm.org/10.1145/1639642.1639646.

Imperial, J. M. (2021). BERT Embeddings for Automatic Readability Assessment. *arXiv preprint arXiv:2106.07935* .

Lee, B. W., Y. S. Jang & J. H.-J. Lee (2021). Pushing on Text Readability Assessment: A Transformer Meets Handcrafted Linguistic Features. *arXiv preprint arXiv:2109.12258* .

Lee, J. & S. Vajjala (2022). A Neural Pairwise Ranking Model for Readability Assessment. *arXiv preprint arXiv:2203.07450* .

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106 (3), 1126–1177.

Levy, R. & E. Gibson (2013). Surprisal, the PDC, and the primary locus of processing difficulty in relative clauses. *Frontiers in Language Sciences* 4:229, 1–3.

Liu, H., S. Li, J. Zhao, Z. Bao & X. Bai (2017). Chinese teaching material readability assessment with contextual information. In *2017 International Conference on Asian Language Processing (IALP)*. IEEE, pp. 66–69.

Madrazo Azpiazu, I. & M. S. Pera (2020a). An Analysis of Transfer Learning Methods for Multilingual Readability Assessment. *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* pp. 95–100.

Madrazo Azpiazu, I. & M. S. Pera (2020b). Is cross-lingual readability assessment possible? *Journal of the Association for Information Science and Technology* .

Martinc, M., S. Pollak & M. Robnik-Šikonja (2021). Supervised and unsupervised neural approaches to text readability. *Computational Linguistics* 47(1), 141–179.

McNamara, D. S., M. M. Louwerse, P. M. McCarthy & A. C. Graesser (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes* 47(4), 292–330. URL https://umdrive.memphis.edu/pmmccrth/public/Papers/DP_Cohesion%20%28McNamara%29.doc.

Nikolova, L. (2015). Readability Classification for Bulgarian. Master thesis in computational linguistics, Department of Linguistics, University of Tübingen.

Petersen, S. E. & M. Ostendorf (2006a). Assessing the Reading Level of Web Pages. In *Ninth International Conference on Spoken Language Processing (Interspeech-ICSLP)*. Pittsburgh, Pennsylvania. URL http://sarahpetersen.net/portfolio/petersen_is2006.pdf.

Petersen, S. E. & M. Ostendorf (2006b). Assessing the Reading Level of Web Pages (Poster). In *Proceedings of Interspeech 2006*.

Petersen, S. E. & M. Ostendorf (2007). Text Simplification for Language Learners: A Corpus Analysis. In *Speech and Language Technology for Education (SLaTE)*. pp. 69–72. URL http://isca-speech.org/archive_open/archive_papers/slate_2007/sle7_069.pdf.

Petersen, S. E. & M. Ostendorf (2009). A machine learning approach to reading level assessment. *Computer Speech and Language* 23, 86–106.

Pilán, I., S. Vajjala & E. Volodina (2015). A Readable Read: Automatic Assessment of Language Learning Materials based on Linguistic Complexity. In *Proceedings of CICLING 2015- Research in Computing Science Journal Issue (to appear)*. https://arxiv.org/abs/1603.08868.

Pitler, E., A. Louis & A. Nenkova (2010). Automatic evaluation of linguistic quality in multi-document summarization. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Stroudsburg, PA, USA: Association for Computational Linguistics, ACL '10, pp. 544–554. URL http://dl.acm.org/citation.cfm?id=1858681.1858737.

Pitler, E. & A. Nenkova (2008). Revisiting readability: a unified framework for predicting text quality. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Stroudsburg, PA, USA: Association for Computational Linguistics, EMNLP '08, pp. 186–195. URL http://dl.acm.org/citation.cfm?id=1613715.1613742.

Reynolds, R. (2016). Russian natural language processing for computer-assisted language learning: capturing the benefits of deep morphological analysis in real-life applications. Ph.D. thesis, UiT - The Arctic University of Norway. URL https://munin.uit.no/handle/10037/9685.

Schwarm, S. & M. Ostendorf (2005). Reading Level Assessment Using Support Vector Machines and Statistical Language Models. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL-05)*. Ann Arbor, Michigan, pp. 523–530.

Shain, C., M. van Schijndel, R. Futrell, E. Gibson & W. Schuler (2016). Memory access during incremental sentence processing causes reading time latency. In *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC)*. Osaka, pp. 49–58. URL https://aclweb.org/anthology/W16-4106.

Si, L. & J. Callan (2001). A Statistical Model for Scientific Readability. In *Proceedings of the 10th International Conference on Information and Knowledge Management (CIKM)*. ACM, pp. 574–576.

Tseng, H. C., H. T. Hung, Y. T. Sung & B. Chen (2016). Classification of text readability based on deep neural network and representation learning techniques. In *28th Conference on Computational Linguistics and Speech Processing, ROCLING 2016*. The Association for Computational Linguistics and Chinese Language . . . , pp. 255–270.

Vajjala, S. (2021). Trends, limitations and open challenges in automatic readability assessment research. *arXiv preprint arXiv:2105.00973* .

Van Oosten, P., V. Hoste & D. Tanghe (2011). A posteriori agreement as a quality measure for readability prediction systems. In *Proceedings of the 12th international conference on Computational linguistics and intelligent text processing - Volume Part II*. Berlin,

Heidelberg: Springer-Verlag, CICLing'11, pp. 424–435. URL http://dl.acm.org/citation.cfm?id=1964750.1964790.

van Oosten, P., D. Tanghe & V. Hoste (2010). Towards an Improved Methodology for Automated Readability Prediction. In *LREC'10*. pp. −1–1. URL http://www.lrec-conf.org/proceedings/lrec2010/pdf/286_Paper.pdf.

van Schijndel, M. & W. Schuler (2016). Addressing surprisal deficiencies in reading time models. In *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CL4LC) at COLING*. Osaka.

Vor der Brück, T., S. Hartrumpf & H. Helbig (2008a). A Readability Checker with Supervised Learning using Deep Syntactic and Semantic Indicators. *Informatica* 32(4), 429–435.

Vor der Brück, T., H. Helbig & J. Leveling (2008b). *The readability checker DeLite*. Tech. Rep. Technical Report 345-5/2008, Fakultät für Mathematik und Informatik, FernUniversität in Hagen.

Weiss, Z. & D. Meurers (2018). Modeling the Readability of German Targeting Adults and Children: An Empirically Broad Analysis and its Cross-Corpus Validation. In *Proceedings of the 27th International Conference on Computational Linguistics (COLING)*. Santa Fe, New Mexico, USA. https://www.aclweb.org/anthology/C18-1026.

Yancey, K., A. Pintard & T. Francois (2021). Investigating readability of French as a foreign language with deep learning and cognitive and pedagogical features. *Lingue e linguaggio* 20(2), 229–258.