

Please *ReaderBench* This Text: A Multi-Dimensional Textual Complexity Assessment Framework

Mihai Dascalu, Scott A. Crossley, Danielle S. McNamara, Philippe Dessus, and Stefan Trausan-Matu

Corresponding Author: mihai.dascalu@cs.pub.ro

Abstract

A critical task for tutors is to provide learners with suitable reading materials in terms of difficulty. The challenge of this endeavor is increased by students' individual variability and the multiple levels in which complexity can vary, thus arguing for the necessity of automated systems to support teachers. This chapter describes *ReaderBench*, an open-source multi-dimensional and multi-lingual system that uses advanced Natural Language Processing techniques to assess textual complexity at multiple levels including surface-based, syntax, semantics and discourse structure. In contrast to other existing approaches, *ReaderBench* is centered on cohesion and makes extensive usage of two complementary models, i.e., Cohesion Network Analysis and the polyphonic model inspired from dialogism. The first model provides an in-depth view of discourse in terms of cohesive links, whereas the second one highlights interactions between points of view spanning throughout the discourse. In order to argue for its wide applicability and extensibility, two studies are introduced. The first study investigates the degree to which *ReaderBench* textual complexity indices differentiate between high and low cohesion texts. The *ReaderBench* indices led to a higher classification accuracy than those included in prior studies using *Coh-Matrix* and *TAACO*. In the second study, *ReaderBench* indices are used to predict the difficulty of a set of various texts. Although the high number of predictive indices (50 plus) accounted for less variance than previous studies, they make valuable contributions to our understanding of text due to their wide coverage.

Keywords: comprehension modeling, learning analytics, automated essay scoring, data analytics, Natural Language Processing

[Back to Table of Contents](#)

[Link to Book](#)

APA citation information

Dascalu, M., Crossley, S. A., McNamara, D. S., Dessus, P., & Trausan-Matu, S. (2018). Please ReaderBench this text: A multi-dimensional textual complexity assessment framework. In S. D. Craig (Ed.). *Tutoring and Intelligent Tutoring Systems* (pp. 251-272). New York, NY: Nova Science Publishers.

References

Allen, L. K., Dascalu, M., McNamara, D. S., Crossley, S., & Trausan-Matu, S. (2016). Modeling Individual Differences among Writers Using ReaderBench. In *8th Int. Conf. on Education and New Learning Technologies (EduLearn16)* (pp. 5269–5279). Barcelona, Spain: IATED.

Balint, M., Dascalu, M., & Trausan-Matu, S. (2016a). Classifying Written Texts through Rhythmic Features. In *15th Int. Conf. on Artificial Intelligence: Methodology, Systems, and Applications (AIMSA 2016)* (pp. 121–129). Varna, Bulgaria: Springer.

Balint, M., Dascalu, M., & Trausan-Matu, S. (2016b). The Rhetorical Nature of Rhythm. In *15th Int. Conf. on Networking in Education and Research (RoEduNet)* (pp. 48–53). Bucharest, Romania: IEEE.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*(1), 73–79.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*(4-5), 993–1022.

Botarleanu, R. M., Dascalu, M., Sirbu, M. D., Crossley, S. A., & Trausan-Matu, S. (2018). README – Generating Personalized Feedback for Essay Writing using the ReaderBench Framework. In *3rd Int. Conf. on Smart Learning Ecosystems and Regional Development (SLERD 2018)* (pp. 133–145). Aalborg, Denmark: Springer.

Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, *39*, 324–345.

Cambria, E., Grassi, M., Poria, S., & Hussain, A. (2013). Sentic computing for social media analysis, representation, and retrieval. In Ramzan, N., Zwol, R., Lee, J. S., Clüver, K. & Hua, X. S. (Eds.), *Social Media Retrieval* (pp. 191–215). New York, NY: Springer.

Chall, J. S., & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Northampton, MA: Brookline Books.

Crossley, S. A., Dascalu, M., Trausan-Matu, S., Allen, L., & McNamara, D. S. (2016). Document Cohesion Flow: Striving towards Coherence. In *38th Annual Meeting of the Cognitive Science Society* (pp. 764–769). Philadelphia, PA: Cognitive Science Society.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, 48(4), 1227–1237.

Crossley, S. A., Paquette, L., Dascalu, M., McNamara, D. S., & Baker, R. S. (2016). Combining Click-Stream Data with NLP Tools to Better Understand MOOC Completion. In *6th Int. Conf. on Learning Analytics & Knowledge (LAK '16)* (pp. 6–14). Edingurgh, UK: ACM.

Crossley, S. A., Roscoe, R. D., McNamara, D. S., & Graesser, A. (2011). Predicting human scores of essay quality using computational indices of linguistic and textual features. In Biswas, G., Bull, S., Kay, J., & Mitrovic, A. (Eds.), *15th Int. Conf. on Artificial Intelligence in Education* (pp. 438–440). Christchurch, New Zealand: Springer.

Crossley, S. A., Skalicky, S. C., Dascalu, M., Kyle, K., & McNamara, D. S. (2017). Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes*, 54(5-6), 340–359.

Dascalu, M. (2014). *Analyzing discourse and text complexity for learning and collaborating*. Cham, Switzerland: Springer.

Dascalu, M., Allen, K. A., McNamara, D. S., Trausan-Matu, S., & Crossley, S. A. (2017). Modeling Comprehension Processes via Automated Analyses of Dialogism. In *39th Annual Meeting of the Cognitive Science Society (CogSci 2017)* (pp. 1884–1889). London, UK: Cognitive Science Society.

Dascalu, M., Dessus, P., Bianco, M., & Trausan-Matu, S. (2014). Are Automatically Identified Reading Strategies Reliable Predictors of Comprehension? In S. Trausan-

Matu, K. E. Boyer, M. Crosby & K. Panourgia (Eds.), *12th Int. Conf. on Intelligent Tutoring Systems (ITS 2014)* (pp. 456–465). Honolulu, USA: Springer.

Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., & Nardy, A. (2014). Mining texts, learner productions and strategies with ReaderBench. In Peña-Ayala, A. (Ed.), *Educational Data Mining: Applications and Trends* (pp. 345–377). Cham, Switzerland: Springer.

Dascalu, M., Gifu, D., & Trausan-Matu, S. (2016). What Makes your Writing Style Unique? Significant Differences between Two Famous Romanian Orators. In Nguyen, N. T., Manolopoulos, Y., Iliadis, L., & Trawinski, B. (Eds.), *8th Int. Conf. on Computational Collective Intelligence (ICCCI 2016)* (pp. 143–152). Halkidiki, Greece: Springer.

Dascalu, M., Gutu, G., Ruseti, S., Paraschiv, I. C., Dessus, P., McNamara, D. S., Crossley, S., & Trausan-Matu, S. (2017). ReaderBench: A Multi-Lingual Framework for Analyzing Text Complexity. In Lavoué, E., Drachler, H., Verbert, K., Broisin, J., & Pérez-Sanagustín, M. (Eds.), *12th European Conference on Technology Enhanced Learning (EC-TEL 2017)* (pp. 495–499). Tallinn, Estonia: Springer.

Dascalu, M., McNamara, D. S., Crossley, S. A., & Trausan-Matu, S. (2015). Age of Exposure: A Model of Word Learning. In *30th AAAI Conference on Artificial Intelligence* (pp. 2928–2934). Phoenix, AZ: AAAI Press.

Dascalu, M., McNamara, D. S., Trausan-Matu, S., & Allen, L.K. (2018). Cohesion Network Analysis of CSCL Participation. *Behavior Research Methods*, *50*(2), 604–619. doi: 10.3758/s13428-017-0888-4

Dascalu, M., Popescu, E., Becheru, A., Crossley, S. A., & Trausan-Matu, S. (2016). Predicting Academic Performance Based on Students' Blog and Microblog Posts. In *11th European Conference on Technology Enhanced Learning (EC-TEL 2016)* (pp. 370–376). Lyon, France: Springer.

Dascalu, M., Stavarache, L. L., Trausan-Matu, S., Dessus, P., & Bianco, M. (2014). Reflecting Comprehension through French Textual Complexity Factors. In *26th Int. Conf. on Tools with Artificial Intelligence (ICTAI 2014)* (pp. 615–619). Limassol, Cyprus: IEEE.

Dascalu, M., Trausan-Matu, S., McNamara, D. S., & Dessus, P. (2015). ReaderBench – Automated Evaluation of Collaboration based on Cohesion and Dialogism. *International Journal of Computer-Supported Collaborative Learning*, *10*(4), 395–423. doi: 10.1007/s11412-015-9226-y.

- Dascalu, M., Westera, W., Ruseti, S., Trausan-Matu, S., & Kurvers, H. (2017). ReaderBench Learns Dutch: Building a Comprehensive Automated Essay Scoring System for Dutch. In Baker, A. E. R., Hu, X., Rodrigo, M. M. T. & du Boulay, B., (Eds.), *18th Int. Conf. on Artificial Intelligence in Education (AIED 2017)* (pp. 52–63). Wuhan, China: Springer.
- Gervasi, V., & Ambriola, V. (2002). Quantitative assessment of textual complexity. In M. L. Barbaresi (Ed.), *Complexity in language and text* (pp. 197–228). Pisa, Italy: Plus.
- Gifu, D., Dascalu, M., Trausan-Matu, S., & Allen, L. K. (2016). Time Evolution of Writing Styles in Romanian Language. In *28th Int. Conf. on Tools with Artificial Intelligence (ICTAI 2016)* (pp. 1048–1054). San Jose, CA: IEEE.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers*, *36*(2), 193–202.
- Grömping, Ulrike. (2006). Relative importance for linear regression in R: the package relaimpo. *Journal of statistical software*, *17*(1), 1–27.
- Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, *21*(2), 203–225.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354.
- Kate, R. J., Luo, X., Patwardhan, S., Franz, M., Florian, R., Mooney, R. J., Roukos, S., & Welty, C. (2010). Learning to predict readability using diverse linguistic features. In *23rd Int. Conf. on Computational Linguistics* (pp. 546–554): Association for Computational Linguistics.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge, MA: Cambridge University Press.
- Kuhn, M. (2008). Caret package. *Journal of statistical software*, *28*(5), 1–26.
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.
- Kyle, K., & Crossley, S. A. (2018). Measuring Syntactic Complexity in L2 Writing Using Fine-Grained Clausal and Phrasal Indices. *Modern Language Journal*, *102* (2), 333-349.

Kyle, K., Crossley, S. A., & Berger, C. (in press). The Tool for the Automatic Analysis of Lexical Sophistication Version 2.0. *Behavior Research Methods*.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*(2), 211–240.

Landauer, T. K., Kireyev, K., & Panaccione, C. (2011). Word maturity: A new metric for word knowledge. *Scientific Studies of Reading*, *15*(1), 92–108.

Lasswell, H. D., & Namenvirth, J. Z. (1969). *The Lasswell Value Dictionary*. New Haven: Yale University Press.

Lee, H., Peirsman, Y., Chang, A., Chambers, N., Surdeanu, M., & Jurafsky, D. (2011). Stanford's Multi-Pass Sieve Coreference Resolution System at the CoNLL-2011 Shared Task. In *Fifteenth Conference on Computational Natural Language Learning: Shared (TaskCONLL Shared Task '11)* (pp. 28–34). Portland, OR: ACL.

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 55–60). Baltimore, MA: ACL.

Marcus, S. (1970). *Poetica matematică*. Bucharest, Romania: Editura Acad. Rep. Soc. Romania.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). The linguistic features of quality writing. *Written Communication*, *27*(1), 57–86.

McNamara, D. S., Graesser, A. C., & Louwerson, M. M. (2012). Sources of text difficulty: Across the ages and genres. In Sabatini, J. P., Albro, E. & O'Reilly, T. (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 89–116). Lanham, MD: R&L Education.

McNamara, D. S., Louwerson, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Matrix: Capturing linguistic features of cohesion. *Discourse Processes*, *47*(4), 292–330.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representation in Vector Space. In *Workshop at ICLR*. Scottsdale, AZ.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3), 436–465.

Newbold, N., & Gillam, L. (2010). The linguistics of readability: the next step for word processing. In *NAACL HLT 2010 Workshop on Computational Linguistics and Writing: Writing Processes and Authoring Aids* (pp. 65–72). Los Angeles, CA: Association for Computational Linguistics.

Page, E. (1966). The imminence of grading essays by computer. *Phi Delta Kappan*, 47, 238–243.

Page, E. (1968). Analyzing student essays by computer. *International Review of Education*, 14(2), 210–225.

Pennebaker, James W, Booth, Roger J, & Francis, Martha E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. Austin, TX: University of Texas.

Powers, D. E., Burstein, J., Chodorow, M., Fowles, M. E., & Kukich, K. (2001). Stumping e-rater®: Challenging the validity of automated essay scoring. Princeton, NJ: Educational Testing Service.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: Foundation for Statistical Computing.

Roscoe, R. D., Varner, L. K., Weston, J. L., Crossley, S. A., & McNamara, D. S. (2014). The Writing Pal intelligent tutoring system: Usability testing and development. *Computers and Composition*, 34, 39–59.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social science information*, 44(4), 695–729.

Schock, J., Cortese, M. J., Khanna, M. M., & Toppi, S. (2012). Age of acquisition estimates for 3,000 disyllabic words. *Behavior Research Methods*, 44(4), 971–977.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 379–423 & 623–656.

Shannon, C. E. (1951). Prediction and entropy of printed English. *The Bell System Technical Journal*, 30, 50–64.

Stone, P., Dunphy, D. C., Smith, M. S., Ogilvie, D. M., & associates. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: The MIT Press.

Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). PolyCAFe—automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning*, 9(2), 127–156. doi: 10.1007/s11412-014-9190-y.

Trausan-Matu, S., Stahl, G., & Sarmiento, J. (2007). Supporting polyphonic collaborative learning. *E-service Journal*, 6(1), 58–74.

Wresch, W. (1993). The imminence of grading essays by computer—25 years later. *Computers and Composition*, 10(2), 45–58.

Wu, Z., & Palmer, M. (1994). Verb semantics and lexical selection. In *32nd Annual Meeting of the Association for Computational Linguistics, ACL '94* (pp. 133–138). New Mexico, USA: ACL.