ISCL Hauptseminar (Winter semester 2019)

Computational Linguistic Analysis of Linguistic Complexity for Readability and Proficiency Assessment

Abstract:

Notions of complexity surface in a number of different contexts: In theoretical linguistics, syntactic structures are analyzed in terms of their complexity and constraints such as the complex-NP constraint are formulated on this basis. In cognitive psychology, the complexity involved in cognitively processing language input in human sentence processing is studied. In second language acquisition research, the analysis of complexity (together with accuracy and fluency) is used to gain insights into the process and product of acquisition. In language testing and learner corpus research, the linguistic complexity of learner language is related to proficiency levels. For readability research, the linguistic complexity is used to determine who a given text is readable for.

In this seminar, we will discuss the empirical and conceptual nature of these notions of complexity and explore where the formalization and automatic analysis offered by computational linguistics can lead to applications such as automatic readability measures, proficiency classification, and search engines supporting the filtering of results by complexity for particular target audiences.

Instructors:

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

- Dr. Xiaobin Chen (HIB - Empirical Educational Science)
  - *Email:* xiaobin.chen@uni-tuebingen.de
  - *Office hours:* Fridays 14–16 (please arrange slot by email beforehand) in Room 1.29, Blochbau (Wilhelmstr. 19)

- Zarah Weiss
  - *Office:* Room 1.29, Blochbau (Wilhelmstr. 19)
  - *Email:* zarah-leonie.weiss@uni-tuebingen.de
  - *Office hours:* Wednesdays 12–14 (please arrange slot by email beforehand)

Course meets: 4 SWS

- Wed/Fr 10 – 12, SfS 1.13

Credit Points:

- Core Computatonal Linguistics Hauptseminar 6 CP, with project term paper 9 CP.

Syllabus (this file):

- html-Version (http://purl.org/dm/19/ws/hs)
- pdf-Version (http://purl.org/dm/19/ws/hs/syllabus.pdf)

Moodle page: https://moodle.zdv.uni-tuebingen.de/course/view.php?id=543

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 implementing analyses. Substantial programming experience (at least at the level of the second Data Structures and Algorithms course) is required; permission may be granted for teams of two people combining complementary expertise. Everyone is expected to

1. successfully complete the regular exercises and small projects assigned during the semester and present the results to the seminar,
2. 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.
3. 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 instructors by email and discuss them with me in a half hour slot during my office hour *at least 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
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 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.

Sketch of assignments

1. Traditional readability: Flesch-Kincaid formula
    - processing required: tokenization, sentence segmentation, syllable identification
    - input: corpus (one text per file), output: csv file with grades
2. Lexical complexity: lexical richness, frequency
    - processing required: tokenization, use of frequency data, predictions using supervised machine learning
3. Syntactic complexity
4. Psycholinguistic perspective (Ted Gibson's DLT)
5. Discourse: aspects of cohesion (connectives, overlap, coreference, ...)

Data sources to be used include:

- EFCamDat (English L2): course levels, mostly cross-sectional
- MERLIN (German L2): proficiency classes
- LitKey (German L1): longitudinal within-student design

Weeks

1. 23./25.10. Detmar: Introduction
2. 30.10. Zarah: Readability Formulas *(no class on 1.11. holiday)*
3. 6./8.11. Detmar: Introduction (cont.)
    - Deadline for second homework assignment (6.11.)
4. 13.11. Xiaobin: Aggregating lexical-level complexity information to predict the text-level complexity (Chen & Meurers 2017)
5. 15.11. Xiaobin: Linking text readability and learner proficiency using linguistic complexity feature vector distance (Chen & Meurers 2019)
    - Deadline for second homework assignment (17.11.)
6. 20.11 Tanja Heck: Lexical complexiy (Laufer & Nation 1995)
7. 22.11. Eva Huber: Morphological complexity (Paquot 2019)
8. 27.11. Xiaobin Chen, Zarah Weiss: Longitudinal Development of Complexity and Accuracy
9. 29.11. Haemanth Santhi-Ponnusamy: Syntactic complexity in college-level English writing and L1 differences (Lu & Ai 2015)

10. 4.12. Elizabeth Bear: The Development of Second Language Writing Complexity in Groups and Individuals: A Longitudinal Learner Corpus Study (Vyatkina 2012)
11. 6.12. Zarah Weiss: Analyzing linguistic complexity and accuracy in academic language development (Weiss & Meurers 2019a)
12. 11.12 Hebah Ahmed: Automatic Measurement of Syntactic Complexity Using the Revised Developmental Level Scale (Lu 2009; Voss 2005)
13. 13.12. Daniela Rossman: IPSyn (Sagae et al. 2005; Lubetich & Sagae 2014)
14. 18.12. Jana Murasová: Comparing child L2 development with adult L2 development (Unsworth 2008)
15. 20.12. Sarah Neuhaus: Readability assessment for aphasia: (Aleligay et al. 2008; Abou-Diab et al. 2019) (and Discussion of Christmas break project)
16. 8.1 Zarah Weiss: Dependency Locality Theory (Gibson 2000; Shain et al. 2016)
17. 10.1 Mareile Winkler: Propositional Idea Density (Brown et al. 2008)
18. 15.1. Masoumeh Moradipour-tari: Discourse/Cohesion (Graesser et al. 2004)
19. 17.1 Xiaobin Chen: CTAP (**?**) and its UIMA architecture
20. 22.1. Nelly Sagirov: Testing target text fluency
21. 24.1. Denise Loefflad: Redability analysis for French as a foreign language (François & Fairon 2012)
22. 29.1. Mohamed Ouji: Evaluation (Huenerfauth et al. 2009; van Oosten et al. 2010; Van Oosten et al. 2011)
23. 31.1
24. 5.2.

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

- Traditional readability measures
  - (DuBay 2004, 2006; François & Miltsakaki 2012)
- Complexity in Second Language Acquisition research (background: Skehan 1989; Wolfe-Quintero et al. 1998; Ortega 2003; Housen & Kuiken 2009)
  1. Lexical complexity (Laufer & Nation 1995) (more: Malvern et al. 2004; Read & Nation 2004; McCarthy & Jarvis 2010; Lu 2012; Kyle & Crossley 2015)
  2. Phraseological complexity (Paquot 2019)
  3. Distinguishing clausal from phrasal complexity (Kyle 2016)
  - Morphological complexity (Brezina & Pallotti 2019)

  4./5. Syntactic complexity (Cheung & Kemper 1992; Covington et al. 2006; Lu 2010)

6. Discourse/Cohesion (McNamara et al. 2002; Graesser et al. 2004; Crossley et al. 2007, 2008)

- Psycholinguistic/Psychological Measures

7. Dependency Locality Theory (Gibson 2000; Shain et al. 2016) (Zarah?)
8. Surprisal (Levy 2008; Boston et al. 2011; Levy & Gibson 2013)
9. Propositional Idea Density (Brown et al. 2008)

  - Reading corpora as evidence (Demberg & Keller 2008; Demberg & Sayeed 2011; van Schijndel & Schuler 2016; Boston et al. 2008)

- Child Language Development

10. Revised D-Level (Lu 2009; Voss 2005)
11. IPSyn (Sagae et al. 2005; Lubetich & Sagae 2014)

  - Analyzing linguistic complexity and accuracy in academic language development (Weiss & Meurers 2019a) (Zarah)

- Writing development (Lu 2011; Ai & Lu 2013; Lu & Ai 2015; Crossley & McNamara 2014)
- Tutoring system based on dependent clause extraction and exercise generation (Xiaobin)
- Languages other than English:
  - German
    * DeLite (Vor der Brück et al. 2008a,b)
    * readability (Hancke et al. 2012; Weiss & Meurers 2018)
    * proficiency (Hancke 2013; Hancke & Meurers 2013; Weiss 2017; Weiss & Meurers 2019b)
    * academic language development (Weiss & Meurers 2019a)
    * human grading (Weiss et al. 2019)
  - Bulgarian (Nikolova 2015)
  - French (François & Fairon 2012)
  - Greek (Georgatou 2016)
  - Italian (Dell'Orletta et al. 2011)
  - Russian (Reynolds 2016)
  - Swedish (Pilán et al. 2015)
- Longitudional analysis (Vyatkina 2012)
- Task effects (Alexopoulou et al. 2017) (Detmar?)
- Evaluation (Huenerfauth et al. 2009; van Oosten et al. 2010; Van Oosten et al. 2011)

References

Abou-Diab, S. N., D. C. Moser & S. R. Atcherson (2019). Evaluation of the readability, validity, and user-friendliness of written web-based patient education materials for aphasia. *Aphasiology* 33(2), 187–199.

Ai, H. & X. Lu (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier & P. Thompson (eds.), *Automatic Treatment and Analysis of Learner Corpus Data*, John Benjamins, pp. 249–264.

Aleligay, A., L. E. Worrall & T. A. Rose (2008). Readability of written health information provided to people with aphasia. *Aphasiology* 22(4), 383–407.

Alexopoulou, T., M. Michel, A. Murakami & D. Meurers (2017). Task Effects on Linguistic Complexity and Accuracy: A Large-Scale Learner Corpus Analysis Employing Natural Language Processing Techniques. *Language Learning* 67, 181–209. URL https://doi.org/10.1111/lang.12232.

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.

Brezina, V. & G. Pallotti (2019). Morphological complexity in written L2 texts. *Second language research* 35(1), 99–119.

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.

Chen, X. & D. Meurers (2017). Word frequency and readability: Predicting the text-level readability with a lexical-level attribute. *Journal of Research in Reading* 41(3), 486–510.

Chen, X. & D. Meurers (2019). Linking text readability and learner proficiency using linguistic complexity feature vector distance. *Computer-Assisted Language Learning* https://doi.org/10.1080/09588221.2018.1527358.

Cheung, H. & S. Kemper (1992). Competing complexity metrics and adults' production of complex sentences. *Applied Psycholinguistics* 13(01), 53–76. URL http://dx.doi.org/10.1017/S0142716400005427.

Covington, M. A., C. He, C. Brown, L. Naçi & J. Brown (2006). *How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level Scale*. Computer Analysis of Speech for Psychological Research (CASPR) Research Report 2006-01, The University of Georgia, Artificial Intelligence Center, Athens, GA. URL http://www.ai.uga.edu/caspr/2006-01-Covington.pdf.

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.

Crossley, S. A., J. Greenfield & D. S. McNamara (2008). *Assessing text readability using cognitively based indices*, Teachers of English to Speakers of Other Languages, Inc. 700 South Washington Street Suite 200, Alexandria, VA 22314, pp. 475–493.

Crossley, S. A. & D. S. McNamara (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing* 26, 66–79.

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*.

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.

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.

Graesser, A. C., D. S. McNamara, M. M. Louweerse & Z. Cai (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments and Computers* 36, 193–202. URL http://home.autotutor.org/graesser/publications/bsc505.pdf.

Hancke, J. (2013). Automatic Prediction of CEFR Proficiency Levels Based on Linguistic Features of Learner Language. Master's thesis, International Studies in Computational Linguistics. Seminar für Sprachwissenschaft, Universität Tübingen.

Hancke, J. & D. Meurers (2013). Exploring CEFR classification for German based on rich linguistic modeling. In *Learner Corpus Research 2013, Book of Abstracts*. Bergen, Norway. URL http://purl.org/dm/papers/Hancke.Meurers-13.html.

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)*. Mumbay, India, pp. 1063–1080. http://aclweb.org/anthology-new/C/C12/C12-1065.pdf.

Housen, A. & F. Kuiken (2009). Complexity, Accuracy and Fluency in Second Language Acquisition. *Applied Linguistics* 30(4), 461–473.

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.

Kyle, K. (2016). Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of Syntactic Sophistication. Ph.D. thesis, Georgia State University. URL http://scholarworks.gsu.edu/alesl_diss/35.

Kyle, K. & S. A. Crossley (2015). Automatically Assessing Lexical Sophistication: Indices, Tools, Findings, and Application. *TESOL Quarterly* 49(4), 757–786.

Laufer, B. & P. Nation (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics* 16(3), 307–322. URL http://applij.oxfordjournals.org/content/16/3/307.abstract.

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.

Lu, X. (2009). Automatic measurement of syntactic complexity in child language acquisition. *International Journal of Corpus Linguistics* 14(1), 3–28.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4), 474–496.

Lu, X. (2011). A Corpus-Based Evaluation of Syntactic Complexity Measures as Indices of College-Level ESL Writers' Language Development. *TESOL Quarterly* 45(1), 36–62.

Lu, X. (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Languages Journal* pp. 190–208.

Lu, X. & H. Ai (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing* 29, 16–27.

Lubetich, S. & K. Sagae (2014). Data-driven Measurement of Child Language Development with Simple Syntactic Templates. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin, Ireland: Dublin City University and Association for Computational Linguistics, pp. 2151–2160. URL http://aclweb.org/anthology/C14-1203.

Malvern, D. D., R. B. J., C. N. & D. P. (2004). *Lexical diversity and language development: Quantification and assessment*. Palgrave Macmillan.

McCarthy, P. & S. Jarvis (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity

assessment. *Behavior Research Methods* 42(2), 381–392.
McNamara, D. S., M. M. Louwerse & A. C. Graesser (2002). Coh-Metrix: Automated Cohesion and Coherence Scores to Predict Text Readability and Facilitate Comprehension. Proposal of Project funded by the Office of Educational Research and Improvement, Reading Program. URL http://cohmetrix.memphis.edu/cohmetrixpr/archive/Coh-MetrixGrant.pdf.
Nikolova, L. (2015). Readability Classification for Bulgarian. Master thesis in computational linguistics, Department of Linguistics, University of Tübingen.
Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics* 24(4), 492–518.
Paquot, M. (2019). The phraseological dimension in interlanguage complexity research. *Second Language Research* 35(1), 121–145.
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.
Read, J. & P. Nation (2004). Measurement of formulaic sequences. *Formulaic sequences: Acquisition, processing and use* pp. 23–35.
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.
Sagae, K., A. Lavie & B. MacWhinney (2005). Automatic measurement of syntactic development in child language. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL-05)*. Ann Arbor, MI.
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.
Skehan, P. (1989). *Individual Differences in Second Language Learning*. Edward Arnold.
Unsworth, S. (2008). Comparing child L2 development with adult L2 development. In *Current trends in child second language acquisition: A generative perspective*, John Benjamins Publishing, pp. 301–346.
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.
Voss, M. J. (2005). *Determining Syntactic Complexity Using Very Shallow Parsing*. Research Report 2005-01, Computer Analysis of Speech for Psychological Research (CASPR), Institute for Artificial Intelligence, The University of Georgia. URL http://www.ai.uga.edu/caspr/2005-01-Voss.pdf. Published verison of MSc thesis.
Vyatkina, N. (2012). The Development of Second Language Writing Complexity in Groups and Individuals: A Longitudinal Learner Corpus Study. *The Modern Language Journal* 96(4), 576–598. URL https://doi.org/10.1111/j.1540-4781.2012.01401.x.
Weiss, Z. (2017). Using Measures of Linguistic Complexity to Assess German L2 Proficiency in Learner Corpora under Consideration of Task-Effects. Master's thesis, University of Tübingen, Germany. URL http://www.sfs.uni-tuebingen.de/~zweiss/ma-thesis/weiss2017-distr.pdf.
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.
Weiss, Z. & D. Meurers (2019a). Analyzing Linguistic Complexity and Accuracy in Academic Language Development of German across Elementary and Secondary School. In *Proceedings of the 14th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)*. Florence, Italy: Association for Computational Linguistics.
Weiss, Z. & D. Meurers (2019b). Broad Linguistic Modeling is Beneficial for German L2 Proficiency Assessment. In A. Abel, A. Glaznieks, V. Lyding & L. Nicolas (eds.), *Widening the Scope of Learner Corpus Research. Selected Papers from the Fourth Learner Corpus Research Conference*. Louvain-La-Neuve: Presses Universitaires de Louvain.
Weiss, Z., A. Riemenschneider, P. Schröter & D. Meurers (2019). Computationally Modeling the Impact of Task-Appropriate Language Complexity and Accuracy on Human Grading of German Essays. In *Proceedings of the 14th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)*. Florence, Italy.
Wolfe-Quintero, K., S. Inagaki & H.-Y. Kim (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy & Complexity*. Honolulu: Second Language Teaching & Curriculum Center, University of Hawaii at Manoa. URL https://doi.org/10.2307/3587656.