ISCL Hauptseminar
Summer Semester 2013

Computational Approaches to Text Simplification

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 is correlated with stages of acquisition (together with accuracy and fluency). On the applied side, complexity measures have long been used to determine the readability of a given text, and some readability measures have recently been automated in computational linguistics.

For several of these notions of complexity, researchers have raised the question how a given text or sentence could be simplified. This includes a range of application contexts, from shortening sentences to make parsing more efficient to making texts accessible to children, second language learners, or people with disabilities.

In this seminar, we will discuss the empirical and conceptual nature of these notions of complexity and explore computational approaches to text simplification building on these.

Instructors:

- Detmar Meurers
  - *Office:* Room 1.28, Blochbau (Wilhelmstr. 19)
  - *Email:* dm@sfs.uni-tuebingen.de
  - *Office hours:* Mondays 18:00 (best to email beforehand)
- Sowmya V. B.
  - *Office:* Room 1.29, Blochbau (Wilhelmstr. 19)
  - *Email:* sowmya@sfs.uni-tuebingen.de
  - *Office hours:* arrange by email

Course meets: in Seminarraum 1.13, Blochbau (Wilhelmstr. 19)

- Mondays and Wednesdays, 16ct-18

Credits:

- Credit Points: 10 (MA ISCL)
- Credit Points: 15 (including BA thesis for BA ISCL)

Syllabus (this file):

- html-Version (http://purl.org/dm/13/ss/simplification)
- pdf-Version (http://purl.org/dm/13/ss/simplification/syllabus.pdf)

Moodle page: https://moodle02.zdv.uni-tuebingen.de/course/view.php?id=470

Nature of course and our expectations: This is a Hauptseminar (and BA thesis seminar) which on the one hand intends to provide an overview of current perspectives and approaches and on the other hand, offers computational linguistics students the opportunity to define and implement a text simplification approach, which typically forms the basis of the term paper or BA thesis project.

Students enrolled in the course are expected to

1. regularly and actively participate in class, read the papers assigned by any of the presenters and post a question on Moodle to the"Reading Discussion Forum" on each reading *at the latest on the day before it is discussed* in class. (30% of grade)

   Note: Following the general university rules, missing more than two meetings unexcused, automatically results in failing the class.

2. explore and present a topic (30% of grade):

   - select one of the sub-topics during the first week of the semester
   - thoroughly research the topic, taking our literature pointers *as a starting point*
   - start a new Moodle thread on the "Reading Discussion Forum" specifying what every course participant should read to prepare for your presentation *a week before your presentation*
   - prepare the presentation with slides and send them to both instructors *at least two days before the presentation*
   - present the topic in class

3. work out a project term paper (40% of grade)

   - select a topic and submit a one-page abstract *by July 22*
     - For computational linguistics students, the topic of the paper in general will be the exploration and implementation of an approach analyzing the complexity or performing text simplification.
   - email the term paper in pdf format to both instructors *before the beginning of the next semester, i.e., by September 30, 2013.*
     - Note for ISCL students: The term paper must be produced in LaTeX using the Computational Linguistics journal style (http://cljournal.org/style.html), and 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: In principle, most of this is obvious, but to be explicit and clear: Please do not read or work on materials for other classes in our seminar. Come to class on time and do not pack up early. All portable electronic devices such as cell phones should be switched off for the entire length of the flight, oops, class. Laptops should not be open in class unless there is a concrete, assigned activity. If for some reason, you must leave early or you have an important call coming in, or you have to miss class for an important reason, please let the instructor know *before* class.

Topics:

1. **Introduction**

   What is readability? What is simplification? Applications, Motivations, Issues

   - Reading assignments:
     - April 22: (Vajjala & Meurers 2012)
     - April 24/29: (Feng 2008)

   Corpus studies, comparison of easy and difficult texts, psycholinguistic perspective on sentence comprehension

   - Reading assignments: (Allen 2009b)
   - Other relevant reading: (Petersen & Ostendorf 2007; Allen 2009a; Blum & Levenston 1978; Fukazawa 1994; Bautista 2010)
   - Chap 10 in (Aitchison 2011)
   - Simple Wiki guidelines[1]
   - (Bautista et al. 2011)
   - http://www.boeing.com/boeing/phantom/sechecker/features.page

2. Background: Early Work

   - Text simplification for Machines (Chandrasekar et al. 1996; Chandrasekar & Srinivas 1996)
   - Text Simplification for Humans (Carroll et al. 1998, 1999; Canning & Tait 1999; Canning et al. 2000; Inui et al. 2003; Devlin & Unthank 2006)
     (Devlin 1999; Canning 2002) are PhD theses on this topic, which are not freely available online.

3. A real-world application: The Porsimples project (Aluísio et al. 2008; Candido et al. 2009; Aluísio & Gasperin 2010a,b; Belder & Moens 2010; Gasperin et al. 2009a)

4. Lexical Simplification:

[1] http://simple.wikipedia.org/wiki/Wikipedia:Rules

- Corpus creation (De Belder & Moens 2012)
- SemEval 2012 Lexical Simplification task systems (Specia et al. 2012, ...)
- Lexical Simplification via WordNet (Thomas & Anderson 2012) and via synonym replacement (Keskisärkkä 2012, for Swedish)
- A relatively full-fledged system for Spanish Lexical Simplification (Bott et al. 2012a; Drndarevic & Saggion 2012b,a; Drndarevic et al. 2012)
- More Advanced Lexical Simplification (Yatskar et al. 2010; Biran et al. 2011)

5. Syntactic Simplification

   - Introduction (Gasperin et al. 2010)
   - Coh-Metrix-based research (Crossley et al. 2012)
   - Syntactic simplification with rules (Seretan 2012; Junior et al. 2011; Siddharthan 2011; Aranzabe et al. 2012b)
   - Syntactic simplification using statistical machine translation (Specia 2010; Zhu et al. 2010; Coster & Kauchak 2011a,b; Woodsend & Lapata 2011a,b; Bach et al. 2011)

6. Advanced Simplification issues and deeper linguistic aspects in text simplification, including preservation of discourse structure and cohesion (Siddharthan 2002a,b, 2003, 2004, 2006; Siddharthan & Katsos 2010; Siddharthan & Copestake 2002; Kandula et al. 2010) Siddharthan (2006) is a brief version of (Siddharthan 2004) thesis.

7. Identifying targets for simplification (Gasperin et al. 2009b; Medero & Ostendorf 2011; Bott & Saggion 2011a; Štajner, Drndarevic & Saggion 2013)

8. Corpus Creation (and systems) for other languages

   - Basque (Aranzabe et al. 2012a)
   - Brazilian Portuguese (Caseli et al. 2009)
   - Danish (Klerke 2012; Klerke & Søgaard 2012)
   - Italian (Barlacchi & Tonelli 2013)
   - Spanish (Bott & Saggion 2011b; Bott et al. 2012b)
   - Swedish (Smith & Jönsson 2011; Keskisärkkä 2012)

9. Application to other fields:

   - Question Generation (Heilman & Smith 2010)
   - Relation Extraction (Miwa et al. 2010)
   - Spoken Language Understanding (Tur et al. 2011)
   - Information Extraction (Jonnalagadda & Gonzalez 2010; J.Evans 2011)
   - Ontology Mapping (Lu & Parameswaran 2009)
   - Subtitling (Daelemans et al. 2004; Bouayad-Agha et al. 2006)
   - Biomedical (Jonnalagadda & Gonzalez 2009; Jonnalagadda et al. 2009)
   - Medical Literature (Damay et al. 2006)

- Crisis Management Domain (Temnikova 2012)
- Summarization for Simplification and Simplification for Summarization (Lal & Rüger 2002; Blake et al. 2007; Margarido et al. 2008; Smith & Jönsson 2011; Advaith Siddharthan & McKeown 2004)

Scheduling

Note that the following session plan is subject to change; it only constitutes the current state of our planning as the semester unfolds.

1. Wednesday, April 17: Organization and Introduction [Detmar Meurers]
2. Monday, April 22: Overview [Sowmya V.B.]
3. Wednesday, April 24: cont.
4. Monday, April 29: cont.
5. *Wednesday, May 1: Labor Day (the real one)*
6. Monday, May 6: cont.rr [Detmar Meurers]
7. *Wednesday, May 8: LEAD grad school*
8. Monday, May 13: TBD [NN]
9. Wednesday, May 15: TBD [NN]
10. *Monday, May 20: Pentecost holiday*
11. Wednesday, May 22: TBD [NN]
12. Monday, May 27: TBD [NN]
13. Wednesday, May 29: TBD [NN]
14. Monday, June 3: TBD [NN]
15. Wednesday, June 6: TBD [NN]
16. Monday, June 3: TBD [NN]
17. Wednesday, June 6: TBD [NN]
18. Monday, June 10: Project session
19. Wednesday, June 12: Project session
20. Monday, June 17: TBD [NN]
21. Wednesday, June 19: TBD [NN]
22. Monday, June 24: TBD [NN]
23. Wednesday, June 26: TBD [NN]
24. Monday, July 1: TBD [NN]
25. Wednesday, July 3: TBD [NN]
26. Monday, July 8: TBD [NN]
27. Wednesday, July 10: TBD [NN]
28. Monday, July 15: TBD [NN]
29. Wednesday, July 17: TBD [NN]

Last update: April 28, 2013

References

Advaith Siddharthan, A. N. & K. McKeown (2004). Syntactic Simplification for Improving Content Selection in Multi-Document Summarization. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*.

Aitchison, J. (2011). *The Articulate Mammal: An Introduction to Psycholinguistics*. Routledge Classics.

Allen, D. (2009a). A study of the role of relative clauses in the simplification of news texts for learners of English. *System* 37(4), 58–599. URL http://www.sciencedirect.com/science/article/B6VCH-4XHJX3D-1/2/c8c06bd6afeddc1ec4e8df219f39a83e.

Allen, D. (2009b). Using a corpus of simplified news texts to investigate features of the intuitive approach to simplification. In *Proceedings of the Corpus Linguistics Conference 2009*.

Aluísio, S. M., L. Specia, T. A. Pardo, E. G. Maziero & R. P. Fortes (2008). Towards Brazilian Portuguese automatic text simplification systems. In *Proceeding of the eighth ACM symposium on Document engineering*. New York, NY, USA, DocEng '08, pp. 240–248. URL http://doi.acm.org/10.1145/1410140.1410191.

Aluísio, S. M. & C. Gasperin (2010a). Fostering Digital Inclusion and Accessibility: The PorSimples project for Simplification of Portuguese Texts. In *Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas*.

Aluísio, S. M. & C. Gasperin (2010b). PorSimples: Simplification of Portuguese Texts Fostering Digital Inclusion and Accessibility. A presentation on Porsimples project.

Aranzabe, M. J., A. D. de Ilarraza & I. Gonzalez-Dios (2012a). First Approach to Automatic Text Simplification in Basque. In *The worshop Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*.

Aranzabe, M. J., A. D. de Ilarraza & I. Gonzalez-Dios (2012b). Transforming Complex Sentences using Dependency Trees for Automatic Text Simplification in Basque. In *SEPLN Journal*. URL http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/4660/2762.

Bach, N., Q. Gao, S. Vogel & A. Waibel (2011). TriS: A Statistical Sentence Simplifier with Log-linear Models and Margin-based Discriminative Training. In *Proceedings of 5th International Joint Conference on Natural Language Processing*. Chiang Mai, Thailand: Asian Federation of Natural Language Processing, pp. 474–482. URL http://aclweb.org/anthology/I11-1053.

Barlacchi, G. & S. Tonelli (2013). ERNESTA: A Sentence Simplification Tool for Children's Stories in Italian. In *14th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2013*.

Bautista, S. (2010). Semiautomatic Simplification to Improve Readability of Texts for People with Special Needs. NLG Group Meeting September 9th 2010 The Open University - Presentation.

Bautista, S., C. León, R. Hervás & P. Gervás (2011). Empirical Identification of Text Simplification Strategies for Reading-Impaired People. In *European Conference for the Advancement of Assistive Technology*.

Belder, J. D. & M.-F. Moens (2010). Text Simplification For Children. In *SIGIR workshop*

on accessible search systems, 2010.

Biran, O., S. Brody & N. Elhadad (2011). Putting it Simply: a Context-Aware Approach to Lexical Simplification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA: Association for Computational Linguistics, pp. 496–501. URL http://aclweb.org/anthology/P11-2087.

Blake, C., J. Kampov, A. K. Orphanides, D. West & C. Lown (2007). UNC-CH at DUC 2007: Query Expansion, Lexical Simplification and Sentence Selection Strategies for Multi-Document Summarization. In *Proceedings of Document Understanding Conference*.

Blum, S. & E. A. Levenston (1978). Universals of Lexical Simplification. *Language Learning* 28, 399–415.

Bott, S., L. Rello, B. Drndarevic & H. Saggion (2012a). Can Spanish Be Simpler? LexSiS: Lexical Simplification for Spanish. In *In Proceedings of the 24th International Conference on Computational Linguistics (COLING)*.

Bott, S. & H. Saggion (2011a). Spanish Text Simplification: An Exploratory Study. In *27th CONFERENCE OF THE SPANISH SOCIETY FOR NATURAL LANGUAGE PROCESSING*.

Bott, S. & H. Saggion (2011b). An Unsupervised Alignment Algorithm for Text Simplification Corpus Construction. In *ACL Workshop on Monolingual Text-to-Text Generation*.

Bott, S., H. Saggion & D. Figueroa (2012b). A Hybrid System for Spanish Text Simplification. In *NAACL-HLT 2012 Workshop on Speech and Language Processing for Assistive Technologies (SLPAT)*.

Bouayad-Agha, N., A. Gil, O. Valentin & V. Pascual (2006). A Sentence Compression Module for Machine-Assisted Subtitling. In *CICLing*. pp. 490–501.

Candido, Jr., A., E. Maziero, C. Gasperin, T. A. S. Pardo, L. Specia & S. M. Aluisio (2009). Supporting the adaptation of texts for poor literacy readers: a text simplification editor for Brazilian Portuguese. In *Proceedings of the Fourth Workshop on Innovative Use of NLP for Building Educational Applications*. Stroudsburg, PA, USA, EdAppsNLP '09, pp. 34–42. URL http://portal.acm.org/citation.cfm?id=1609843.1609848.

Canning, Y. (2002). Syntactic Simplification of Text. Ph.D. thesis, University of Sunderland.

Canning, Y. & J. Tait (1999). Syntactic Simplification of Newspaper Text for Aphasic Readers. In *Proceedings of SIGIR-99 Workshop on Customised Information Delivery*. Berkeley, CA, pp. 6–11.

Canning, Y., J. Tait, J. Archibald & R. Crawley (2000). Cohesive Generation of Syntactically Simplified Newspaper Text. In *Third International Workshop on Text, Speech and Dialogue, TSD 2000, Brno, Czech Republic, September 13-16, 2000*.

Carroll, J., G. Minnen, Y. Canning, S. Devlin & J. Tait (1998). Practical Simplification of English Newspaper Text to Assist Aphasic Readers. In *Proceedings of the AAAI-98 Workshop on Integrating Artificial Intelligence and Assistive Technology*. Madison, Wisconsin: Association for the Advancement of Artificial Intelligence (AAAI). URL http://www.informatics.susx.ac.uk/research/groups/nlp/carroll/papers/aaai98.pdf.

Carroll, J., G. Minnen, D. Pearce, Y. Canning, S. Devlin & J. Tait (1999). Simplifying Text

for Language-Impaired Readers. In *Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. pp. 269–270.

Caseli, H., T. Pereira, L. Specia, T. Pardo, C. Gasperin & S. Aluísio (2009). Building a Brazilian Portuguese parallel corpus of original and simplified texts. In *In Proceedings of the 10th Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2009)*.

Chandrasekar, R., C. Doran, B. Srinivas & R. Ch (1996). Motivations and Methods for Text Simplification. URL http://aclweb.org/anthology/C96-2183.pdf.

Chandrasekar, R. & B. Srinivas (1996). *Automatic Induction of Rules for Text Simplification*. Tech. Rep. IRCS Report 96–30, Upenn, NSF Science and Technology Center for Research in Cognitive Science.

Coster, W. & D. Kauchak (2011a). Learning to Simplify Sentences Using Wikipedia. In *Proceedings of the Workshop on Monolingual Text-To-Text Generation*. Portland, Oregon: Association for Computational Linguistics, pp. 1–9. URL http://aclweb.org/anthology/W11-1601.

Coster, W. & D. Kauchak (2011b). Simple English Wikipedia: A New Text Simplification Task. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA: Association for Computational Linguistics, pp. 665–669. URL http://aclweb.org/anthology/P11-2117.pdf.

Crossley, S. A., D. Allen & D. S. McNamara (2012). Text simplification and comprehensible input: A case for an intuitive approach. In *Language Teaching Research*. vol. 16. URL http://ltr.sagepub.com/content/16/1/89.

Daelemans, W., A. Höthker & E. T. K. Sang (2004). Automatic Sentence Simplification for Subtitling in Dutch and English. In *In Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC)*.

Damay, J. J. S., G. J. D. Lojico, K. A. L. Lu, D. B. Tarantan & E. C. Ong (2006). SIMTEXT: Text Simplification of Medical Literature. In *3rd National Natural Language Processing Symposium - Building Language Tools and Resources*.

De Belder, J. & M.-F. Moens (2012). A dataset for the evaluation of lexical simplification. *Lecture Notes in Computer Science* 7182, 426–437.

Devlin, S. (1999). Simplifying Natural Language for Aphasic Readers. Ph.D. thesis, University of Sunderland.

Devlin, S. & G. Unthank (2006). Helping aphasic people process online information. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. New York, NY, USA: ACM, Assets '06, pp. 225–226. URL http://doi.acm.org/10.1145/1168987.1169027.

Drndarevic, B. & H. Saggion (2012a). Reducing Text Complexity through Automatic Lexical Simplification: an Empirical Study for Spanish. *The Spanish Society for Natural Language Processing (SEPLN)* 49, –.

Drndarevic, B. & H. Saggion (2012b). Towards Automatic Lexical Simplification in Spanish: An Empirical Study. In *Proceedings of the First Workshop on Predicting and Improving Text Readability for target reader populations*. Montréal, Canada: Association for Computational Linguistics, pp. 8–16.

Drndarevic, B., S. Štajner & H. Saggion (2012). Reporting Simply: A Lexical Simplifi-

cation Strategy for Enhancing Text Accessibility. In *Proceedings of "Easy to read on the web" online symposium*. URL http://www.w3.org/WAI/RD/2012/easy-to-read/paper7/.

Feng, L. (2008). *Text Simplification: A Survey*. Tech. rep., CUNY. URL http://lijun.symptotic.com/files/TextSimplification.pdf.

Fukazawa, S. (1994). A Study of Simplification Strategies by Native Speakers of English Use of Discourse Markers. In *Bulletin of Research Center for Educational Study and Practice*.

Gasperin, C., E. Maziero & S. M. Alusio (2010). Challenging Choices for Text Simplification. In *Proceedings of the 9th International Conference on the Computational Processing of the Portuguese Language*.

Gasperin, C., E. Maziero, L. Specia, P. T.S.P. & S. Aluisio (2009a). Natural language processing for social inclusion: a text simplification architecture for different literacy levels. In *XXXVI Seminário Integrado de Software e Hardware (SEMISH-2009)*. Bento Gonçalves, Brazil, pp. 387–401.

Gasperin, C., L. Specia, T. F. Pereira & S. M. Aluisio (2009b). Learning When to Simplify Sentences for Natural Text Simplification. In *Encontro Nacional de Inteligência Artificial (ENIA-2009)*.

Heilman, M. & N. Smith (2010). Extracting Simplified Statements for Factual Question Generation. In *In Proceedings of the Third Workshop on Question Generation*.

Inui, K., A. Fujita, T. Takahashi, R. Iida & T. Iwakura (2003). Text Simplification for Reading Assistance: A Project Note. In *Proceedings of the Second International Workshop on Paraphrasing, held at ACL 2003*. URL http://aclweb.org/anthology/W03-1602.

J.Evans, R. (2011). Comparing methods for the syntactic simplification of sentences in information extraction. In *Literary and Linguistic Computing*, Oxford University Press.

Jonnalagadda, S. & G. Gonzalez (2009). Sentence Simplification Aids Protein-Protein Interaction Extraction. In *Proceedings of The 3rd International Symposium on Languages in Biology and Medicine, Jeju Island, South Korea, November 8-10, 2009*.

Jonnalagadda, S. & G. Gonzalez (2010). BioSimplify: an open source sentence simplification engine to improve recall in automatic biomedical information extraction. In *AMIA Annual Symposium Proceedings*.

Jonnalagadda, S., L. Tari, J. Hakenberg, C. Baral & G. Gonzalez (2009). Towards Effective Sentence Simplification for Automatic Processing of Biomedical Text. In *Proceedings of the NAACL-HLT 2009, Boulder, USA, June*.

Junior, A. C., A. Copestake, L. Specia & S. M. Aluísio (2011). Towards an on-demand Simple Portuguese Wikipedia. In *Proceedings of the 2nd Workshop on Speech and Language Processing for Assistive Technologies*.

Kandula, S., D. Curtis & Q. Zeng-Treitler (2010). A semantic and syntactic text simplification tool for health content. In *In Proceedings of AMIA Annual Symposium*.

Keskisärkkä, R. (2012). Automatic Text Simplification via Synonym Replacement. Master's thesis, Linköping University.

Klerke, S. (2012). Automatic Text Simplification in Danish: Sampling a restricted space of rewrites to optimize readability using lexical substitutions and dependency analyses. Master's thesis, University of Copenhagen.

Klerke, S. & A. Søgaard (2012). Danish parallel corpus for text simplification. In *In Proceedings of Language Resources and Evaluation Conference (LREC), 2012*.

Lal, P. & S. Rüger (2002). Extract-based Summarization with Simplification. In *In Proceedings of Document Understanding Conference (DUC) 2002*.

Lu, L. & N. Parameswaran (2009). Sentence Simplification Based Ontology Mapping. In *Proceedings of the Twenty-Second International FLAIRS Conference (2009)*.

Margarido, P. R. A., T. A. S. Pardo, G. M. Antonio, V. B. Fuentes, R. Aires, S. M. Aluísio & R. P. M. Fortes (2008). Automatic summarization for text simplification: evaluating text understanding by poor readers. In *WebMedia '08 Companion Proceedings of the XIV Brazilian Symposium on Multimedia and the Web*.

Medero, J. & M. Ostendorf (2011). Identifying Targets for Syntactic Simplification. In *ISCA International Workshop on Speech and Language Technology in Education (SLaTE 2011)*. URL http://project.cgm.unive.it/events/SLaTE2011/papers/Medero_ostendorf_SLaTE2011.pdf.

Miwa, M., R. Sætre, Y. Miyao & J. Tsujii (2010). Entity-Focused Sentence Simplification for Relation Extraction. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*.

Petersen, S. E. & M. Ostendorf (2007). Text Simplification for Language Learners: A Corpus Analysis. In *Speech and Language Technology for Education (SLaTE)*. URL http://sarahpetersen.net/portfolio/Petersen_Ostendorf_SLaTE2007_final.pdf.

Seretan, V. (2012). Acquisition of Syntactic Simplification Rules for French. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*.

Siddharthan, A. (2002a). An Architecture for a Text Simplification System. In *In Proceedings of the Language Engineering Conference 2002 (LEC 2002)*.

Siddharthan, A. (2002b). Resolving Attachment and Clause Boundary Ambiguities for Simplifying Relative Clause Constructs. In *Proceedings of the Student Research Workshop, 40th Meeting of the Association for Computational Linguistics (ACL 2002)*.

Siddharthan, A. (2003). Preserving Discourse Structure when Simplifying Text. In *Proceedings of the European Natural Language Generation Workshop (ENLG)*.

Siddharthan, A. (2004). *Syntactic simplification and text cohesion*. Tech. Rep. UCAM-CL-TR-597, University of Cambridge Computer Laboratory. URL http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-597.pdf.

Siddharthan, A. (2006). Syntactic Simplification and Text Cohesion. *Research on Language and Computation* 4(1), 77–109.

Siddharthan, A. (2011). Text Simplification using Typed Dependencies: A Comparison of the Robustness of Different Generation Strategies. In *Proceedings of the 13th European Workshop on Natural Language Generation (ENLG)*.

Siddharthan, A. & A. Copestake (2002). Generating Anaphora for Simplifying Text. In *Proceedings of the 4th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2002)*.

Siddharthan, A. & N. Katsos (2010). Reformulating Discourse Connectives for Non-Expert Readers. In *In Proceedings of Human Language Technologies: the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics*.

Smith, C. & A. Jönsson (2011). Automatic Summarization As Means Of Simplifying

Texts, An Evaluation For Swedish. In *Proceedings of the 18th Nordic Conference of Computational Linguistics NODALIDA 2011*.

Specia, L. (2010). Translating from complex to simplified sentences. In *Proceedings of the 9th international conference on Computational Processing of the Portuguese Language (PROPOR'10)*.

Specia, L., S. K. Jauhar & R. Mihalcea (2012). Semeval-2012 task 1: English lexical simplification. In *In Proceedings of the 6th International Workshop on Semantic Evaluation (SemEval 2012)*.

Temnikova, I. (2012). Text Complexity and Text Simplification in the Crisis Management domain. Ph.D. thesis, University of Wolverhampton, UK.

Thomas, S. R. & S. Anderson (2012). WordNet-based lexical simplification of a document. In *Proceedings of KONVENS 2012*.

Tur, G., D. Hakkani-Tür, L. Heck & S. Parthasarathy (2011). Sentence simplification for spoken language understanding. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*.

Vajjala, S. & D. Meurers (2012). On Improving the Accuracy of Readability Classification using Insights from Second Language Acquisition. In J. Tetreault, J. Burstein & C. Leacock (eds.), *Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA7) at NAACL-HLT*. Montréal, Canada: Association for Computational Linguistics, pp. 163—-173. URL http://aclweb.org/anthology/W12-2019.pdf.

Woodsend, K. & M. Lapata (2011a). Learning to Simplify Sentences with Quasi-Synchronous Grammar and Integer Programming. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. URL http://aclweb.org/anthology/D11-1038.

Woodsend, K. & M. Lapata (2011b). WikiSimple: Automatic Simplification of Wikipedia Articles. In *In Proceedings of the 25th National Conference on Artificial Intelligence*.

Yatskar, M., B. Pang, C. Danescu-Niculescu-Mizil & L. Lee (2010). For the sake of simplicity: Unsupervised extraction of lexical simplifications from Wikipedia. In *Proceedings of the NAACL*. pp. 365–368.

Zhu, Z., D. Bernhard & I. Gurevych (2010). A Monolingual Tree-based Translation Model for Sentence Simplification. In *Proceedings of The 23rd International Conference on Computational Linguistics (COLING), August 2010. Beijing, China*.
ΩŠtajner et al.

Štajner, S., B. Drndarevic & H. Saggion (2013). Corpus-based Sentence Deletion and Split Decisions for Spanish Text Simplification. In *CICLing 2013: The 14th International Conference on Intelligent Text Processing and Computational Linguistics*.