



Yannick Versley  
Institut für Computerlinguistik  
Im Neuenheimer Feld 325  
69120 Heidelberg

Email: [versley@cl.uni-heidelberg.de](mailto:versley@cl.uni-heidelberg.de)  
Telephone: +49-6221-54 3591  
WWW: <http://www.versley.de>

# Yannick Versley

## *Diplom-Informatiker, Dr. Phil. Computerlinguistik*

### General Information

- Date of Birth** • September 14th, 1979  
**Place of Birth** • Hamburg, Germany  
**Citizenship** • German  
**Languages** • German (native), English (near-native), French (fluent), Italian (basic), Spanish (very basic)  
**Research Interests** • Use of common sense knowledge in the context of the micro- and macrostructure of discourse; Methods for Lexical Acquisition; Machine Learning methods for structured data; Natural Language Processing techniques for German

### Education

#### University

- 2005–2010** • University of Tübingen, Seminar für Sprachwissenschaft.  
PhD in Computational Linguistics  
Thesis title: *Resolving Coreferent Bridging in German Newspaper text*  
Grade: *Magna cum laude*; Thesis Advisor: Prof. Erhard Hinrichs
- 1999–2004** • University of Hamburg, Department of Computer Science  
Degree obtained: Informatik-Diplom  
Thesis title: *Tagging kausaler Relationen*  
Grade: 1.4 (sehr gut); Thesis Advisor: Prof. Christopher Habel

#### School

- 1990–1998** • Gymnasium Osterbek, Hamburg  
Abitur; Grade: 1.4 (sehr gut)
- 1986–1990** • Lycée Français de Hambourg, Hamburg

### Work Experience

- 2013-current** • University of Heidelberg, Institute for Computational Linguistics  
Visiting professor (“Professurvertretung”)
- 2009–2013** • University of Tübingen, Collaborative Research Center 833  
Research associate in the project *A3: Desambiguierung von Diskurskon-  
nektoren mit korpusinduzierten semantischen Relationen*
- 2009** • University of Trento, Center for Mind/Brain Sciences (CiMeC)  
Research fellow in the project *LiveMemories*
- 2005–2008** • University of Tübingen, Collaborative Research Center 441

## Work Experience (continued)

Research associate in the project *A1: Representation and Automatic Acquisition of Linguistic Data*

### Part time / student employment

- 2007
  - Johns Hopkins Summer Workshop, Project *Encyclopedic and Lexical Knowledge for Entity Disambiguation*
  - Graduate Research Team Member
- 2003–2004
  - University of Hamburg, Knowledge and Language Processing group: Student research assistant
- 2001
  - Internship at Mummert+Partner, Hamburg
  - Customer-specific ABAP programming (SAP R/3)
- 2000–2002
  - Bitsdontbyte GbR, Hamburg
  - Lotus Notes programming in LotusScript and Java; Java Servlets; Apple WebObjects
- 2000–2003
  - University of Hamburg
  - Tutor *Praktische Informatik I* (FB Inf.), *Java-Programmierung* (RRZ)
- 1997–1998
  - Hamburger Bildungsserver, Hamburg
  - Linux installation in schools
- 1997
  - Internship at Ergole Informatique, Grenoble
  - GUI programming for Windows using C

## Teaching

- 2015
  - Mathematical foundations for CL
  - Structured Inference for NLP applications (*Hauptseminar*)
- 2014
  - Mathematical foundations for CL
  - Introduction to Computational Linguistics
  - Multimodal Semantics (*Hauptseminar*)
  - NLP methods for Digital Humanities (*Proseminar*)
  - Software project (SoSe, WiSe)
- 2013
  - Introduction to Computational Linguistics
  - Statistical Parsing (*Hauptseminar*)
  - Computational Linguistics in Context (*Proseminar*)
- 2009
  - Anaphora Resolution
  - (with Prof. Massimo Poesio, Kepa Rodriguez)
  - Kurs bei der *5th DGfS Fall School*,
  - September 2009, Universität Konstanz.

### Teaching Assistant / Tutor

- 2000–2003
  - Praktische Informatik I
  - (Prof. Wolfgang Menzel, Prof. Leonie Dreschler Fischer)
  - Fachbereich Informatik, Universität Hamburg*
- 2001
  - Java-Programmierung
  - (Bernd Eggink)
  - Regionales Rechenzentrum (RRZ), Universität Hamburg*

## Administration

- 2000–2003 • Studienreformausschuss (SRA; studentisches Mitglied)  
*Fachbereich Informatik, Universität Hamburg*
- 2001–2004 • Prüfungsausschuss (PA; studentisches Mitglied)  
*Fachbereich Informatik, Universität Hamburg*

## Publications

### Journal Articles

- Yannick Versley (2013): A graph-based approach for implicit discourse relations. *CLIN Journal* 3:148–173.
- Yannick Versley and Anna Gastel (2013): Linguistic Tests for Discourse Relations. Stefanie Dipper, Bonnie Webber and Heike Zinsmeister (eds.): *Dialogue and Discourse* 4(2). Special Issue on *Beyond Semantics: The Challenges of Annotating Pragmatic and Discourse Phenomena*.  
Contributions: [YV] General conception of the paper, writing; [AG] writing; example selection
- Heike Telljohann, Yannick Versley, Kathrin Beck, Erhard Hinrichs and Thomas Zastrow (2013): STTS als Part-of-Speech-Tagset in Tübinger Baumbanken (in German). *Journal for Language Technology and Computational Linguistics* 28(1):1–16.  
Contributions: [HT] Details on treebanks and treebank annotation schemes, writing; [YV] Experimental part of the paper, general conception; writing; [KB, EH] Comments on details of annotation schemes.
- Yannick Versley (2008): Vagueness and Referential Ambiguity in a Large-scale Annotated Corpus. Massimo Poesio and Ron Artstein (eds.): *Ambiguity and Semantic Judgement*. Special Issue of the Journal on Research in Language and Computation.

### Conference papers

- Michael Haas and Yannick Versley (to appear): Subsentential Sentiment on a Shoestring: A Crosslingual Analysis of Compositional Classification. Accepted for: *2015 Conference of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL HLT 2015)*.  
Contributions: [MH] Implementation and experimental work; [YV] General conception for the conference paper, research supervision, writing.
- Yannick Versley (2013): Graph-based Classification of Explicit and Implicit Discourse Relations. *International Conference on Computational Semantics (IWCS 2013)*, Potsdam, Germany.
- Yannick Versley (2011): Multilabel tagging of discourse relations in ambiguous temporal connectives. *Proceedings of Recent Advances in Natural Language Processing (RANLP 2011)*.
- Samuel Broscheid, Simone Ponzetto, Yannick Versley and Massimo Poesio (2010): Extending BART to provide a coreference resolution system for German. *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010)*  
Contributions: [SB] Implementation of coreference features for German, experiments; [SP] supervision of SB, writing; [YV] Data conversions, German preprocessing and mention extraction, writing; [MP] general ideas and comments

## Publications (continued)

- Massimo Poesio, Olga Uryupina and Yannick Versley (2010): Creating a Coreference Resolution System for Italian. *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010)*  
Contributions: [MP] supervision of work, general ideas, writing; [OU] Implementation of coreference features for Italian, experiments, writing; [YV] Preprocessing for Italian; Italian-specific adaptations for the BART framework.
- Yannick Versley, Kathrin Beck, Erhard Hinrichs and Heike Telljohann (2010): A Syntax-first approach to High-quality Morphological Analysis and Lemma Disambiguation for the TüBa-D/Z Treebank. *Proceedings of the 9th Conference on Treebanks and Linguistic Theories (TLT9)*.  
Contributions: [YV] Lemmatizer implementation and experiments, general conception writing; [KB, EH, HT] annotation guidelines for closed-class lemmas, general description on the treebank, supervision of the lemma annotation of the gold standard used.
- Yannick Versley and Ines Rehbein (2009): Scalable Discriminative Parsing for German. *International Conference on Parsing Technology (IWPT'09)*.  
Contributions: [YV] Parser implementation, experiments, paper conception, writing; [IR] General discussion, insights on the Tiger annotation scheme.
- Yannick Versley, Alessandro Moschitti, Massimo Poesio and Xiaofeng Yang (2008): Coreference Systems based on Kernel Methods. *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*.  
[YV] Integration of Kernel-based learning into BART, expletive kernel, experiments, general conception, writing; [AM] word sequence kernel, writing; [XY] binding kernel; [MP] general discussion, general conception, comments on paper.
- Yannick Versley (2007): Antecedent Selection Techniques for High-Recall Coreference Resolution. *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*.
- Yannick Versley (2007): Using the Web to Resolve Coreferent Bridging in German Newspaper Text. *Proceedings der GLDV-Frühjahrestagung 2007*.

## Workshop Papers

- Yannick Versley (2014): Experiments with Easy-first nonprojective constituent parsing. *Proceedings of the First Joint Workshop on Statistical Parsing of Morphologically Rich Languages and Syntactic Analysis of Non-Canonical Languages*.
- Yannick Versley (2013): SFS-TUE: Compound Paraphrasing with a Language Model and Discriminative Reranking. *Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*, Atlanta, US.
- Yannick Versley (2012): Supervised Learning of German Qualia Relations. *ACL 2012 Joint Workshop on Statistical Parsing and Semantic Processing of Morphologically Rich Languages (SP-Sem-MRL 2012)*
- Yannick Versley and Yana Panchenko (2012): Not Just Bigger: Towards Better-Quality Web Corpora. *Proceedings of the 7th Web as Corpus Workshop at WWW2012 (WAC7)*.
- Yannick Versley (2011): Towards finer-grained tagging of discourse connectives. *AG Beyond Semantics, Deutsche Gesellschaft für Sprachwissenschaft (DGfS 2011)*.
- Yannick Versley (2010): Discovery of Ambiguous and Unambiguous Discourse Connectives via Annotation Projection. *Workshop on the Annotation and Exploitation of Parallel Corpora (AEPC)*.

## Publications (continued)

- Marta Recasens, Lluís Màrquez, Emili Sapena, M. Antònia Martí, Mariona Taulé, Véronique Hoste, Massimo Poesio, and Yannick Versley (2010): SemEval-2010 Task 1: Coreference Resolution in Multiple Languages. In *Proceedings of the ACL Workshop on Semantic Evaluations (SemEval-2010)*.
- Reut Tsarfaty, Djamé Seddah, Yoav Goldberg, Sandra Kuebler, Yannick Versley, Marie Candito, Jennifer Foster, Ines Rehbein and Lamia Tounsi (2010): Statistical Parsing of Morphologically Rich Languages (SPMRL): What, How and Whither. *Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages*.
- Yannick Versley (2008): Decorrelation and Shallow Semantic Patterns for Distributional Clustering of Nouns and Verbs. Stefan Evert and Marco Baroni (eds.), *Proceedings of the ESSLLI'08 Workshop on Distributional Lexical Semantics*.
- Yannick Versley, Simone Paolo Ponzetto, Massimo Poesio, Vladimir Eidelman, Alan Jern, Jason Smith, Xiaofeng Yang, Alessandro Moschitti (2008): BART: A Modular Toolkit for Coreference Resolution. In *Companion Volume of the Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL 2008)*.
- Yannick Versley, Holger Wunsch and Heike Zinsmeister (2007): A Pilot Study on Computer-aided Coreference Annotation. Constantin Orasan and Sandra Kübler (eds.) *Proceedings of the International Workshop on Computer Aided Language Processing (CALP) 2007*.
- Yannick Versley and Heike Zinsmeister (2006): From Dependency Parsing to Deep(er) Semantics. *Proceedings of the Fifth International Workshop on Treebanks and Linguistic Theories (TLT 2006)*.
- Yannick Versley (2006): A Constraint-based Approach to Noun Phrase Coreference Resolution in German Newspaper Text. *Konferenz zur Verarbeitung Natürlicher Sprache (KONVENS 2006)*.
- Yannick Versley (2006): Disagreement Dissected: Vagueness as a Source of Ambiguity in Nominal (Co-)Reference. Ron Artstein and Massimo Poesio (eds.), *Proceedings of the ESSLLI 2006 Workshop on Ambiguity in Anaphora*
- Yannick Versley (2005): Parser Evaluation across Text Types. *Proceedings of the Fourth Workshop on Treebanks and Linguistic Theories (TLT 2005)*.
- Schilder, F., Versley, Y., and Habel, Ch. (2004) Extracting spatial information: grounding, classifying and linking spatial expressions. Ross Purves and Christopher B. Jones (eds.), *SIGIR Workshop on Geographic Information Retrieval*.

## Edited Volumes

- Yoav Goldberg, Yuval Marton, Yannick Versley, Özlem Cetinoğlu, Ines Rehbein, Joel Tetrault, Sandra Kübler, Djamé Seddah and Reut Tsarfaty (2014): *Proceedings of the First Joint Workshop on Statistical Parsing of Morphologically Rich Languages and Syntactic Analysis of Non-Canonical Languages (SPMRL-SANCL 2014)*.
- Yoav Goldberg, Yuval Marton, Ines Rehbein, Yannick Versley, Sandra Kübler, Djamé Seddah and Reut Tsarfaty (2013): *Proceedings of the Fourth Workshop on Statistical Parsing of Morphologically Rich Languages (SPMRL 2013)*.
- Yves Peirsman, Yannick Versley and Tim Van de Cruys (2009): *Proceedings of the CogSci 2009 Workshop on Distributional Semantics beyond Concrete Concepts (DisCo 2009)*.

## Publications (continued)

- Massimo Poesio, Roland Stuckardt and Yannick Versley (in preparation): *Anaphora Resolution*. Book in preparation, to be published by Springer.
- Sam Featherston and Yannick Versley (in preparation): *Firm Foundations: Quantitative Studies of Sentence Grammar and Grammatical Change in Germanic*. Book in preparation, to be published by De Gruyter in the *Trends in Linguistics. Studies and Monographs (TiLSM)* series.

## Theses

- Yannick Versley (2010) Resolving Coreferent Bridging in German Newspaper Text. PhD Thesis, Seminar für Sprachwissenschaft, Universität Tübingen.  
<http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-57748>
- Yannick Versley (2004) Tagging kausaler Relationen (in German). Diploma Thesis. Fachbereich Informatik, Universität Hamburg