Curriculum Vitae

Dr. Amir Zeldes

Postal address:	Department of Linguistics Georgetown University 1421 37th St. NW Poulton Hall 257 Washington, DC 20057 USA
E-Mail:	amir.zeldes@georgetown.edu
Website:	http://corpling.uis.georgetown.edu/amir

Education

12/2007 – 1/2012	Doctorate in General Linguistics (*summa cum laude*) at Humboldt University, Berlin. Supervisors: Prof. Anke Lüdeling, HU Berlin, Prof. Stefan Th. Gries, UC Santa Barbara. Thesis (2011): "Productivity in Argument Selection: A Usage-Based Approach to Lexical Choice in Syntactic Slots". (Published as a book 2012, see Publications).
10/2006 – 12/2007	Magister Artium (1st honors, *mit Auszeichnung*) at Humboldt University, Berlin (major in German Linguistics, minor in Comparative Historical Linguistics) and at Potsdam University (minor in Computational Linguistics). Thesis: "Data-Based Methods for Historical Grammar and Lexicon Extraction in a Diachronic Corpus". Final grade of 1,0.
10/2005 – 9/2006	DAAD Exchange Scholarship Program, Humboldt University, Berlin, Institute for German Language and Linguistics. Research project: Application of Machine Translation Techniques to Historical Corpora.
2002 – 2005	B.A. (*summa cum laude*) at the Hebrew University of Jerusalem, Israel, Majors in General Linguistics and Cognitive Science. GPA of 97.42%.

Work Experience

8/2014 – present	Assistant Professor of Computational Linguistics, Department of Linguistics, Georgetown University, Washington, DC.
3/2014 – 7/2014	Director of the BMBF funded eHumanities Junior Research Group KOMeT (corpus linguistics methods for e-Philology with TEI) at HU Berlin.
1/2012 – 2/2014	Post-doc at Collaborative Research Center 632 on Information Structure, Project D1 "Linguistic Database for Information Structure" (Prof. A. Lüdeling, HU Berlin and Prof. M. Stede, Potsdam University).
12/2007 – 1/2012	Doctoral researcher at Collaborative Research Center 632 on Information Structure, Project D1 "Linguistic Database for Information Structure" (Prof. A. Lüdeling, HU Berlin and Prof. M. Stede, Potsdam University).
10/2006 – 12/2007	Student Assistant at Humboldt University, Berlin, Institute fir German Language and Linguistics, Chair of Corpus Linguistics and Morphology (Prof. A. Lüdeling).

10/2004 – 7/2005	Student Assistant, Hebrew University of Jerusalem, Israel, Department of Egyptology and Ancient Near Eastern Studies (Prof. O. Goldwasser).
2/2004 – 7/2005	Research Assistant, Hebrew University of Jerusalem, Israel, Institute of Asian and African Studies (Prof. Em. P. Shinar).
10/2003 – 10/2004	Cataloguer of Arabic and Japanese Books, Israeli National Library, Jerusalem, Israel. Department of Asian and African Studies.

Teaching Experience

Summer Term 2014	52 20063 SE Introduction to Coptic Grammar (Wed 14-16 DOR24)
	52 20061 KO Research Colloquium Corpus Linguistics (Wed 18-20 DOR24, together with Hagen Hirschmann)
	MGK Workshop Mixed-effects Models and Logistic Regression with R (Universität Potsdam, Campus Golm, Bdg. 14, R. 2.15/16, 3.7.2014)
Winter Term 2013/14	Multilayer and Parallel Corpora: An Introduction Using ANNIS (Hong Kong City University, Block Seminar, 13.11.2013)
Summer Term 2013	Theoretical and Empirical Approaches to Argument Structure (HU Berlin, Seminar 52 20015)
	Corpus Linguistics and Non-Compositionality (Minerva Non-Compositionality Summer School, Jerusalem 23-28 June 2013)
	Statistical Methods in Linguistics (1-5 July 2013, Hebrew University of Jerusalem intensive MA course 10817)
Winter Term 2012/13	Syntactic Annotation (HU Berlin, Seminar 52 20077)
	Multifactorial Methods with R (HU Berlin, 4th Emp. Methods Workshop)
	Introduction to Statistics (Tutorial at 35th Annual Conference of the German Association for Linguistics [DGfS], Potsdam University)
Summer Term 2012	Revival of the Hebrew Language (HU Berlin, Seminar 52 20004)
	Statistics for the Digital Humanities (NEH Institute 'Working with Text in a Digital Age', Tufts University)
Winter Term 2011/12	R II - Advanced Course in Statistical Methods with R (HU Berlin, Empirical Methods Workshop 2012)
Summer Term 2011	R Workshop, Binary Logistic Regression & Linear Mixed Effects Models (Centre for General Linguistics [ZAS], Berlin)
Winter Term 2010/11	Introduction to Linguistics (HU Berlin, Intro Course 52 20104)
	R II - Advanced Course in Statistical Methods with R (HU Berlin, Empirical Methods Workshop 2011)
Summer Term 2010	Statistical Methods in Linguistics (HU Berlin, Tutorial 52 20146)
	Corpus Linguistics (HU Berlin, Empirical Methods Workshop 2010)

Summer Term 2009	Introduction to Morphology (HU Berlin, Intro Course 52 20120)
	Corpus Linguistics (University of Oldenburg, block seminar together with Hagen Hirschmann)
Winter Term 2008/09	Parallel Corpora (HU Berlin, Seminar 52 20154)
Summer Term 2008	Introduction to Corpus Linguistics (HU Berlin, Tutorial 52 185)

Selected Publications (see website above for full list)

Monograph

Zeldes, Amir (2012), *Productivity in Argument Selection. From Morphology to Syntax.* (Trends in Linguistics: Studies and Monographs 260.) Berlin/New York: De Gruyter. xx, 282 pp.

Peer reviewed journal articles

Romary, Laurent, Zeldes, Amir & Zipser, Florian (to appear), <tiger2/> – Serialising the ISO SynAF Syntactic Object Model. *Language Resources and Evaluation*.

Zeldes, Amir (2013), Is Modern Hebrew Standard Average European? The View from European. *Linguistic Typology* 17(3), 439-470.

Zeldes, Amir (2013), Productive Argument Selection: Is Lexical Semantics enough? *Corpus Linguistics and Linguistic Theory* 9(2), 263–291.

Zeldes, Amir (2013), Komposition als Konstruktionsnetzwerk im fortgeschrittenen L2-Deutsch [=Compounding as a network of constructions in advanced L2 German]. *Zeitschrift für germanistische Linguistik* 41(2), 240–276 [=Journal of German Linguistics]. (in German)

Hirschmann, Hagen, Lüdeling, Anke & Zeldes, Amir (2012), Measuring and Coding Language Change: An Evolving Study in a Multi-Layer Corpus Architecture. *ACM Journal of Computing and Cultural Heritage (JOCCH)* 5(1), 4:1–4:16.

Chiarcos, Christian, Fiedler, Ines, Grubic, Mira, Hartmann, Katharina, Ritz, Julia, Schwarz, Anne, Zeldes, Amir & Zimmermann, Malte (2011), Information Structure in African Languages: Corpora and Tools. *Language Resources and Evaluation* 45(3), 361–374.

Crane, Gregory, Babeu, Alison, Bamman, David, Breuel, Thomas, Cerrato, Lisa, Deckers, Daniel, Lüdeling, Anke, Mimno, David, Singhal, Rashmi, Smith, David A. & Zeldes, Amir (2009), Classics in the Million Book Library. *Digital Humanities Quarterly* 3(1).

Petrova, Svetlana, Solf, Michael, Ritz, Julia, Chiarcos, Christian & Zeldes, Amir (2009), Building and Using a Richly Annotated Interlinear Diachronic Corpus: The Case of Old High German Tatian. *Traitement automatique des langues* [=Automatic Language Processing] 50(2), 47–71.

Lüdeling, Anke & Zeldes, Amir (2008), Three Views on Corpora: Corpus Linguistics, Literary Computing, and Computational Linguistics. *Jahrbuch für Computerphilologie* [=Yearbook of Computational Philology] 2007, 149–178.

Zeldes, Amir (2007), Abstracting Suffixes: A Morphophonemic Approach to Polish Morphological Analysis. *Zeitschrift für Sprachwissenschaft* [=German Journal of Linguistics] 26(2), 347–370.

Book chapters

Zeldes, Amir (to appear), The Case for Caseless Prepositional Constructions with *voller* in German. In: Boas, Hans C. & Ziem, Alexander (eds.), Constructional Approaches to Argument Structure in German. (Trends in Linguistics: Studies and Monographs.) Berlin: De Gruyter.

Lüderling, Anke, Ritz, Julia, Stede, Manfred & Zeldes, Amir (to appear). Corpus Linguistics and Information Structure Research. In: Féry, Caroline & Ishihara, Shinichiro (eds.), *The Oxford Handbook of Information Structure*. Oxford: Oxford University Press.

Zeldes, Amir & Kanbar, Ghazwan (2014), Arabisch und Hebräisch [=Arabic and Hebrew]. In: Krifka, Manfred, Błaszczak, Joanna, Leßmöllmann, Annette, Meinunger, André, Stiebels, Barbara, Tracy,

Rosemarie & Truckenbrodt, Hubert. *Das mehrsprachige Klassenzimmer* [=The Multilingual Classroom]. Heidelberg: Springer, 135–174.

Gaeta, Livio & Zeldes, Amir (2012), Deutsche Komposita zwischen Syntax und Morphologie: Ein korpusbasierter Ansatz [=German compounds between syntax and morphology. A corpus-based approach]. In: Gaeta, Livio & Schlücker, Barbara (eds.), *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte* [=German as a compounding-happy language: structural properties and system-related aspects]. (Linguistik - Impulse und Tendenzen.) Berlin: De Gruyter, 197–217. (in German)

Lüdeling, Anke, Hirschmann, Hagen & Zeldes, Amir (2011), Variationism and Underuse Statistics in the Analysis of the Development of Relative Clauses in German. In: Kawaguchi, Yuji, Minegishi, Makoto & Viereck, Wolfgang (eds.), *Corpus Analysis and Diachronic Linguistics*. (Tokyo University of Foreign Studies, Studies in Linguistics 3.) Amsterdam: John Benjamins, 37–57.

Conference papers

Hirschmann, Hagen, Lüdeling, Anke, Rehbein, Ines, Reznicek, Marc & Zeldes, Amir (2013), Underuse of Syntactic Categories in Falko. A Case Study on Modification. In: Granger, Sylviane and Meunier, Fanny (eds.), *20 Years of Learner Corpus Research. Looking Back, Moving Ahead*. (Corpora and Language in use.) Louvain: Presses universitaires de Louvain, 223–234.

Petrova, Svetlana & Zeldes, Amir (2012), How Exceptional is CP Recursion in Germanic OV Languages? Corpus-based Evidence from Middle Low German. In: *International Conference on Historical Corpora 2012*, December 6-9, Frankfurt, Germany.

Zeldes, Amir (2012), Novel Argument Realization: Semantic, Pragmatic and Conventional Productivity Effects. In: *Linguistic Evidence 2012. Empirical, Theoretical and Computational Perspectives*. Tübingen, 9–11 February 2012, 327–331

Krause, Thomas, Lüdeling, Anke, Odebrecht, Carolin & Zeldes, Amir (2012), Multiple Tokenizations in a Diachronic Corpus. In: *Exploring Ancient Languages through Corpora*, 14–16 June 2012, Oslo.

Zeldes, Amir (2011), On the Productivity and Variability of the Slots in German Comparative Correlative Constructions. In: Konopka, Marek, Kubczak, Jacqueline, Mair, Christian, Štícha, František & Waßner, Ulrich H. (eds.), *Grammar & Corpora. Third International Conference*. Tübingen: Narr, 429–449.

Krause, Thomas, Ritz, Julia, Zeldes, Amir & Zipser, Florian (2011), Topological Fields, Constituents and Coreference: A New Multi-layer Architecture for TüBa-D/Z. In: Hedeland, Hanna, Schmidt Thomas & Wörner, Kai (eds.), *Multilingual Resources and Multilingual Applications. Proceedings of the Conference of the German Society for Computational Linguistics and Language Technology (GSCL) 2011*. (Working Papers in Multilingualism 96.) Hamburg: Hamburg University, 259–262.

Zeldes, Amir (2009), Quantifying Constructional Productivity with Unseen Slot Members. In: *Proceedings of the NAACL HLT Workshop on Computational Approaches to Linguistic Creativity*, June 5, Boulder CO, 47–54.

Zeldes, Amir, Ritz, Julia, Lüdeling, Anke & Chiarcos, Christian (2009), ANNIS: A Search Tool for Multi-Layer Annotated Corpora. In: *Proceedings of Corpus Linguistics 2009*, July 20–23, Liverpool, UK.

Petrova, Svetlana, Chiarcos, Christian, Ritz, Julia & Zeldes, Amir (2009), The Tatian Corpus of Old High German: Information-Structural and Grammatical Annotation. In: *Proceedings of Corpus Linguistics 2009*, July 20–23, Liverpool, UK.

Zeldes, Amir, Lüdeling, Anke & Hirschmann, Hagen (2008), What's Hard? Quantitative Evidence for Difficult Constructions in German Learner Data. In: *Proceedings of Quantitative Investigations in Theoretical Linguistics 3 (QITL-3)*. Helsinki, Finland, 74–77.

Zeldes, Amir (2007), Machine Translation between Language Stages: Extracting Historical Grammar from a Parallel Diachronic Corpus of Polish. In: *Proceedings of Corpus Linguistics 2007*, July 27–30, Birmingham, UK.

Book reviews

Zeldes, Amir (to appear), Review of: Schmidt, Thomas & Wörner, Kai, "Multilingual Corpora and Multilingual Corpus Analysis". *Languages in Contrast*.

Zeldes, Amir (2010), Review of: McEnery, Tony, Xiao, Richard & Tono, Yukio, "Corpus-Based Language Studies: An Advanced Resource Book". *Zeitschrift für Rezensionen zur germanistischen Sprachwissenschaft* [=Review Journal of German Linguistics] 2(2), 268–274. (in German)

Academic Awards

8/2005 – 10/2006	DAAD Scholarship for Doctoral Candidates and Young Scientists (A/05/38058), German Academic Exchange Service
12/2004	Rector's Prize 2003/4, Hebrew University of Jerusalem, Israel
11/2004	Dean's List and Dean's Prize 2003/4, Hebrew University of Jerusalem, Israel
7/2004	National Commendation for Excellent Student Achievement, Committee of the University Presidents of Israel
3/2004	Rector's Prize 2002/3, Hebrew University of Jerusalem, Israel
3/2004	Dean's List and Dean's Prize 2002/3, Hebrew University of Jerusalem Jerusalem, Israel
10/2002	Scholarship for Excellent Academic Achievement, Hebrew University of Jerusalem Jerusalem, Israel

Work on funded projects

- *Coptic SCRIPTORIUM* – an NEH funded digital humanities project on resources for Sahidic Coptic. Co-directed with Prof. Caroline T. Schroeder, University of the Pacific.

- *KOMeT* – Korpuslinguistische Methoden für ePhilologie mit TEI – A BMBF funded junior researcher group on corpus linguistics methods in the digital humanities (under my direction)

- *RIDGES* – winner of two Google Digital Humanities Research Awards – Project on Register in Diachronic German Science, 16th–19th century German: http://korpling.german.hu-berlin.de/ridges/ (directed by Prof. Anke Lüdeling)

- *D1: Linguistic Database for Information Structure* – a project developing tools for multi-layer corpus search and visualization and annotation of information structure (ANNIS, see http://www.sfb632.uni-potsdam.de/~d1/annis/), as part of the Collaborative Research Center on Information Structure (project directors Prof. Anke Lüdeling and Prof. Manfred Stede, funded by the German Research Foundation DFG, current funding 2011–2015)

- *Textual Revision in Second Language Writing* – A joint project on English and German learner language in cooperation with City University of Hong Kong (together with Marc Reznicek, HU Berlin and Prof. John Lee, City University of Hong Kong, funded by DAAD, German Academic Exchange Service for 2012–2013)

Service and unfunded voluntary project work

- Editorial board of *Trends in Linguistics: Studies and Monographs* (De Gruyter Mouton, since June 2013)

- Member of CLARIN (http://www.clarin.eu/) Working Group 5.7 on Interoperability and Standards for language resources (since June 2010)

- Co-lead of <tiger2/>, a project developing an XML serialization of the ISO SynAF syntactic model for ISO/TC37/SC4: http://korpling.german.hu-berlin.de/tiger2/ (since October 2009)

- Co-organizer of the departmental reading circle on usage-based grammar models: http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/events-en/reading-circle-ubg/

- Counseling for the Arbeitskreis neue Erziehung (Association for New Education) Project Multilingual Letters for Parents: http://www.ane.de/homepage-english/letters-to-parents-media-of-parent-education.html (July 2009)

- Committee to revise the doctoral regulations of Philosophical Faculty II, HU Berlin (January 2009)

- Reviewing work for scholarly journals:
    - *ACM Journal on Computing and Cultural Heritage*
    - *Cognitive Science*
    - *Corpus Linguistics and Linguistic Theory*
    - *Interdisciplinary Studies in Information Structure*
    - *Languages in Contrast*
    - *Zeitschrift für germanistische Linguistik* [=Journal of German Linguistics]
    - *Zeitschrift für Rezensionen zur Germanistischen Sprachwissenschaft* [=Review Journal of German Linguistics]
- Reviewing for publishing houses (book proposals and chapters):
    - Cambridge University Press
    - De Gruyter Mouton

Conference and workshop organization

Organizing committee
- Volkswagen Stiftung Herrenhausen Conference on Big Data, Hanover 2015.
- Workshop *Coptic Resources in the Digital Age*, Humboldt University Berlin, 14 May 2013.
- Workshop on *Usage-Based Approaches to Morphology* at the 35th Annual Meeting of the German Society for Linguistics (DGfS), Potsdam University 13–15 March 2013.
- International conference on *Information Structure: Empirical Perspectives on Theory*. Potsdam University, 2–3 December 2011.
- International conference *Quantitative Investigations in Theoretical Linguistics 4* (QITL-4), Humboldt University Berlin, 28 February – 31 March 2011.
- *ANNIS/PAULA User and Developer Workshop* at the Annual Meeting of the German Society for Computational Linguistics (GSCL), Potsdam University, 2 October 2009.

Program committee / reviewer
- *Fall School in Computational Linguistics 2015*, German Society for Linguistics (DGfS), Special Interest Group on Computational Linguistics. Trier, 2015.
- *KONVENS2014*, 12th Conference on Natural Language Processing, Hildesheim, 8-10.10.2014.
- *Tenth International Tbilisi Symposium on Language, Logic and Computation*, Gudauri, Georgia, September 23-27, 2013.
- International conference *Quantitative Investigations in Theoretical Linguistics 5* (QITL-5), University of Leuven, 12–14 Sept. 2013.
- *Fall School in Computational Linguistics 2013*, German Society for Linguistics (DGfS), Special Interest Group on Computational Linguistics. Heidelberg, September, 2013.
- *KONVENS2012*, 11th Conference on Natural Language Processing, Vienna, 19-21.9.2012.
- LREC 2012, Workshop *Challenges in the Management of Large Corpora*, Istanbul, 22 May 2012.

- *Fall School in Computational Linguistics 2011*, German Society for Linguistics (DGfS), Special Interest Group on Computational Linguistics. Zurich, 29 August –9 September, 2011.

Membership in professional societies

- Deutsche Gesellschaft für Sprachwissenschaft (DGfS) – The German Society for Linguistics, and the Special Interest Group on Computational Linguistics (DGfS-CL) (since 2008)
- German Society for Computational Linguistics & Language Technology (GSCL) (since 2009)
- Association for Computational Linguistics (ACL, European chapter, since 2009)

Computer Skills

- **Operating systems:** Experience with both Windows and Linux (esp. Ubuntu) desktop and server environments.
- **Databases:** Extensive experience with PostgreSQL, SQL Server, MS Access, basic Oracle.
- **Programming languages:** Java, Visual Basic, JavaScript, PHP, Perl, R, basic knowledge of Python, Matlab.
- **Web design:** HTML/CSS, XML/XSLT, JSP/ASP, SVG, Graphviz, Photoshop, oXygen.
- **Office etc.:** Excellent knowledge of MS Office, including VBA API, basic LaTeX.

Languages

Native:	Hebrew, English
Near-native fluency:	German
Advanced knowledge:	Arabic, French, Italian, Japanese, Polish
Ancient languages (reading):	Ancient Egyptian, Church Slavonic, Classical Greek, Coptic, Sanskrit