

# Contents

<b>Preface</b> .....	9
<b>Martin Durrell:</b> ‘Representativeness’, ‘Bad Data’, and legitimate expectations. What can an electronic historical corpus tell us that we didn’t actually know already (and how)?.....	13
<b>Karin Donhauser:</b> Das Referenzkorpus Altdeutsch. Das Konzept, die Realisierung und die neuen Möglichkeiten .....	35
<b>Claudine Moulin / Iryna Gurevych / Natalia Filatkina / Richard Eckart de Castilho:</b> Analyzing formulaic patterns in historical corpora.....	51
<b>Roland Mittmann:</b> Automated quality control for the morphological annotation of the Old High German text corpus. Checking the manually adapted data using standardized inflectional forms.....	65
<b>Timothy Blaine Price:</b> Multi-faceted alignment. Toward automatic detection of textual similarity in Gospel-derived texts .....	77
<b>Gaye Detmold / Helmut Weiß:</b> Historical corpora and word formation. How to annotate a corpus to facilitate automatic analyses of noun-noun compounds.....	91
<b>Augustin Speyer:</b> Object order and the Thematic Hierarchy in older German .....	101
<b>Marco Coniglio / Eva Schlachter:</b> The properties of the Middle High German “Nachfeld”. Syntax, information structure, and linkage in discourse .....	125
<b>Stefanie Dipper / Julia Krasselt / Simone Schultz-Balluff:</b> Creating synopses of ‘parallel’ historical manuscripts and early prints. Alignment guidelines, evaluation, and applications.....	137
<b>Svetlana Petrova / Amir Zeldes:</b> How exceptional is CP recursion in Germanic OV languages? Corpus-based evidence from Middle Low German .....	151

<b>Alexander Geyken / Thomas Gloning:</b> A living text archive of 15 <sup>th</sup> -19 <sup>th</sup> -century German. Corpus strategies, technology, organization .....	165
<b>Christian Thomas / Frank Wiegand:</b> Making great work even better. Appraisal and digital curation of widely dispersed electronic textual resources (c. 15 <sup>th</sup> -19 <sup>th</sup> centuries) in CLARIN-D.....	181
<b>Bryan Jurish / Henriette Ast:</b> Using an alignment-based lexicon for canonicalization of historical text .....	197
<b>Armin Hoenen / Franziska Mader:</b> A new LMF schema application. An Austrian lexicon applied to the historical corpus of the writer Hugo von Hofmannsthal.....	209
<b>Thomas Efer / Jens Blecher / Gerhard Heyer:</b> Leipziger Rektoratsreden 1871-1933. Insights into six decades of scientific practice .....	229
<b>Stefania Degaetano-Ortlieb / Ekaterina Lapshinova-Koltunski / Elke Teich / Hannah Kermes:</b> Register contact: an exploration of recent linguistic trends in the scientific domain.....	241
<b>Esther Rinke / Svetlana Petrova:</b> The expression of thetic judgments in Older Germanic and Romance .....	255
<b>Richard Ingham:</b> Spoken and written register differentiation in pragmatic and semantic functions in two Anglo-Norman corpora.....	269
<b>Ana Paula Banza / Irene Rodrigues / José Saias / Filomena Gonçalves:</b> A historical linguistics corpus of Portuguese (16 <sup>th</sup> -19 <sup>th</sup> centuries) .....	281
<b>Natália Resende:</b> Testing the validity of translation universals for Brazilian Portuguese by employing comparable corpora and NLP techniques .....	291
<b>Jost Gippert / Manana Tandashvili:</b> Structuring a diachronic corpus. The Georgian National Corpus project.....	305
<b>Marina Beridze / Liana Lortkipanidze / David Nadaraia:</b> The Georgian Dialect Corpus: problems and prospects .....	323
<b>Claudia Schneider:</b> Integrating annotated ancient texts into databases. Technical remarks on a corpus of Indo-European languages tagged for information structure .....	335

<b>Giuseppe Abrami / Michael Freiberg / Paul Warner:</b> Managing and annotating historical multimodal corpora with the eHumanities desktop. An outline of the current state of the LOEWE project “Illustrations of Goethe’s Faust” .....	353
<b>Manuel Raaf:</b> A web-based application for editing manuscripts .....	365
<b>Gerhard Heyer / Volker Boehlke:</b> Text mining in the Humanities – A plea for research infrastructures.....	373