

Workshop on Processing Historical Language 2017

Held in conjunction with NoDaLiDa 2017

NEALT Proceedings Series Volume 32

Gothenburg, Sweden
22 May 2017

Editors:

Gerlof Bouma
Yvonne Adesam

ISBN: 978-1-5108-4336-3

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2017) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2017)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Contents

– *Invited Talk* –

Variance in Historical Data: How bad is it and how can we profit from it for historical linguistics? <i>Stefanie Dipper</i>	1
Improving POS Tagging in Old Spanish Using TEITOK <i>Maarten Janssen, Josep Ausensi, and Josep M. Fontana</i>	2
The Making of the Royal Society Corpus <i>Jörg Knappen, Stefan Fischer, Hannah Kermes, Elke Teich, and Peter Fankhauser</i>	7
Normalizing Medieval German Texts: from rules to deep learning <i>Natalia Korchagina</i>	12
Ambiguity in Semantically Related Word Substitutions: an investigation in historical Bible translations <i>Maria Moritz and Marco Büchler</i>	18
The Lemlat 3.0 Package for Morphological Analysis of Latin <i>Marco Passarotti, Marco Budassi, Eleonora Litta, and Paolo Ruffolo</i>	24
HistoBankVis: Detecting Language Change via Data Visualization <i>Christin Schätzle, Michael Hund, Frederik L. Dennig, Miriam Butt, and Daniel A. Keim</i>	32
Comparing Rule-based and SMT-based Spelling Normalisation for English Historical Texts <i>Gerold Schneider, Eva Pettersson, and Michael Percillier</i>	40
Data-driven Morphology and Sociolinguistics for Early Modern Dutch <i>Marijn Schraagen, Marjo van Koppen, and Feike Dietz</i>	47
Applying BLAST to Text Reuse Detection in Finnish Newspapers and Journals, 1771–1910 <i>Aleksi Vesanto, Asko Nivala, Heli Rantala, Tapio Salakoski, Hannu Salmi, and Filip Ginter</i>	54