

EACL 2012

**Second Workshop on Computational Linguistics and Writing
(CL&W 2012):
Linguistic and Cognitive Aspects of
Document Creation and Document Engineering**

Proceedings of the Workshop

April 23, 2012
Avignon, France

© 2012 The Association for Computational Linguistics

ISBN 978-1-937284-19-0

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

Introduction

Writing, whether professional, academic, or private, needs editors, input tools and display devices, and involves the coordination of cognitive, linguistic, and technical aspects. Most texts composed in the 21st century are probably created on electronic devices; people compose texts in word processors, text editors, content management systems, blogs, wikis, e-mail clients, and instant messaging applications. Texts are rendered and displayed on very small and very large screens, they are meant to be read by experts and laypersons, and they are supposed to be interactive and printable all at the same time.

The production of documents has been researched from various perspectives:

- Writing research has been concerned with text processing tools and cognitive processes since the 1970s. The current rise of new writing environments and genres (e.g., blogging), as well as new possibilities to observe text production in the workplace, has prompted new studies in this area of research.
- Document engineering is concerned with aspects of rendering and displaying textual and other resources for the creation, maintenance, and management of documents. Writers today use tools for layout design, collaborating with co-authors, and tracking changes in the production process with versioning systems—all of these are active research areas in document engineering.
- Computational linguistics has mostly been concerned with static or finished texts. There is now a growing need to explore how computational linguistics can support human text production and interactive text processing. Methods from natural language processing can also provide support for exploring data relevant for writing research (e.g., keystroke-logging data) and document engineering (e.g., tailoring documents to specific user needs).

CL&W 2010, held at NAACL 2010 in Los Angeles, was a successful workshop, offering researchers from different but related disciplines a platform for sharing findings and ideas. This follow-on Workshop on Computational Linguistics and Writing brings together researchers from the communities listed above to stimulate discussion and cooperation between these areas of research.

We received 9 submissions from both computational linguistics and writing researchers. After a rigorous review process we selected 6 papers for the workshop. We would like to thank the members of the Program Committee for their excellent work—the reviews were all very thorough, carefully written, and detailed, and helped the authors to improve their papers.

The papers included here present research that explores writing processes, text production, and document engineering principles as well as actual working systems that support novice and expert writers in one or more aspects when producing a document. We are pleased to present these papers in this volume.

We hope the work presented at CL&W 2012 will foster discussion and collaboration between researchers, bringing together expertise and interest from different but related fields.

Michael Piotrowski, Cerstin Mahlow, and Robert Dale

Organizers:

Michael Piotrowski, University of Zurich (Switzerland)
Cerstin Mahlow, University of Basel (Switzerland)
Robert Dale, Macquarie University (Australia)

Program Committee:

Gerd Bräuer, University of Education Freiburg (Germany)
Jill Burstein, ETS (USA)
Rickard Domeij, The Language Council of Sweden (Sweden)
Kevin Egan, University of Southern California (USA)
Caroline Hagège, Xerox Research Centre Europe (France)
Sofie Johansson Kokkinakis, University of Gothenburg (Sweden)
Ola Karlsson, The Language Council of Sweden (Sweden)
Ola Knutsson, KTH (Sweden)
Eva Lindgren, Umeå University (Sweden)
Aurélien Max, LIMSI (France)
Guido Nottbusch, University of Bielefeld (Germany)
Martin Reynaert, Tilburg University (The Netherlands)
Koenraad de Smedt, University of Bergen (Norway)
Sylvana Sofkova Hashemi, University West (Sweden)
Eric Wehrli, University of Geneva (Switzerland)
Carl Whithaus, UC Davis (USA)
Michael Zock, CNRS (France)

Table of Contents

<i>From character to word level: Enabling the linguistic analyses of Inputlog process data</i> Mariëlle Leijten, Lieve Macken, Veronique Hoste, Eric Van Horenbeeck and Luuk Van Waes	1
<i>From Drafting Guideline to Error Detection: Automating Style Checking for Legislative Texts</i> Stefan Höfler and Kyoko Sugisaki	8
<i>Summary of Research Based on My Reviewers & the Benefits of Aggregated, Crowd-Sourced Assessment</i> Joe Moxley	18
<i>Google Books N-gram Corpus used as a Grammar Checker</i> Rogelio Nazar and Irene Renau	23
<i>LELIE: a Tool dedicated to Procedure and Requirement Authoring (Demo paper)</i> Camille Albert, Flore Barcellini, Corinne Grosse and Patrick Saint-Dizier	31
<i>Focus Group on Computer Tools Used for Professional Writing and Preliminary Evaluation of Linguis-Tech</i> Marie-Josée Goulet and Annie Duplessis	35

Workshop Program

April 23, 2012

14:00 Opening

Session 1

14:15–14:40 *From character to word level: Enabling the linguistic analyses of Inputlog process data*

Mariëlle Leijten, Lieve Macken, Veronique Hoste, Eric Van Horenbeeck and Luuk Van Waes

14:40–15:05 *From Drafting Guideline to Error Detection: Automating Style Checking for Legislative Texts*

Stefan Höfler and Kyoko Sugisaki

15:05–15:30 *Summary of Research Based on My Reviewers & the Benefits of Aggregated, Crowd-Sourced Assessment*

Joe Moxley

15:30 Coffee Break

Session 2

16:00–16:25 *Google Books N-gram Corpus used as a Grammar Checker*

Rogelio Nazar and Irene Renau

16:25–16:50 *LELIE: a Tool dedicated to Procedure and Requirement Authoring (Demo paper)*

Camille Albert, Flore Barcellini, Corinne Grosse and Patrick Saint-Dizier

16:50–17:15 *Focus Group on Computer Tools Used for Professional Writing and Preliminary Evaluation of LinguisTech*

Marie-Josée Goulet and Annie Duplessis

17:30 Discussion and Closing

