

Shall We Play the Festschrift Game?

Diana Santos • Krister Lindén • Wanjiku Ng'ang'a
Editors

Shall We Play the Festschrift Game?

Essays on the Occasion of Lauri Carlson's
60th Birthday

Editors

Diana Santos
Department of Literature, Area Studies
and European Languages
University of Oslo
Oslo, Norway

Wanjiku Ng'ang'a
School of Computing & Informatics
University of Nairobi
Nairobi, Kenya

Krister Lindén
Department of Modern Languages
University of Helsinki
Helsinki, Finland

ISBN 978-3-642-30772-0

ISBN 978-3-642-30773-7 (eBook)

DOI 10.1007/978-3-642-30773-7

Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012946035

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Playing the Festschrift Game

1 Presentation

There are not many people who can be said to have influenced and impressed researchers in so many disparate areas and language-geographic fields as Lauri Carlson, as is evidenced in the present Festschrift.

Lauri's insight and acute linguistic sensitivity and linguistic rationality have spawned findings and research work in many areas, from non-standard etymology to hard-core formal linguistics, not forgetting computational areas such as parsing, terminological databases, and, last but not least, machine translation.

In addition to his renowned and widely acknowledged insights on tense and aspect, and their relationship with nominal quantification, and his ground-breaking work in dialogue using game-theoretic machinery, Lauri has in the last fifteen years as Professor of Language Theory and Translation Technology contributed immensely to areas such as translation, terminology and general applications of computational linguistics.

The three editors of the present volume have successfully performed doctoral studies under Lauri's supervision, and wish with this volume to pay tribute to his supervision and to his influence in matters associated with research, and scientific, linguistic and philosophical inquiry, as well as to his humanity and friendship as a definitely non-bossy supervisor.

Diana still recalls with gratitude the short periods either in Helsinki or in Lisbon, where discussions on tense and aspect matters took place:

The pertinence of Lauri's questions, questioning what at the time seemed unquestionable matters to me. The need to try to explain seemingly evident things in my understanding of my own native language was a sobering and extremely useful exercise. His genuine interest in language in general, and languages in particular, was also contagious, and the turning point in my PhD study—and possibly my whole research career—came about by Lauri buying a cheap edition of Steinbeck's *The Pearl* and its translation into Portuguese, and challenging me to prove that my translation rules were backed by real-world evidence. Naturally, as soon as we started looking at the texts themselves, a whole new world opened. Parallel corpora studies can now be said to constitute a well-established and productive field. Lauri does not even claim to have played a serious part in it. But he will clearly remain for me one of its founders, or rather one of its most important sources of inspiration.

After his graduation, Krister was involved in an English-to-Finnish Machine Translation project under the supervision of Lauri Carlson, who later also became the primary supervisor of his thesis:

His vast knowledge of both applied and formal linguistics combined with his down-to-earth remarks paved my way to linguistics. Later, Lauri again proved his worth in gold when he, with a few keystrokes, fleshed out a definition of synonymy and how to implement it in practice for our joint project on creating a Finnish wordnet, which is now the second largest wordnet in the world after the famous Princeton WordNet.

Given her non-linguistic background at the start of her PhD study, Wanjiku is hugely appreciative of Lauri's mentorship and patience that saw her acquire the requisite linguistic knowledge, that enabled her to work with Swahili:

Working with a then resource-scarce language required clever combination of resources and methods, and Lauri was very supportive of my lines of inquiry, a testament to his liberal thinking and open-mindedness, and I remember our meetings fondly. Our varied discussions on language, culture and religion helped me understand and appreciate the intricacies of language and the strengths and limitations of machine learning approaches to language processing. In addition, I benefited immensely from exposure to some of his networks—collaborative linkages which were crucial for my research, and others which persisted after my doctoral studies and which continue to shape my career to date.

2 Book Contents

Here, we provide a brief overview of the book's contents by describing, for each paper, the pertinent research questions and the points of contact with the other texts. This brief overview serves to harmonise the various contributions, and should in no way diminish the need to read each chapter individually.

The general nature of linguistic inquiry and how methods and empirical data should be tackled is the subject of Fred Karlsson's contribution, *Is There a Crisis in Generative Linguistics?*, an answer to a paper by Lauri on a similar subject. A quite original and interesting related subject is the way etymology is construed in different traditions, with David Nash's thought-provoking piece on language history in Australia, *"It's Etymology Captain, But Not As We Know It": 'Pump' in North Australia*.

Translation in History by Nicholas Ostler takes us on the challenging endeavor of making sense of translation diachronically and provides a new dimension to an ever-present activity in our times. Andrew Chesterman's chapter, *Catford Revisited*, provides a new view on Catford, an early theorist of translation who deserves to be rediscovered and reappraised, in his view. Diana Santos's contribution, *The Next Step for the Translation Network* delineates a computational program for studying human translations based on large corpora, as a direct follower to the proposal sketched in her dissertation.

An intriguing investigation on the concept of core vocabulary is provided by Lars Borin's *Core Vocabulary: A Useful But Mystical Concept in Some Kinds of Linguistics*, which provides thought-provoking material on some basic concepts in need

of rethinking. On a practical note, Krister Lindén, Jyrki Niemi and Mirka Hyvärinen provide an interesting evaluation of the Finnish wordnet and its coverage of the Finnish core vocabulary in *Extending and Updating the Finnish Wordnet*. As mentioned, Lauri Carlson was instrumental when creating the initial version of the Finnish wordnet. An exciting related subject, in a different frame of mind, is Janet Pierrehumbert's work on *Burstiness of Verbs and Derived Nouns*, which merges formal semantics and statistical insights to Lauri's probably dearest subject of all, discourse.

Creating cutting-edge linguistic-computational resources, both for Finnish and for other languages, is the subject of *Outsourcing Parsebanking: The FinnTreeBank Project* by Atro Voutilainen, Tanja Purtonen and Kristiina Muhonen, where the authors suggest and demonstrate a novel way for developing large treebanks. In order to assist the treebanking effort, Anssi Yli-Jyrä's article *On Dependency Analysis via Contractions and Weighted FSTs* deals with an improved parsing technique for dependency parsers, presenting a system which is a radical improvement on his dissertation work, of which Lauri was one of the supervisors.

New analyses of linguistic phenomena round up this widely varied book: Patrick McConvell and Jane Simpson discuss some intriguing features of case marking in *Fictive Motion Down Under: The Locative-Allative Case Alternation in Some Australian Indigenous Languages*, describing several forms to mark the different location of events and of participants, and proposing a grammaticalisation path, while also comparing with parallel phenomena in Finnic. Aet Lees studies the expression of obligation and need in this last language family using Bible translations, in her chapter *Necessive Expressions in Finnic Bible Translations*, which handles the subject diachronically.

Wanjiku N'gang'a's *Building Swahili Resource Grammars for the Grammatical Framework*, in addition to the theoretical part of resource grammar writing, reports on an aid for visually-impaired mobile phone users that has been developed using the Grammatical Framework. Her work builds on the Grammatical Framework of Aarne Ranta, whose chapter *On the Syntax and Translation of Finnish Discourse Clitics* not only discusses Finnish discourse clitics but also describes a Finnish–English machine translation system that handles this interesting issue in Finnish grammar. In fact, it should be mentioned that this last chapter achieves the highest number of connections to Lauri's main interests: discourse, translation, formal analysis, and computation.

3 Contributions

In addition to the authors, a number of people have also contributed to this book in various capacities. Maria Vilkuna and Antti Arppe helped in reviewing the contributions. Jyrki Niemi undertook the job of converting all the non- \LaTeX contributions to \LaTeX and unifying the ones provided in \LaTeX . Finally, we are grateful to Seija Carlson for her help and for keeping it all a secret!

We would also like to acknowledge Lauri's friends and colleagues who would have wanted to write a chapter but did not manage to do so—either because we were unable to contact them and they consequently were unaware of this noble initiative, or due to pressing time constraints.

We are also grateful to all the authors, who willingly participated in the cross-review process, providing valuable feedback which has undoubtedly resulted in the high quality and readability of this book. We as editors are therefore extremely grateful for this added value.

Acknowledgements Diana Santos acknowledges the support of the University of Oslo in preparing this volume, and Krister Lindén would like to thank the FINCLARIN infrastructure project at the University of Helsinki.

Oslo
Helsinki
Nairobi

Diana Santos
Krister Lindén
Wanjiku Ng'ang'a

Contents

Is There a Crisis in Generative Linguistics?	1
Fred Karlsson	
“It’s Etymology Captain, but Not as We Know It”: <i>Pump</i> in North Australia	7
David Nash	
Translation in History	19
Nicholas Ostler	
Catford Revisited	25
Andrew Chesterman	
The Next Step for the Translation Network	35
Diana Santos	
Core Vocabulary: A Useful But Mystical Concept in Some Kinds of Linguistics	53
Lars Borin	
Extending and Updating the Finnish Wordnet	67
Krister Lindén, Jyrki Niemi, and Mirka Hyvärinen	
Burstiness of Verbs and Derived Nouns	99
Janet B. Pierrehumbert	
Outsourcing Parsebanking: The FinnTreeBank Project	117
Atro Voutilainen, Tanja Purtonen, and Kristiina Muhonen	
On Dependency Analysis via Contractions and Weighted FSTs	133
Anssi Yli-Jyrä	
Fictive Motion Down Under: The Locative-Allative Case Alternation in Some Australian Indigenous Languages	159
Patrick McConvell and Jane Simpson	

Necessive Expressions in Finnic Bible Translations	181
Aet Lees	
Building Swahili Resource Grammars for the Grammatical Framework .	215
Wanjiku Ng'ang'a	
On the Syntax and Translation of Finnish Discourse Clitics	227
Aarne Ranta	

Contributors

Lars Borin Språkbanken, Department of Swedish, University of Gothenburg, Gothenburg, Sweden

Andrew Chesterman Translation Studies, Department of Modern Languages, University of Helsinki, Helsinki, Finland

Mirka Hyvärinen Department of Modern Languages, University of Helsinki, Helsinki, Finland

Fred Karlsson General Linguistics, Department of Modern Languages, University of Helsinki, Helsinki, Finland

Aet Lees Department of Linguistics, University of Sydney, Sydney, Australia

Krister Lindén Department of Modern Languages, University of Helsinki, Helsinki, Finland

Patrick McConvell School of Language Studies, Australian National University, Canberra, Australia

Kristiina Muhonen Department of Modern Languages, University of Helsinki, Helsinki, Finland

David Nash School of Language Studies, The Australian National University, Canberra, Australia

Wanjiku Ng'ang'a School of Computing and Informatics, University of Nairobi, Nairobi, Kenya

Jyrki Niemi Department of Modern Languages, University of Helsinki, Helsinki, Finland

Nicholas Ostler Foundation for Endangered Languages, Bath, England, UK; School of Oriental and African Studies, London, England, UK

Janet B. Pierrehumbert Department of Linguistics, and Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL, USA

Tanja Purtonen Department of Modern Languages, University of Helsinki, Helsinki, Finland

Aarne Ranta University of Gothenburg, Gothenburg, Sweden

Diana Santos University of Oslo, Oslo, Norway

Jane Simpson School of Language Studies, Australian National University, Canberra, Australia

Atro Voutilainen Department of Modern Languages, University of Helsinki, Helsinki, Finland

Anssi Yli-Jyrä Department of Modern Languages, University of Helsinki, Helsinki, Finland