

Lecture Notes in Artificial Intelligence 5509

Edited by R. Goebel, J. Siekmann, and W. Wahlster

Subseries of Lecture Notes in Computer Science

Michael Kipp Jean-Claude Martin
Patrizia Paggio Dirk Heylen (Eds.)

Multimodal Corpora

From Models of Natural Interaction
to Systems and Applications

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Michael Kipp
Deutsches Forschungszentrum für künstliche Intelligenz (DFKI)
Campus D3.2, 66123 Saarbrücken, Germany
E-mail: kipp@dfki.de

Jean-Claude Martin
Laboratoire d'Informatique pour la Mécanique
et les Sciences de l'Ingénieur (LIMSI-CNRS)
BP 133, 91403 Orsay Cedex, France
E-mail: martin@limsi.fr

Patrizia Paggio
University of Copenhagen, Faculty of Humanities
Centre for Language Technology
Njalsgade 140-142, 2300 Copenhagen, Denmark
E-mail: paggio@hum.ku.dk

Dirk Heylen
University of Twente, Computer Science, Human Media Interaction
PO Box 217, 7500 AE Enschede, The Netherlands
E-mail: d.k.j.heylen@ewi.utwente.nl

Library of Congress Control Number: 2009935170

CR Subject Classification (1998): I.5, I.4, H.5.2, I.3, I.2, J.5, H.3.1

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-642-04792-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-04792-3 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

springer.com

© Springer-Verlag Berlin Heidelberg 2009
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12747986 06/3180 5 4 3 2 1 0

Preface

The book presents a cross-section of state-of-the-art research on multimodal corpora, a highly interdisciplinary area that is a prerequisite for various specialized disciplines. A number of the papers included are revised and expanded versions of papers accepted to the International Workshop on Multimodal Corpora: From Models of Natural Interaction to Systems and Applications, held in conjunction with the 6th International Conference for Language Resources and Evaluation (LREC) on May 27, 2008, in Marrakech, Morocco. This international workshop series started in 2000 and has since then grown into a regular satellite event of the bi-annual LREC conference, attracting researchers from fields as diverse as psychology, artificial intelligence, robotics, signal processing, computational linguistics and human-computer interaction. To complement the selected papers from the 2008 workshop, we invited well-known researchers from corpus collection initiatives to contribute to this volume. We were able to obtain seven invited research articles, including contributions from major international multimodal corpus projects like AMI and SmartWeb, which complement the six selected workshop contributions. All papers underwent a special review process for this volume, resulting in significant revisions and extensions based on the experts' advice. While we were pleased that the 2006 edition of the workshop resulted in a special issue of the *Journal of Language Resources and Evaluation*, published in 2007, we felt that this was the time for another major publication, given not only the rapid progress and increased interest in this research area but especially in order to acknowledge the difficulty of disseminating results across discipline borders. The Springer LNAI series is the perfect platform for doing so. We also created the website www.multimodal-corpora.org to serve as a permanent information hub for the workshop series, including links to research in multimodality and a mailing list for the announcement of future events.

Empirically based research in multimodality can be viewed as a work pipeline. It starts with collecting video data that are relevant for a specific research question (e.g., how do people express emotions using gestures?), either by filming interactions in a controlled (both technically and scenario-wise) setup or by taking existing material from TV or the movies. These data must then be enriched by human coders, who annotate the corpus with concepts only human competence can recognize and apply (if possible supported by automatic devices such as trackers), and who also explore the potential interconnections between modalities (face, gesture, speech, posture etc.). Since these data will potentially be distributed to a wide spectrum of consumers for research and development, the human-encoded data must be validated for consistency using quantitative measures. Systematic analysis of the observed multimodal behaviors is the central characteristic of multimodal corpora. This allows for the subsequent use of the corpora in creating analytic or generative models. For analysis and

modeling tasks, exploratory qualitative and quantitative analysis tools are needed for browsing, viewing, extraction and modeling. Finally, since multimodal interaction is always embedded in a social setting, the intercultural dimension must be taken into account. This means investigating how multimodal interactions change across cultures with different social protocols and possibly different standards and norms.

In this volume, we have carefully selected representative contributions for each step in this pipeline. Accordingly, we have organized the papers along this pipeline and, for the sake of clarity, not assembled all invited papers in one section, but freely interspersed papers from both categories according to their contents. For the topic of collection and distribution of corpora, we have two invited contributions about the SmartWeb corpus (Schiel) and the IFADV corpus (van Son et al.). In the area of coding and analyzing corpora, we invited Blache et al. for an account of the ToMA project. Additionally, three extended workshop papers deal with the multimodal behavior of children (Colletta et al.), politicians (Poggi and Vincze), and multiparty meetings (Matsusaka et al.). Two further included workshop papers deal with the important topic of validating multimodal corpora: Cavicchio and Poesio discuss how to assess coding scheme reliability and Reidsma et al. discuss agreement scores in conjunction with contextual information as part of the AMI project. As multimodal behavior must often be differentiated with regard to the intercultural dimension, we invited Rehm et al. to report on standardized video recordings for this purpose. This is complemented by an analysis of intercultural differences in information and communication technology by Allwood and Ahlsén. In an era of growing availability of large resources, tools for browsing, coding and exchanging data are becoming ever more important. We have therefore included three contributions that present existing tools for visualizing conversational-speech interaction (Campbell), for accessing a large multimodal corpus (Popescu-Belis et al.), and for exchanging data between different special purpose tools (Schmidt et al.).

Overall, we hope that this “overview by example” volume will provide a unique opportunity for researchers of all related disciplines to gain insight into this active research area where empiricism meets application, and the humanities meet technology. We would like to especially thank all reviewers who contributed their precious time to the excellent quality of the papers.

July 2009

Michael Kipp
Jean-Claude Martin
Patrizia Paggio
Dirk Heylen

Organization

Referees

Jan Alexandersson
Jens Allwood
Elisabeth Ahlsén
Elisabeth André
Gerard Bailly
Stéphanie Buisine
Susanne Burger
Loredana Cerrato
Piero Cosi
Morena Danieli

Nicolas Ech Chafai
John Glauert
Kostas Karpouzis
Alfred Kranstedt
Peter Kuehnlein
Daniel Loehr
Maurizio Mancini
Costanza Navarretta
Catherine Pelachaud
Fabio Pianesi

Isabella Poggi
Matthias Rehm
Laurent Romary
Thomas Schmidt
Rob van Son
Ielka van der Sluis
Rainer Stiefelhagen
Peter Wittenburg
Massimo Zancanaro

Table of Contents

Collecting and Distributing Multimodal Corpora

- The SmartWeb Corpora: Multimodal Access to the Web in Natural Environments 1
Florian Schiel
- Promoting *free* Dialog Video Corpora: The IFADV Corpus Example 18
R.J.J.H. van Son, Wieneke Wesseling, Eric Sanders, and Henk van den Heuvel

Coding and Analyzing Multimodal Corpora

- Creating and Exploiting Multimodal Annotated Corpora: The ToMA Project 38
Philippe Blache, Roxane Bertrand, and Gaëlle Ferré
- Multi-track Annotation of Child Language and Gestures 54
Jean-Marc Colletta, Ramona N. Kunene, Aurélie Venouil, Virginie Kaufmann, and Jean-Pascal Simon
- Gesture, Gaze and Persuasive Strategies in Political Discourse 73
Isabella Poggi and Laura Vincze
- Unsupervised Clustering in Multimodal Multiparty Meeting Analysis 93
Yosuke Matsusaka, Yasuhiro Katagiri, Masato Ishizaki, and Mika Enomoto

Validating Multimodal Corpora

- Multimodal Corpora Annotation: Validation Methods to Assess Coding Scheme Reliability 109
Federica Cavicchio and Massimo Poesio
- On the Contextual Analysis of Agreement Scores 122
Dennis Reidsma, Dirk Heylen, and Rieks op den Akker

The Intercultural Dimension of Multimodal Corpora

- Creating Standardized Video Recordings of Multimodal Interactions across Cultures 138
Matthias Rehm, Elisabeth André, Nikolaus Bee, Birgit Endrass, Michael Wissner, Yukiko Nakano, Afia Akhter Lipi, Toyoaki Nishida, and Hung-Hsuan Huang

Multimodal Intercultural Information and Communication
Technology – A Framework for Designing and Evaluating Multimodal
Intercultural Communicators 160
Jens Allwood and Elisabeth Ahlsén

**Tools for Browsing, Coding and Exchanging Data in
Multimodal Corpora**

Tools and Resources for Visualising Conversational-Speech
Interaction 176
Nick Campbell

Accessing a Large Multimodal Corpus Using an Automatic Content
Linking Device 189
*Andrei Popescu-Belis, Jean Carletta, Jonathan Kilgour, and
Peter Poller*

An Exchange Format for Multimodal Annotations 207
*Thomas Schmidt, Susan Duncan, Oliver Ehmer, Jeffrey Hoyt,
Michael Kipp, Dan Loehr, Magnus Magnusson, Travis Rose, and
Han Sloetjes*

Author Index 223