Steve Renals
Samy Bengio (Eds.)

LNCS 3869

Machine Learning for Multimodal Interaction

Second International Workshop, MLMI 2005
Edinburgh, UK, July 2005
Revised Selected Papers

TP181-53
M149.2
2005

Steve Renals Samy Bengio (Eds.)

Machine Learning for Multimodal Interaction

Second International Workshop, MLMI 2005
Edinburgh, UK, July 11-13, 2005
Revised Selected Papers

Springer

Volume Editors

Steve Renals
University of Edinburgh, Centre for Speech Technology Research
2 Buccleuch Place, Edinburgh EH8 9LW, UK
E-mail: s.renals@ed.ac.uk

Samy Bengio
IDIAP Research Institute
Rue du Simplon 4, Case Postale 592, 1920 Martigny, Switzerland
E-mail: bengio@idiap.ch

Library of Congress Control Number: 2006920577

CR Subject Classification (1998): H.5.2-3, H.5, I.2.6, I.2.10, I.2, I.7, K.4, I.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-540-32549-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-32549-9 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 is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2006
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11677482 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3869

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK

Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler
University of Surrey, Guildford, UK

Jon M. Kleinberg
Cornell University, Ithaca, NY, USA

Friedemann Mattern
ETH Zurich, Switzerland

John C. Mitchell
Stanford University, CA, USA

Moni Naor
Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz
University of Bern, Switzerland

C. Pandu Rangan
Indian Institute of Technology, Madras, India

Bernhard Steffen
University of Dortmund, Germany

Madhu Sudan
Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos
New York University, NY, USA

Doug Tygar
University of California, Berkeley, CA, USA

Moshe Y. Vardi
Rice University, Houston, TX, USA

Gerhard Weikum
Max-Planck Institute of Computer Science, Saarbruecken, Germany

Lecture Notes in Computer Science

For information about Vols. 1–3779

please contact your bookseller or Springer

Vol. 3884: B. Durand, W. Thomas (Eds.), STACS 2006. XIV, 714 pages. 2006.

Vol. 3879: T. Erlebach, G. Persinao (Eds.), Approximation and Online Algorithms. X, 349 pages. 2006.

Vol. 3878: A. Gelbukh (Ed.), Computational Linguistics and Intelligent Text Processing. XVII, 589 pages. 2006.

Vol. 3874: R. Missaoui, J. Schmidt (Eds.), Formal Concept Analysis. X, 309 pages. 2006. (Sublibrary LNAI).

Vol. 3872: H. Bunke, A. L. Spitz (Eds.), Document Analysis Systems VII. XIII, 630 pages. 2006.

Vol. 3870: S. Spaccapietra, P. Atzeni, W.W. Chu, T. Catarci, K.P. Sycara (Eds.), Journal on Data Semantics V. XIII, 237 pages. 2006.

Vol. 3869: S. Renals, S. Bengio (Eds.), Machine Learning for Multimodal Interaction. XIII, 490 pages. 2006.

Vol. 3868: K. Römer, H. Karl, F. Mattern (Eds.), Wireless Sensor Networks. XI, 342 pages. 2006.

Vol. 3863: M. Kohlhase (Ed.), Mathematical Knowledge Management. XI, 405 pages. 2006. (Sublibrary LNAI).

Vol. 3861: J. Dix, S.J. Hegner (Eds.), Foundations of Information and Knowledge Systems. X, 331 pages. 2006.

Vol. 3860: D. Pointcheval (Ed.), Topics in Cryptology – CT-RSA 2006. XI, 365 pages. 2006.

Vol. 3858: A. Valdes, D. Zamboni (Eds.), Recent Advances in Intrusion Detection. X, 351 pages. 2006.

Vol. 3857: M. Fossorier, H. Imai, S. Lin, A. Poli (Eds.), Applied Algebra, Algebraic Algorithms and Error-Correcting Codes. XI, 350 pages. 2006.

Vol. 3855: E. A. Emerson, K.S. Namjoshi (Eds.), Verification, Model Checking, and Abstract Interpretation. XI, 443 pages. 2005.

Vol. 3853: A.J. Ijspeert, T. Masuzawa, S. Kusumoto (Eds.), Biologically Inspired Approaches to Advanced Information Technology. XIV, 388 pages. 2006.

Vol. 3852: P.J. Narayanan, S.K. Nayar, H.-Y. Shum (Eds.), Computer Vision - ACCV 2006, Part II. XXXI, 977 pages. 2005.

Vol. 3851: P.J. Narayanan, S.K. Nayar, H.-Y. Shum (Eds.), Computer Vision - ACCV 2006, Part I. XXXI, 973 pages. 2006.

Vol. 3850: R. Freund, G. Păun, G. Rozenberg, A. Salomaa (Eds.), Membrane Computing. IX, 371 pages. 2006.

Vol. 3848: J.-F. Boulicaut, L. De Raedt, H. Mannila (Eds.), Constraint-Based Mining and Inductive Databases. X, 401 pages. 2006. (Sublibrary LNAI).

Vol. 3847: K.P. Jantke, A. Lunzer, N. Spyratos, Y. Tanaka (Eds.), Federation over the Web. X, 215 pages. 2006. (Sublibrary LNAI).

Vol. 3846: H. J. van den Herik, Y. Björnsson, N.S. Netanyahu (Eds.), Computers and Games. XIV, 333 pages. 2006.

Vol. 3844: J.-M. Bruel (Ed.), Satellite Events at the MoDELS 2005 Conference. XIII, 360 pages. 2006.

Vol. 3843: P. Healy, N.S. Nikolov (Eds.), Graph Drawing. XVII, 536 pages. 2006.

Vol. 3842: H.T. Shen, J. Li, M. Li, J. Ni, W. Wang (Eds.), Advanced Web and Network Technologies, and Applications. XXVII, 1057 pages. 2006.

Vol. 3841: X. Zhou, J. Li, H.T. Shen, M. Kitsuregawa, Y. Zhang (Eds.), Frontiers of WWW Research and Development - APWeb 2006. XXIV, 1223 pages. 2006.

Vol. 3840: M. Li, B. Boehm, L.J. Osterweil (Eds.), Unifying the Software Process Spectrum. XVI, 522 pages. 2006.

Vol. 3839: J.-C. Filliâtre, C. Paulin-Mohring, B. Werner (Eds.), Types for Proofs and Programs. VIII, 275 pages. 2006.

Vol. 3838: A. Middeldorp, V. van Oostrom, F. van Raamsdonk, R. de Vrijer (Eds.), Processes, Terms and Cycles: Steps on the Road to Infinity. XVIII, 639 pages. 2005.

Vol. 3837: K. Cho, P. Jacquet (Eds.), Technologies for Advanced Heterogeneous Networks. IX, 307 pages. 2005.

Vol. 3836: J.-M. Pierson (Ed.), Data Management in Grids. X, 143 pages. 2006.

Vol. 3835: G. Sutcliffe, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIV, 744 pages. 2005. (Sublibrary LNAI).

Vol. 3834: D.G. Feitelson, E. Frachtenberg, L. Rudolph, U. Schwiegelshohn (Eds.), Job Scheduling Strategies for Parallel Processing. VIII, 283 pages. 2005.

Vol. 3833: K.-J. Li, C. Vangenot (Eds.), Web and Wireless Geographical Information Systems. XI, 309 pages. 2005.

Vol. 3832: D. Zhang, A.K. Jain (Eds.), Advances in Biometrics. XX, 796 pages. 2005.

Vol. 3831: J. Wiedermann, G. Tel, J. Pokorný, M. Bieliková, J. Štuller (Eds.), SOFSEM 2006: Theory and Practice of Computer Science. XV, 576 pages. 2006.

Vol. 3829: P. Pettersson, W. Yi (Eds.), Formal Modeling and Analysis of Timed Systems. IX, 305 pages. 2005.

Vol. 3828: X. Deng, Y. Ye (Eds.), Internet and Network Economics. XVII, 1106 pages. 2005.

Vol. 3827: X. Deng, D.-Z. Du (Eds.), Algorithms and Computation. XX, 1190 pages. 2005.

Vol. 3826: B. Benatallah, F. Casati, P. Traverso (Eds.), Service-Oriented Computing - ICSOC 2005. XVIII, 597 pages. 2005.

此为试读,需要完整PDF请访问: www.ertongbook.com

Vol. 3824: L.T. Yang, M. Amamiya, Z. Liu, M. Guo, F.J. Rammig (Eds.), Embedded and Ubiquitous Computing – EUC 2005. XXIII, 1204 pages. 2005.

Vol. 3823: T. Enokido, L. Yan, B. Xiao, D. Kim, Y. Dai, L.T. Yang (Eds.), Embedded and Ubiquitous Computing – EUC 2005 Workshops. XXXII, 1317 pages. 2005.

Vol. 3822: D. Feng, D. Lin, M. Yung (Eds.), Information Security and Cryptology. XII, 420 pages. 2005.

Vol. 3821: R. Ramanujam, S. Sen (Eds.), FSTTCS 2005: Foundations of Software Technology and Theoretical Computer Science. XIV, 566 pages. 2005.

Vol. 3820: L.T. Yang, X.-s. Zhou, W. Zhao, Z. Wu, Y. Zhu, M. Lin (Eds.), Embedded Software and Systems. XXVIII, 779 pages. 2005.

Vol. 3819: P. Van Hentenryck (Ed.), Practical Aspects of Declarative Languages. X, 231 pages. 2005.

Vol. 3818: S. Grumbach, L. Sui, V. Vianu (Eds.), Advances in Computer Science – ASIAN 2005. XIII, 294 pages. 2005.

Vol. 3817: M. Faundez-Zanuy, L. Janer, A. Esposito, A. Satue-Villar, J. Roure, V. Espinosa-Duro (Eds.), Nonlinear Analyses and Algorithms for Speech Processing. XII, 380 pages. 2006. (Sublibrary LNAI).

Vol. 3816: G. Chakraborty (Ed.), Distributed Computing and Internet Technology. XXI, 606 pages. 2005.

Vol. 3815: E.A. Fox, E.J. Neuhold, P. Premsmit, V. Wuwongse (Eds.), Digital Libraries: Implementing Strategies and Sharing Experiences. XVII, 529 pages. 2005.

Vol. 3814: M. Maybury, O. Stock, W. Wahlster (Eds.), Intelligent Technologies for Interactive Entertainment. XV, 342 pages. 2005. (Sublibrary LNAI).

Vol. 3813: R. Molva, G. Tsudik, D. Westhoff (Eds.), Security and Privacy in Ad-hoc and Sensor Networks. VIII, 219 pages. 2005.

Vol. 3811: C. Bussler, M.-C. Shan (Eds.), Technologies for E-Services. VIII, 127 pages. 2006.

Vol. 3810: Y.G. Desmedt, H. Wang, Y. Mu, Y. Li (Eds.), Cryptology and Network Security. XI, 349 pages. 2005.

Vol. 3809: S. Zhang, R. Jarvis (Eds.), AI 2005: Advances in Artificial Intelligence. XXVII, 1344 pages. 2005. (Sublibrary LNAI).

Vol. 3808: C. Bento, A. Cardoso, G. Dias (Eds.), Progress in Artificial Intelligence. XVIII, 704 pages. 2005. (Sublibrary LNAI).

Vol. 3807: M. Dean, Y. Guo, W. Jun, R. Kaschek, S. Krishnaswamy, Z. Pan, Q.Z. Sheng (Eds.), Web Information Systems Engineering – WISE 2005 Workshops. XV, 275 pages. 2005.

Vol. 3806: A.H. H. Ngu, M. Kitsuregawa, E.J. Neuhold, J.-Y. Chung, Q.Z. Sheng (Eds.), Web Information Systems Engineering – WISE 2005. XXI, 771 pages. 2005.

Vol. 3805: G. Subsol (Ed.), Virtual Storytelling. XII, 289 pages. 2005.

Vol. 3804: G. Bebis, R. Boyle, D. Koracin, B. Parvin (Eds.), Advances in Visual Computing. XX, 755 pages. 2005.

Vol. 3803: S. Jajodia, C. Mazumdar (Eds.), Information Systems Security. XI, 342 pages. 2005.

Vol. 3802: Y. Hao, J. Liu, Y.-P. Wang, Y.-m. Cheung, H. Yin, L. Jiao, J. Ma, Y.-C. Jiao (Eds.), Computational Intelligence and Security, Part II. XLII, 1166 pages. 2005. (Sublibrary LNAI).

Vol. 3801: Y. Hao, J. Liu, Y.-P. Wang, Y.-m. Cheung, H. Yin, L. Jiao, J. Ma, Y.-C. Jiao (Eds.), Computational Intelligence and Security, Part I. XLI, 1122 pages. 2005. (Sublibrary LNAI).

Vol. 3799: M. A. Rodríguez, I.F. Cruz, S. Levashkin, M.J. Egenhofer (Eds.), GeoSpatial Semantics. X, 259 pages. 2005.

Vol. 3798: A. Dearle, S. Eisenbach (Eds.), Component Deployment. X, 197 pages. 2005.

Vol. 3797: S. Maitra, C. E. V. Madhavan, R. Venkatesan (Eds.), Progress in Cryptology - INDOCRYPT 2005. XIV, 417 pages. 2005.

Vol. 3796: N.P. Smart (Ed.), Cryptography and Coding. XI, 461 pages. 2005.

Vol. 3795: H. Zhuge, G.C. Fox (Eds.), Grid and Cooperative Computing - GCC 2005. XXI, 1203 pages. 2005.

Vol. 3794: X. Jia, J. Wu, Y. He (Eds.), Mobile Ad-hoc and Sensor Networks. XX, 1136 pages. 2005.

Vol. 3793: T. Conte, N. Navarro, W.-m.W. Hwu, M. Valero, T. Ungerer (Eds.), High Performance Embedded Architectures and Compilers. XIII, 317 pages. 2005.

Vol. 3792: I. Richardson, P. Abrahamsson, R. Messnarz (Eds.), Software Process Improvement. VIII, 215 pages. 2005.

Vol. 3791: A. Adi, S. Stoutenburg, S. Tabet (Eds.), Rules and Rule Markup Languages for the Semantic Web. X, 225 pages. 2005.

Vol. 3790: G. Alonso (Ed.), Middleware 2005. XIII, 443 pages. 2005.

Vol. 3789: A. Gelbukh, Á. de Albornoz, H. Terashima-Marín (Eds.), MICAI 2005: Advances in Artificial Intelligence. XXVI, 1198 pages. 2005. (Sublibrary LNAI).

Vol. 3788: B. Roy (Ed.), Advances in Cryptology - ASIACRYPT 2005. XIV, 703 pages. 2005.

Vol. 3787: D. Kratsch (Ed.), Graph-Theoretic Concepts in Computer Science. XIV, 470 pages. 2005.

Vol. 3786: J. Song, T. Kwon, M. Yung (Eds.), Information Security Applications. XI, 378 pages. 2006.

Vol. 3785: K.-K. Lau, R. Banach (Eds.), Formal Methods and Software Engineering. XIV, 496 pages. 2005.

Vol. 3784: J. Tao, T. Tan, R.W. Picard (Eds.), Affective Computing and Intelligent Interaction. XIX, 1008 pages. 2005.

Vol. 3783: S. Qing, W. Mao, J. Lopez, G. Wang (Eds.), Information and Communications Security. XIV, 492 pages. 2005.

Vol. 3782: K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick, T.R. Roth-Berghofer (Eds.), Professional Knowledge Management. XXIII, 739 pages. 2005. (Sublibrary LNAI).

Vol. 3781: S.Z. Li, Z. Sun, T. Tan, S. Pankanti, G. Chollet, D. Zhang (Eds.), Advances in Biometric Person Authentication. XI, 250 pages. 2005.

Vol. 3780: K. Yi (Ed.), Programming Languages and Systems. XI, 435 pages. 2005.

¥547.00元

Preface

This book contains a selection of refereed papers presented at the Second Workshop on Machine Learning for Multimodal Interaction (MLMI 2005), held in Edinburgh, Scotland, during 11–13 July 2005.

The workshop was organized and sponsored jointly by two European integrated projects, three European Networks of Excellence and a Swiss national research network:

- AMI, Augmented Multiparty Interaction, http://www.amiproject.org/
- CHIL, Computers in the Human Interaction Loop, http://chil.server.de/
- HUMAINE, Human–Machine Interaction Network on Emotion, http://emotion-research.net/
- PASCAL, Pattern Analysis, Statistical Modeling and Computational Learning, http://www.pascal-network.org/
- SIMILAR, human–machine interfaces similar to human–human communication, http://www.similar.cc/
- IM2, Interactive Multimodal Information Management, http://www.im2.ch/

In addition to the main workshop, MLMI 2005 hosted the NIST (US National Institute of Standards and Technology) Meeting Recognition Workshop. This workshop (the third such sponsored by NIST) was centerd on the Rich Transcription 2005 Spring Meeting Recognition (RT-05) evaluation of speech technologies within the meeting domain. Building on the success of the RT-04 spring evaluation, the RT-05 evaluation continued the speech-to-text and speaker diarization evaluation tasks and added two new evaluation tasks: speech activity detection and source localization.

MLMI 2005 was thus sponsored by the European Commission (Information Society Technologies priority of the Sixth Framework Programme), the Swiss National Science Foundation and the US National Institute of Standards and Technology.

Given the multiple links between the above projects and several related research areas, and the success of the first MLMI 2004 workshop, it was decided to organize once again a joint workshop bringing together researchers from the different communities working around the common theme of advanced machine learning algorithms for processing and structuring multimodal human interaction. The motivation for creating such a forum, which could be perceived as a number of papers from different research disciplines, evolved from an actual need that arose from these projects and the strong motivation of their partners for such a multidisciplinary workshop. This assessment was confirmed this year by a significant increase in the number of sponsoring research projects, and by the success of the workshop itself, which attracted about 170 participants.

The conference program featured invited talks, full papers (subject to careful peer review, by at least three reviewers), and posters (accepted on the basis of

abstracts) covering a wide range of areas related to machine learning applied to multimodal interaction — and more specifically to multimodal meeting processing, as addressed by the various sponsoring projects. These areas included:

- Human–human communication modeling
- Speech and visual processing
- Multimodal processing, fusion and fission
- Multimodal dialog modeling
- Human–human interaction modeling
- Multimodal data structuring and presentation
- Multimedia indexing and retrieval
- Meeting structure analysis
- Meeting summarizing
- Multimodal meeting annotation
- Machine learning applied to the above

Out of the submitted full papers, about 50% were accepted for publication in the present volume, after having been invited to take review comments and conference feedback into account.

In the present book, and following the structure of the workshop, the papers are divided into the following sections:

1. Invited Papers
2. Multimodal Processing
3. HCI and Applications
4. Discourse and Dialog
5. Emotion
6. Visual Processing
7. Speech and Audio Processing
8. NIST Meeting Recognition Evaluation

Based on the successes of MLMI 2004 and MLMI 2005, it was decided to organize MLMI 2006 in the USA, in collaboration with NIST (US National Institute of Standards and Technology), again in conjunction with the NIST meeting recognition evaluation.

Finally, we take this opportunity to thank our Program Committee members, the sponsoring projects and funding agencies, and those responsible for the excellent management and organization of the workshop and the follow-up details resulting in the present book.

November 2005

Steve Renals
Samy Bengio

Organization

General Chairs

Steve Renals	University of Edinburgh
Samy Bengio	IDIAP Research Institute

Local Organization

Caroline Hastings	University of Edinburgh
Avril Heron	University of Edinburgh
Bartosz Dobrzelecki	University of Edinburgh
Jean Carletta	University of Edinburgh
Mike Lincoln	University of Edinburgh

Program Committee

Marc Al-Hames	Munich University of Technology
Tilman Becker	DFKI
Hervé Bourlard	IDIAP Research Institute
Jean Carletta	University of Edinburgh
Franciska de Jong	University of Twente
John Garofolo	NIST
Thomas Hain	University of Sheffield
Lori Lamel	LIMSI
Benoit Macq	UCL-TELE
Johanna Moore	University of Edinburgh
Laurence Nigay	CLIPS-IMAG
Barbara Peskin	ICSI
Thierry Pun	University of Geneva
Marc Schröder	DFKI
Rainer Stiefelhagen	Universitaet Karlsruhe

NIST Meeting Recognition Workshop Organization

Jon Fiscus	NIST
John Garofolo	NIST

Sponsoring Projects and Institutions

Projects:

- AMI, Augmented Multiparty Interaction, http://www.amiproject.org/
- CHIL, Computers in the Human Interaction Loop, http://chil.server.de/
- HUMAINE, Human–Machine Interaction Network on Emotion, http://emotion-research.net/
- SIMILAR, human–machine interfaces similar to human–human communication, http://www.similar.cc/
- PASCAL, Pattern Analysis, Statistical Modeling and Computational Learning, http://www.pascal-network.org/
- IM2, Interactive Multimodal Information Management, http://www.im2.ch/

Institutions :

- European Commission, through the Multimodal Interfaces objective of the Information Society Technologies (IST) priority of the Sixth Framework Programme.
- Swiss National Science Foundation, through the National Center of Competence in Research (NCCR) program.
- US National Institute of Standards and Technology (NIST), http://www.nist.gov/speech/

Table of Contents

此为试读,需要完整PDF请访问: www.ertongbook.com

III HCI and Applications

IV Discourse and Dialogue

V Emotion

VI Visual Processing

VII Speech and Audio Processing

VIII NIST Meeting Recognition Evaluation

Gesture, Gaze, and Ground

David McNeill

University of Chicago

My emphasis in this paper is on floor control in multiparty discourse: the approach is psycholinguistic. This perspective includes turn management, turn exchange and coordination; how to recognize the dominant speaker even when he or she is not speaking, and a theory of all this. The data to be examined comprise a multimodal depiction of a 5-party meeting (a US Air Force war gaming session) and derive from a project carried out jointly with my engineering colleagues, Francis Quek and Mary Harper. See the Chen et al. paper in this volume for details of the recoding session.

Multiparty discourse can be studied in various ways, e.g., signals of turn taking intentions, marking the next 'projected' turn unit and its content, and still others. I adopt a perspective that emphasizes how speakers coordinate their *individual cognitive states* as they exchange turns while acknowledging and maintaining *the dominant speaker's status*. My goals are similar to Pickering & Garrod's interactive alignment account of dialogue (2004), but with the addition of gesture, gaze, posture, F-formations (Kendon 1990) and several levels of coreferential chains—all to be explained below. I adopt a theoretical position agreeing with their portrayal of dialogue as 'alignment' and of alignment as automatic, in the sense of not draining resources, but not their 'mechanistic' (priming) account of it (cf. Krauss et al. 2004 for qualms). The theory I am following is described in the next section. Alignment in this theory is non-mechanistic, does not single out priming, and regards conversational signaling (cf. papers in Ochs et al. 1996) as providing a synchrony of individual cognitive states, or 'growth points'.

1 Theoretical Background

The growth point. A growth point (GP) is a mental package that combines both linguistic categorial and imagistic components. Combining such semiotic opposites, the GP is inherently multimodal, and creates a condition of instability, the resolution of which propels thought and speech forward. The GP concept, while theoretical, is empirically grounded. GPs are inferred from the totality of communication events with special focus on speech-gesture synchrony and co-expressivity (cf. McNeill 2005 for extensive discussion). It is called a growth point because it is meant to be the initial pulse of thinking for and while speaking, out of which a dynamic process of organization emerges. Growth points are brief dynamic processes during which idea units take form. If two individuals share GPs, they can be said to 'inhabit' the same state of cognitive being and this, in the theoretical picture being considered, is what communication aims to achieve, at least in part. The concept of inhabitance was

S. Renals and S. Bengio (Eds.): MLMI 2005, LNCS 3869, pp. 1 – 14, 2006.
© Springer-Verlag Berlin Heidelberg 2006

expressed by Merleau-Ponty (1962) in the following way: "Language certainly has inner content, but this is not self-subsistent and self-conscious thought. What then does language express, if it does not express thoughts? It presents or rather it *is* the subject's taking up of a position in the world of his meanings" (p. 193; emphasis in the original). The GP is a unit of this process of 'taking up a position in the world of meanings'. On this model, an analysis of conversation should bring out how alignments of inhabitance come about and how, as this is taking place, the overall conversational milieu is maintained by the participants.

The hyperphrase. A second theoretical idea—the 'hyperphrase'—is crucial for analyzing how these alignments and maintenances are attained in complex multi-party meetings. A hyperphrase is a nexus of converging, interweaving processes that cannot be totally untangled. We approach the hyperphrase through a multi-modal structure comprising verbal and non-verbal (gaze, gesture) data.

To illustrate the concept, I shall examine one such phrase from a study carried out jointly with Francis Quek and Mary Harper (the 'Wombats study'). This hyperphrase implies a communicative pulse structured on the verbal, gestural, and gaze levels simultaneously. The hyperphrase began part way into the verbal text (# is an audible breath pause, / is a silent pause, * is a self-interruption; F_0 groups are indicated with underlining, and gaze is in italics):

> we're gonna go over to # *thirty-five* *'cause* / *they're ah** / *they're from the neigh borhood* *they know what's* *going on* #".

The critical aspect indicating a hyperphrase is that gaze turned to the listener in the middle of a linguistic clause and remained there over the rest of the selection. This stretch of speech was also accompanied by multiple occurrences of a single gesture type whereby the right hand with its fingers spread moved up and down over the deictic zero point of the spatialized content of speech. Considering the two non-verbal features, gaze and gesture, together with the lexical content of the speech, this stretch of speech is a *single production pulse* organized thematically around the idea unit, 'the people from the neighborhood in thirty-five.' This would plausibly be a growth point. Such a hyperphrase brings together several linguistic clauses. It spans a self-interruption and repair, and spans 9 F_0 groups. The F_0 groups subdivide the thematic cohesion of the hyperphrase, but the recurrence of similar gesture strokes compensates for the oversegmentation. For example, the F_0 break between "what's" and "going on" is spanned by a single gesture down stroke. It is unlikely that a topic shift occured within this gesture. Thus, the hyperphrase is a production domain in which linguistic clauses, prosody and speech repair all play out, each on its own time-scale, and are held together as the hyperphrase nexus.

Thus we have two major theoretical ideas with which to approach the topic of multiparty discourse—the growth point and the hyperphrase. The GP is the theoretical unit of the speaker's state of cognitive being. The hyperphrase is a package of multimodal information that presents a GP. Through hyperphrases GPs can be shared. Multiple speakers can contribute to the same hyperphrases and growth points. Speaker 2 synchronizes growth points with Speaker 1 by utilizing various turn-taking 'signals' to achieve synchrony. This hypothesis assumes that conversationalists align GPs—Speaker 2 emits signals in a hyperphrase until he/she

此为试读,需要完整PDF请访问: www.ertongbook.com