

Anja Belz  
Roger Evans  
Paul Piwek (Eds.)

LNAI 3123

# Natural Language Generation

Third International Conference, INLG 2004  
Brockenhurst, UK, July 2004  
Proceedings



Springer

TP18-53  
135  
2004

Anja Belz Roger Evans Paul Piwek (Eds.)

# Natural Language Generation

Third International Conference, INLG 2004  
Brockenhurst, UK, July 14-16, 2004  
Proceedings



E200404142



Springer

## Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA  
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

## Volume Editors

Anja Belz

Roger Evans

Paul Piwek

University of Brighton, Information Technology Research Institute  
Lewes Road, Brighton BN2 4GJ, UK

E-mail: {Anja.Belz,Roger.Evans,Paul.Piwek}@itri.brighton.ac.uk

Library of Congress Control Number: 2004108214

CR Subject Classification (1998): I.2.7, I.2, F.4.3

ISSN 0302-9743

ISBN 3-540-22340-1 Springer-Verlag 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-Verlag. Violations are liable to prosecution under the German Copyright Law.

Springer-Verlag is a part of Springer Science+Business Media

[springeronline.com](http://springeronline.com)

© Springer-Verlag Berlin Heidelberg 2004

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign

Printed on acid-free paper SPIN: 11017578 06/3142 5 4 3 2 1 0





## Preface

The Third International Conference on Natural Language Generation (INLG 2004) was held from 14th to 16th July 2004 at Careys Manor, Brockenhurst, UK. Supported by the Association for Computational Linguistics Special Interest Group on Generation, the conference continued a twenty-year tradition of biennial international meetings on research into natural language generation. Recent conference venues have included Mitzpe Ramon, Israel (INLG 2000) and New York, USA (INLG 2002). It was our pleasure to invite the thriving and friendly NLG research community to the beautiful New Forest in the south of England for INLG 2004.

INLG is the leading international conference in the field of natural language generation. It provides a forum for the presentation and discussion of original research on all aspects of the generation of language, including psychological modelling of human language production as well as computational approaches to the automatic generation of language. This volume includes a paper by the keynote speaker, Ardi Roelofs of the Max Planck Institute for Psycholinguistics and the F.C. Donders Centre for Cognitive Neuroimaging, 18 regular papers reporting the latest research results and directions, and 4 student papers describing doctoral work in progress. These papers reveal a particular concentration of current research effort on statistical and machine learning methods, on referring expressions, and on variation in surface realisation. The papers were selected from 46 submissions from all over the world (27 from Europe, 13 from North America, 6 from elsewhere), which were subjected to a rigorous double-blind reviewing process undertaken by our hard-working programme committee. In addition, the conference had a poster session giving a snapshot of ongoing research projects in the field (published as ITRI technical report No. ITRI-04-01).

As always, the conference and this volume of proceedings are the result of the combined efforts of many people. First and foremost we would like to thank all the authors who submitted papers, and the Programme Committee who helped us to put together an excellent conference programme. In addition, we thank everyone who helped us put the conference together, especially the ITRI administrative team, staff at Careys Manor Hotel, and members of the SIGGEN and ACL boards, especially Owen Rambow and Kathy McCoy. Finally we thank all the staff and students at ITRI, and in the wider generation community, for their advice, support and participation in INLG 2004.

July 2004

Anja Belz, Roger Evans and Paul Piwek

# Organisation

INLG 2004 was organised by the Information Technology Research Institute, University of Brighton, UK, on behalf of the Association for Computational Linguistics Special Interest Group on Generation (ACL SIGGEN).

## Co-chairs

Anja Belz, ITRI, Brighton  
Roger Evans, ITRI, Brighton  
Paul Piwek, ITRI, Brighton

## Administration

Domino Moore, ITRI, Brighton  
Martyn Haddock, ITRI, Brighton  
Petra Tank, ITRI, Brighton  
Amy Neale, ITRI, Brighton

## ACL SIGGEN Liaison

Owen Rambow, Columbia  
Kathy McCoy, Delaware

## Programme Committee

Anja Belz, ITRI, University of Brighton, UK (Co-chair)  
Roger Evans, ITRI, University of Brighton, UK (Co-chair)  
Paul Piwek, ITRI, University of Brighton, UK (Co-chair)  
Ion Androutsopoulos, Athens University of Economics and Business, Greece  
Srinivas Bangalore, AT&T, USA  
Regina Barzilay, CSAIL, MIT, USA  
John Bateman, Bremen University, Germany  
Tilman Becker, DFKI, Germany  
Sandra Carberry, CIS, University of Delaware, USA  
Alison Cawsey, CEE, Heriot Watt University, UK  
Robert Dale, Macquarie University, Australia  
Kees van Deemter, ITRI, University of Brighton, UK  
Michael Elhadad, Ben-Gurion University of the Negev, Israel  
Barbara di Eugenio, Computer Science, University of Illinois at Chicago, USA  
Nancy Green, University of North Carolina at Greensboro, USA

## VIII Organisation

Catalina Hallett, ITRI, University of Brighton, UK  
Helmut Horacek, Saarland University, Germany  
Eduard Hovy, ISI, University of Southern California, USA  
Aravind Joshi, CIS, University of Pennsylvania, USA  
Min-Yen Kan, National University of Singapore, Singapore  
Emiel Krahmer, Computational Linguistics, Tilburg University, The Netherlands  
Rodger Kibble, Goldsmiths College, University of London, UK  
Inderjeet Mani, Georgetown University, USA  
Daniel Marcu, ISI, University of Southern California, USA  
Colin Matheson, Informatics, University of Edinburgh, UK  
Kathleen McCoy, CIS, University of Delaware, USA  
Kathleen McKeown, CS, Columbia University, USA  
Chris Mellish, Computer Science, University of Aberdeen, UK  
Detmar Meurers, Linguistics, Ohio State University, USA  
Johanna Moore, Informatics, University of Edinburgh, UK  
Mick O'Donnell, Wagsoft Systems, UK  
Jon Oberlander, Informatics, University of Edinburgh, UK  
Shimei Pan, IBM T.J. Watson Research Center, USA  
Richard Power, ITRI, University of Brighton, UK  
Ehud Reiter, Aberdeen University, UK  
Matthew Stone, CS, Rutgers University, USA  
Sebastian Varges, ITRI, University of Brighton, UK  
Nigel Ward, CS, University of Texas at El Paso, USA  
Ingrid Zukerman, CSSE, Monash University, Australia

# Lecture Notes in Artificial Intelligence (LNAI)

- Vol. 3123: A. Belz, R. Evans, P. Piwek (Eds.), *Natural Language Generation*. X, 219 pages. 2004.
- Vol. 3120: J. Shawe-Taylor, Y. Singer (Eds.), *Learning Theory*. X, 648 pages. 2004.
- Vol. 3097: D. Basin, M. Rusinowitch (Eds.), *Automated Reasoning*. XII, 493 pages. 2004.
- Vol. 3070: L. Rutkowski, J. Siekmann, R. Tadeusiewicz, L.A. Zadeh (Eds.), *Artificial Intelligence and Soft Computing - ICAISC 2004*. XXV, 1208 pages. 2004.
- Vol. 3068: E. André, L. Dybkj{\ae} r, W. Minker, P. Heisterkamp (Eds.), *Affective Dialogue Systems*. XII, 324 pages. 2004.
- Vol. 3067: M. Dastani, J. Dix, A. El Fallah-Seghrouchni (Eds.), *Programming Multi-Agent Systems*. X, 221 pages. 2004.
- Vol. 3066: S. Tsumoto, R. S łowiński, J. Komorowski, J.W. Grzymala-Busse (Eds.), *Rough Sets and Current Trends in Computing*. XX, 853 pages. 2004.
- Vol. 3065: A. Lomuscio, D. Nute (Eds.), *Deontic Logic in Computer Science*. X, 275 pages. 2004.
- Vol. 3060: A.Y. Tawfik, S.D. Goodwin (Eds.), *Advances in Artificial Intelligence*. XIII, 582 pages. 2004.
- Vol. 3056: H. Dai, R. Srikant, C. Zhang (Eds.), *Advances in Knowledge Discovery and Data Mining*. XIX, 713 pages. 2004.
- Vol. 3055: H. Christiansen, M.-S. Hacid, T. Andreasen, H.L. Larsen (Eds.), *Flexible Query Answering Systems*. X, 500 pages. 2004.
- Vol. 3040: R. Conejo, M. Urretavizcaya, J.-L. Pérez-de-la-Cruz (Eds.), *Current Topics in Artificial Intelligence*. XIV, 689 pages. 2004.
- Vol. 3035: M.A. Wimmer (Ed.), *Knowledge Management in Electronic Government*. XII, 326 pages. 2004.
- Vol. 3034: J. Favela, E. Menasalvas, E. Chávez (Eds.), *Advances in Web Intelligence*. XIII, 227 pages. 2004.
- Vol. 3030: P. Giorgini, B. Henderson-Sellers, M. Winikoff (Eds.), *Agent-Oriented Information Systems*. XIV, 207 pages. 2004.
- Vol. 3029: B. Orchard, C. Yang, M. Ali (Eds.), *Innovations in Applied Artificial Intelligence*. XXI, 1272 pages. 2004.
- Vol. 3025: G.A. Vouros, T. Panayiotopoulos (Eds.), *Methods and Applications of Artificial Intelligence*. XV, 546 pages. 2004.
- Vol. 3012: K. Kurumatani, S.-H. Chen, A. Ohuchi (Eds.), *Multi-Agnets for Mass User Support*. X, 217 pages. 2004.
- Vol. 3010: K.R. Apt, F. Fages, F. Rossi, P. Szeredi, J. Vánca (Eds.), *Recent Advances in Constraints*. VIII, 285 pages. 2004.
- Vol. 2990: J. Leite, A. Omicini, L. Sterling, P. Torroni (Eds.), *Declarative Agent Languages and Technologies*. XII, 281 pages. 2004.
- Vol. 2980: A. Blackwell, K. Marriott, A. Shimojima (Eds.), *Diagrammatic Representation and Inference*. XV, 448 pages. 2004.
- Vol. 2977: G. Di Marzo Serugendo, A. Karageorgos, O.F. Rana, F. Zambonelli (Eds.), *Engineering Self-Organising Systems*. X, 299 pages. 2004.
- Vol. 2972: R. Monroy, G. Arroyo-Figueroa, L.E. Sucar, H. Sossa (Eds.), *MICAI 2004: Advances in Artificial Intelligence*. XVII, 923 pages. 2004.
- Vol. 2961: P. Eklund (Ed.), *Concept Lattices*. IX, 411 pages. 2004.
- Vol. 2953: K. Konrad, *Model Generation for Natural Language Interpretation and Analysis*. XIII, 166 pages. 2004.
- Vol. 2934: G. Lindemann, D. Moldt, M. Paolucci (Eds.), *Regulated Agent-Based Social Systems*. X, 301 pages. 2004.
- Vol. 2930: F. Winkler (Ed.), *Automated Deduction in Geometry*. VII, 231 pages. 2004.
- Vol. 2926: L. van Elst, V. Dignum, A. Abecker (Eds.), *Agent-Mediated Knowledge Management*. XI, 428 pages. 2004.
- Vol. 2923: V. Lifschitz, I. Niemelä (Eds.), *Logic Programming and Nonmonotonic Reasoning*. IX, 365 pages. 2004.
- Vol. 2915: A. Camurri, G. Volpe (Eds.), *Gesture-Based Communication in Human-Computer Interaction*. XIII, 558 pages. 2004.
- Vol. 2913: T.M. Pinkston, V.K. Prasanna (Eds.), *High Performance Computing - HiPC 2003*. XX, 512 pages. 2003.
- Vol. 2903: T.D. Gedeon, L.C.C. Fung (Eds.), *AI 2003: Advances in Artificial Intelligence*. XVI, 1075 pages. 2003.
- Vol. 2902: F.M. Pires, S.P. Abreu (Eds.), *Progress in Artificial Intelligence*. XV, 504 pages. 2003.
- Vol. 2892: F. Dau, *The Logic System of Concept Graphs with Negation*. XI, 213 pages. 2003.
- Vol. 2891: J. Lee, M. Barley (Eds.), *Intelligent Agents and Multi-Agent Systems*. X, 215 pages. 2003.
- Vol. 2882: D. Veit, *Matchmaking in Electronic Markets*. XV, 180 pages. 2003.
- Vol. 2871: N. Zhong, Z.W. Raś, S. Tsumoto, E. Suzuki (Eds.), *Foundations of Intelligent Systems*. XV, 697 pages. 2003.
- Vol. 2854: J. Hoffmann, *Utilizing Problem Structure in Planning*. XIII, 251 pages. 2003.
- Vol. 2843: G. Grieser, Y. Tanaka, A. Yamamoto (Eds.), *Discovery Science*. XII, 504 pages. 2003.

- Vol. 2842: R. Gavaldá, K.P. Jantke, E. Takimoto (Eds.), *Algorithmic Learning Theory*. XI, 313 pages. 2003.
- Vol. 2838: N. Lavrač, D. Gamberger, L. Todorovski, H. Blockeel (Eds.), *Knowledge Discovery in Databases: PKDD 2003*. XVI, 508 pages. 2003.
- Vol. 2837: N. Lavrač, D. Gamberger, L. Todorovski, H. Blockeel (Eds.), *Machine Learning: ECML 2003*. XVI, 504 pages. 2003.
- Vol. 2835: T. Horváth, A. Yamamoto (Eds.), *Inductive Logic Programming*. X, 401 pages. 2003.
- Vol. 2821: A. Günter, R. Kruse, B. Neumann (Eds.), *KI 2003: Advances in Artificial Intelligence*. XII, 662 pages. 2003.
- Vol. 2807: V. Matoušek, P. Mautner (Eds.), *Text, Speech and Dialogue*. XIII, 426 pages. 2003.
- Vol. 2801: W. Banzhaf, J. Ziegler, T. Christaller, P. Dittrich, J.T. Kim (Eds.), *Advances in Artificial Life*. XVI, 905 pages. 2003.
- Vol. 2797: O.R. Zaiane, S.J. Simoff, C. Djeraba (Eds.), *Mining Multimedia and Complex Data*. XII, 281 pages. 2003.
- Vol. 2792: T. Rist, R.S. Aylett, D. Ballin, J. Rickel (Eds.), *Intelligent Virtual Agents*. XV, 364 pages. 2003.
- Vol. 2782: M. Klusch, A. Omicini, S. Ossowski, H. Laamanen (Eds.), *Cooperative Information Agents VII*. XI, 345 pages. 2003.
- Vol. 2780: M. Dojat, E. Keravnou, P. Barahona (Eds.), *Artificial Intelligence in Medicine*. XIII, 388 pages. 2003.
- Vol. 2777: B. Schölkopf, M.K. Warmuth (Eds.), *Learning Theory and Kernel Machines*. XIV, 746 pages. 2003.
- Vol. 2752: G.A. Kaminka, P.U. Lima, R. Rojas (Eds.), *RoboCup 2002: Robot Soccer World Cup VI*. XVI, 498 pages. 2003.
- Vol. 2741: F. Baader (Ed.), *Automated Deduction – CADE-19*. XII, 503 pages. 2003.
- Vol. 2705: S. Renals, G. Grefenstette (Eds.), *Text- and Speech-Triggered Information Access*. VII, 197 pages. 2003.
- Vol. 2703: O.R. Zaiane, J. Srivastava, M. Spiliopoulou, B. Masand (Eds.), *WEBKDD 2002 - Mining Web Data for Discovering Usage Patterns and Profiles*. IX, 181 pages. 2003.
- Vol. 2700: M.T. Pazzienza (Ed.), *Extraction in the Web Era*. XIII, 163 pages. 2003.
- Vol. 2699: M.G. Hinchey, J.L. Rash, W.F. Truszkowski, C.A. Rouff, D.F. Gordon-Spears (Eds.), *Formal Approaches to Agent-Based Systems*. IX, 297 pages. 2002.
- Vol. 2691: V. Mařík, J.P. Müller, M. Pechoucek (Eds.), *Multi-Agent Systems and Applications III*. XIV, 660 pages. 2003.
- Vol. 2684: M.V. Butz, O. Sigaud, P. Gérard (Eds.), *Anticipatory Behavior in Adaptive Learning Systems*. X, 303 pages. 2003.
- Vol. 2671: Y. Xiang, B. Chaib-draa (Eds.), *Advances in Artificial Intelligence*. XIV, 642 pages. 2003.
- Vol. 2663: E. Menasalvas, J. Segovia, P.S. Szczepaniak (Eds.), *Advances in Web Intelligence*. XII, 350 pages. 2003.
- Vol. 2661: P.L. Lanzi, W. Stolzmann, S.W. Wilson (Eds.), *Learning Classifier Systems*. VII, 231 pages. 2003.
- Vol. 2654: U. Schmid, *Inductive Synthesis of Functional Programs XXII*, 398 pages. 2003.
- Vol. 2650: M.-P. Huget (Ed.), *Communications in Multi-agent Systems*. VIII, 323 pages. 2003.
- Vol. 2645: M.A. Wimmer (Ed.), *Knowledge Management in Electronic Government*. XI, 320 pages. 2003.
- Vol. 2639: G. Wang, Q. Liu, Y. Yao, A. Skowron (Eds.), *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*. XVII, 741 pages. 2003.
- Vol. 2637: K.-Y. Whang, J. Jeon, K. Shim, J. Srivastava, *Advances in Knowledge Discovery and Data Mining*. XVIII, 610 pages. 2003.
- Vol. 2636: E. Alonso, D. Kudenko, D. Kazakov (Eds.), *Adaptive Agents and Multi-Agent Systems*. XIV, 323 pages. 2003.
- Vol. 2627: B. O'Sullivan (Ed.), *Recent Advances in Constraints*. X, 201 pages. 2003.
- Vol. 2600: S. Mendelson, A.J. Smola (Eds.), *Advanced Lectures on Machine Learning*. IX, 259 pages. 2003.
- Vol. 2592: R. Kowalczyk, J.P. Müller, H. Tianfield, R. Unland (Eds.), *Agent Technologies, Infrastructures, Tools, and Applications for E-Services*. XVII, 371 pages. 2003.
- Vol. 2586: M. Klusch, S. Bergamaschi, P. Edwards, P. Petta (Eds.), *Intelligent Information Agents*. VI, 275 pages. 2003.
- Vol. 2583: S. Matwin, C. Sammut (Eds.), *Inductive Logic Programming*. X, 351 pages. 2003.
- Vol. 2581: J.S. Sichman, F. Bousquet, P. Davidsson (Eds.), *Multi-Agent-Based Simulation*. X, 195 pages. 2003.
- Vol. 2577: P. Petta, R. Tolksdorf, F. Zambonelli (Eds.), *Engineering Societies in the Agents World III*. X, 285 pages. 2003.
- Vol. 2569: D. Karagiannis, U. Reimer (Eds.), *Practical Aspects of Knowledge Management*. XIII, 648 pages. 2002.
- Vol. 2560: S. Goronzy, *Robust Adaptation to Non-Native Accents in Automatic Speech Recognition*. XI, 144 pages. 2002.
- Vol. 2557: B. McKay, J. Slaney (Eds.), *AI 2002: Advances in Artificial Intelligence*. XV, 730 pages. 2002.
- Vol. 2554: M. Beetz, *Plan-Based Control of Robotic Agents*. XI, 191 pages. 2002.
- Vol. 2543: O. Bartenstein, U. Geske, M. Hannebauer, O. Yoshie (Eds.), *Web Knowledge Management and Decision Support*. X, 307 pages. 2003.
- Vol. 2541: T. Barkowsky, *Mental Representation and Processing of Geographic Knowledge*. X, 174 pages. 2002.
- Vol. 2533: N. Cesa-Bianchi, M. Numao, R. Reischuk (Eds.), *Algorithmic Learning Theory*. XI, 415 pages. 2002.
- Vol. 2531: J. Padgett, O. Shehory, D. Parkes, N.M. Sadeh, W.E. Walsh (Eds.), *Agent-Mediated Electronic Commerce IV. Designing Mechanisms and Systems*. XVII, 341 pages. 2002.
- Vol. 2527: F.J. Garijo, J.-C. Riquelme, M. Toro (Eds.), *Advances in Artificial Intelligence - IBERAMIA 2002*. XVIII, 955 pages. 2002.

# Table of Contents

## Keynote Paper

The Seduced Speaker: Modeling of Cognitive Control .....	1
<i>Ardi Roelofs</i>	

## Regular Papers

Generating Intensional Answers in Intelligent Question Answering Systems .....	11
<i>Farah Benamara</i>	
Saliency-Driven Text Planning .....	21
<i>Christian Chiarcos, Manfred Stede</i>	
Finetuning NLG Through Experiments with Human Subjects: The Case of Vague Descriptions.....	31
<i>Kees van Deemter</i>	
Indirect Supervised Learning of Content Selection Logic .....	41
<i>Pablo A. Duboue</i>	
Generating Referring Expressions Using Perceptual Groups.....	51
<i>Kotaro Funakoshi, Satoru Watanabe, Naoko Kuriyama, Takenobu Tokunaga</i>	
The Use of a Structural N-gram Language Model in Generation-Heavy Hybrid Machine Translation .....	61
<i>Nizar Habash</i>	
On Referring to Sets of Objects Naturally .....	70
<i>Helmut Horacek</i>	
An ATMS Approach to Systemic Sentence Generation .....	80
<i>Hasan Kamal, Chris Mellish</i>	
A Corpus-Based Methodology for Evaluating Metrics of Coherence for Text Structuring .....	90
<i>Nikiforos Karamanis, Chris Mellish, Jon Oberlander, Massimo Poesio</i>	
Classification-Based Generation Using TAG .....	100
<i>Tomasz Marciniak, Michael Strube</i>	
Resolving Structural Ambiguity in Generated Speech .....	110
<i>Chris Mellish</i>	

A Framework for Stylistically Controlled Generation.....	120
<i>Daniel S. Paiva, Roger Evans</i>	
SEGUE: A Hybrid Case-Based Surface Natural Language Generator ....	130
<i>Shimei Pan, James Shaw</i>	
Modelling Politeness in Natural Language Generation .....	141
<i>Kaśka Porayska-Pomsta, Chris Mellish</i>	
Context-Based Incremental Generation for Dialogue .....	151
<i>Matthew Purver, Ruth Kempson</i>	
Contextual Influences on Near-Synonym Choice.....	161
<i>Ehud Reiter, Somayajulu Sripada</i>	
Overgenerating Referring Expressions Involving Relations and Booleans..	171
<i>Sebastian Varges</i>	
Reining in CCG Chart Realization .....	182
<i>Michael White</i>	
<b>Student Papers</b>	
Categorization of Narrative Semantics for Use in Generative Multidocument Summarization .....	192
<i>David K. Elson</i>	
Corpus-Based Planning of Deictic Gestures in COMIC.....	198
<i>Mary Ellen Foster</i>	
Hybrid NLG in a Generic Dialog System .....	205
<i>Martin Klärner</i>	
Two Kinds of Alternative Sets and a Marking Principle – When to Say Also .....	212
<i>Kristina Striegnitz</i>	
<b>Author Index .....</b>	<b>219</b>



# The Seduced Speaker: Modeling of Cognitive Control

Ardi Roelofs

Max Planck Institute for Psycholinguistics, F.C. Donders Centre for Cognitive Neuroimaging, and Nijmegen Institute for Cognition and Information,  
Wundtlaan 1, 6525 XD, Nijmegen, the Netherlands

[ardi@mpi.nl](mailto:ardi@mpi.nl)

<http://www.mpi.nl/world/persons/profession/ardi.html>

**Abstract.** Although humans are the ultimate “natural language generators”, the area of psycholinguistic modeling has been somewhat underrepresented in recent approaches to Natural Language Generation in computer science. To draw attention to the area and illustrate its potential relevance to Natural Language Generation, I provide an overview of recent work on psycholinguistic modeling of language production together with some key empirical findings, state-of-the-art experimental techniques, and their historical roots. The techniques include analyses of speech-error corpora, chronometric analyses, eyetracking, and neuroimaging. The overview is built around the issue of cognitive control in natural language generation, concentrating on the production of single words, which is an essential ingredient of the generation of larger utterances. Most of the work exploited the fact that human speakers are good but not perfect at resisting temptation, which has provided some critical clues about the nature of the underlying system.

## 1 Introduction

Unlike most Natural Language Generation programs that run on serial, digital computers, human speakers are occasionally distracted while performing a natural language generation task. “I can resist everything except temptation” (p. 5), a play character of Oscar Wilde [1] once said, and this difficulty in resisting temptation holds for most people. Distractibility seems to be the price paid for the parallelism of the human brain. One of the key tasks of the human cognitive system is to select one appropriate action at any given moment and to focus the machinery of planning and movement on that action. Selectivity of attention is required for the coherent control of action. At the same time, the system needs to remain open to events that may happen outside the focus of attention (e.g., to detect possible danger in the background). The opposing forces of the need to focus and the need to remain open make the human system distractible. This raises the issue of cognitive control.

In speaking, the distractibility of the human cognitive system is revealed by speech errors and delays in initiating articulation. The distractibility is also evident from the eye movements that speakers make. By examining speech errors,

delays, and eye movements, researchers have discovered much about the cognitive foundations of speaking. Computer models have been developed that account for the kinds of speech errors that occur and their relative frequencies, and also for the eye movements and the exact duration of the delays caused by distraction. Furthermore, much has been discovered about the brain areas that are involved in speaking. Computer models can even predict the time course of the increase in blood flow to certain brain areas required for speech production. I provide an overview of work in psycholinguistics that tried to shed light on the human language generation system using evidence from speech-error corpora, chronometric experiments, eyetracking, and neuroimaging. The overview is built around the issue of cognitive control in natural language generation. It concentrates on the production of single words, which is an essential component of the generation of larger utterances. Nearly all of the work that is reviewed exploited the fact that human speakers are good but not perfect at resisting temptation, which has provided important evidence about the nature of the underlying system.

## 2 What Speech Errors Say About Speaking

A slip of the tongue or speech error is an unintended, nonhabitual deviation from a speech plan. Meringer and Mayer [2] were among the first to draw attention to speech errors as a data source that might illuminate the mechanisms underlying speech production. In 1895, they published a large corpus of German speech errors along with a theoretical analysis. They made several seminal observations. First, they discovered that slips of the tongue are typically meaning-based or form-based. The substitution of “dog” for “cat” is a meaning-based error and the substitution “cap” for “cat” is a form-based one. The distinction suggests that words are planned at a conceptual level and at a form level. Second, they observed that there is often a form-relation in meaning-based errors (e.g., “calf” for “cat”), suggesting that the planning levels do not operate completely independently, although this is still a hotly debated issue [3]. Third, they observed that contextual errors may be anticipations (e.g., “leading list” for “reading list”), perseverations (e.g., “beef *needle*” for “beef noodle”), exchanges (e.g., “*flow snurries*” for “snow flurries”), or blends (e.g., “clear” combining “close” and “near”).

Although speech error analyses continued to be carried out during the next half century, there was a real revival of interest in the late 1960s. In 1973, Fromkin [4] edited an influential book on speech errors that included an appendix with part of her own speech error corpus. Another important corpus was collected during the early 1970s at MIT by Garrett and colleagues. Garrett [5] discovered that word exchanges such as the exchange of *roof* and *list* in “we completely forgot to add the *list* to the *roof*” tend to involve elements from different phrases and of the same syntactic category, here noun. By contrast, segment exchanges such as “she is a real *rack pat*” for “pack rat” are likely to involve elements from the same phrase and they do not respect lexical category. Garrett explained this finding by assuming a level of syntactic planning (at which the lexical exchanges

occur) that is different from the level of form planning (at which the segment exchanges occur). Garrett argued that the speech errors also provide support for a distinct morphological planning level.

Some morphemic errors appear to happen at the syntactic level, whereas others arise at the form level. For example, in “how many *pies* does it take to make an *apple*?”, the interacting stems (i.e., *pie* and *apple*) belong to the same syntactic category and come from distinct phrases. Note that the plurality of *apple* is realized on *pie*, which suggests that a number parameter is set. The distributional properties of these morpheme exchanges are similar to those of whole-word exchanges. This suggests that these morpheme errors and whole-word errors occur at the same level of planning. They seem to occur when lexical items in a developing syntactic structure trade places. Similarly, errors such as “I’d *hear* one if I *knew* it” for “I’d *know* one if I *heard* it” suggest that syntactically specified lexical representations may trade places independently of their concrete morphophonological specifications. By contrast, the exchanging morphemes in an error such as “*slicely thinned*” for “thinly sliced” belong to different syntactic categories and come from the same phrase, which is also characteristic of segment exchanges. This suggests that this second type of morpheme error and segment errors occur at the same level of planning, namely the level at which morphemes and segments are retrieved and the morphophonological form of the utterance is constructed.

On the basis of his speech error analyses, Garrett [5] proposed an unimplemented model of speech production that distinguished between conceptual, syntactic, morphological, phonological, and phonetic levels of speech planning. Ten years later, Dell [6] developed the first computer model of memory retrieval in sentence production, instantiating several of Garrett’s insights. Following a long associationist tradition that began with Aristotle [7], Dell convincingly argued that our word memory is organized as an associative network that is accessed by spreading activation. The network contains nodes for conceptual, syntactic, morphological, phonological, and phonetic information about words. In retrieving information for concepts to be verbally expressed, activation spreads from the corresponding concept nodes to associated nodes in the network. After fixed periods of time, the highest activated lexical, morpheme, phoneme, and phonetic nodes are selected. Dell’s associative network model of word memory provided quantitative accounts of the major facts about speech errors: the kinds of errors that occur, their relative frequencies, and the constraints on their form and occurrence. On the account, errors occur when, because of noise in the system or influences outside the intended utterance (distraction), another node in the network than the target one is the most highly activated node and becomes erroneously selected.

### 3 What Response Times Say About Speaking

The first person to measure (in milliseconds) speech production latencies – the time between stimulus onset and the initiation of a verbal response – was Don-

ders [8]. Until Donders' work in the 1860s, most scientists had assumed that the mental operations involved in responding to a stimulus occur instantaneously. Donders designed a subtraction technique to time the different mental processes that the brain goes through when faced with different tasks. His chronometric work demonstrated a simple principle: The time it takes to perform a task depends on the number and types of mental stages involved. With this observation, he laid the foundation of a research programme that is still extremely productive today: the componential processing analysis of human task performance. At the end of his seminal article on the measurement of mental processing times [8], Donders reports that "distraction during the appearance of the stimulus is always punished with prolongation of the process" (p. 428). This observation is interesting in the light of later research developments exploiting distraction, in particular, the work of Stroop in the 1930s. Surprisingly, it was only in the 1990s that speech error and chronometric analyses became equal partners in the study of speaking. Most of the chronometric work that has been done addressed the production of single words or simple phrases. This seems to be due to the fact that it is awfully difficult to investigate the generation of more complex utterances (sentences and discourse) in controlled experimental settings. Still, the investigation of word production has provided some key insights into the algorithms that underlie human language generation. The first computer model of word production built on chronometric evidence is *WEAVER++* [9] [10] [11]. This model recognizes the key insights from the speech error analyses, but it was specifically designed to provide a unifying account of the increasing body of chronometric data.

Like Dell's model, *WEAVER++* assumes that word planning involves the retrieval of information from an associative network through spreading activation. In addition, *WEAVER++* assumes that the associations are labeled, because a mere associative link between two nodes in a network tells nothing about the relation between the entities represented. For example, the concept *RED(X)* is strongly associated with both *GREEN(X)* and *FIRE(X)*, but the relationship between *RED(X)* and *GREEN(X)* is very different from the relationship between *RED(X)* and *FIRE(X)*. The importance of representing the relation between entities symbolically was first recognized by Selz [7] in the early 1900s. Labeled links have become a central part of semantic networks in computer science since the seminal work of Quillian in the late 1960s.

To explain even the simplest forms of language generation, like single word production, it is not enough to assume an associative memory and spreading activation. Natural language is a very flexible tool that can be used to achieve various goals. Around 1900, Watt, Ach, and Külpe of the Würzburg school [7], as well as Selz [7], called attention to the importance of understanding the directedness of action in general and verbal action in particular (the problem of cognitive control is also referred to in the literature as the problem of attentional, executive, or willed control.) They convincingly argued that the various associative models that had been developed during the past two millennia failed to explain the directedness of thought and action. Plato already drew attention

to the problem, which he characterized in *Phaedrus* as the problem of a charioteer attempting to manage a number of horses pulling in different directions. Until recently, this aspect of natural language generation was neglected in psycholinguistic research on word production. The directedness of natural language generation has, in its simplest form, perhaps been most intensively studied by using the “gold standard” of attentional measures, the color-word Stroop task [12] and picture-word analogs of it. The Stroop task is one of the most widely used tasks in academic and applied psychology, reviewed by MacLeod [13]. In the classic, color-word version of the task, speakers name the ink color of color words (one basic task variant) or they read the color words aloud (another basic task variant). In performing the Stroop task, speakers are slower naming the ink colors of incongruent color words (e.g., the word BLUE in red ink) than of a series of Xs. Word reading times are unaffected by incongruent ink colors. The correct naming of the colors of incongruent color words shows that goals keep verbal actions on track in the face of distraction, albeit with a temporal cost.

Issues of cognitive control were already explored in the early days of experimental psychology (between 1870-1920) by Cattell, Donders, James, and Wundt, who saw all his work on response times as studies of volition [14]. However, no progress was made in understanding the mechanisms of control. Associationist and behaviorist theories accounted for action selection by postulating associations between stimuli and responses (e.g., Müller in the early 1900s [7], and later Watson and Skinner). However, if all our actions were determined exclusively by stimulus-response associations, goals could not determine which action to make because the strongest association would automatically determine the response. Watt and Ach of the Würzburg school [7] extended the idea of stimulus-response associations to associations between stimuli and an internally represented task goal (“Aufgabe”), on the one hand, and responses, on the other. Later theoretical developments are descendants of this idea.

On the view that currently dominates the attention and performance literature, which was anticipated by Müller [7] and recently implemented in the GRAIN computer model by Botvinick and colleagues [15], goals associatively bias the activation of one response pathway (e.g., for color naming) rather than another (e.g., for oral reading). On another view, implemented in WEAVER++ [16], attentional control arises from explicit, symbolic reference to goals, accomplished by condition-action rules. WEAVER++’s associative network is accessed by spreading activation while the condition-action rules determine what is done with the activated lexical information depending on the task. When a goal symbol (e.g., indicating to name the color) is placed in working memory, attention is focused on those rules that include the goal among their conditions (e.g., those for color naming rather than reading). Words are planned by incrementally extending verbal goals. Lexical nodes are selected for concept nodes, morpheme nodes for lexical nodes, segment nodes for morpheme nodes, and phonetic syllable program nodes for syllabified segment nodes, whereby the syllabification of segments proceeds incrementally from the beginning of a word to its end.

The idea of incrementality in natural language generation was first proposed by Wundt [14].

WEAVER++'s combination of a spreading activation network with a parallel system of goal-factored condition-action rules yields a simple but powerful and efficient device for selecting one line of action among the available options. The crucial role of spreading activation in the model is to provide a relevance heuristic. Spreading activation serves to solve the "frame problem" that confronts any cognitive system. In making decisions, a cognitive system can, in principle, draw on all the information available, but the amount may be indefinitely large in that everything may potentially be relevant. The frame problem is how to get at the relevant information and when to stop thinking and start acting. Spreading activation is a parallel mechanism for making relevant information available and triggering relevant computations, following the heuristic that information associated with the current information is likely of direct relevance too. Triggering condition-action rules by spreading activation prevents the problem of all rules having to test their conditions at any one moment in time. Only the rules that are linked to a sufficiently activated piece of associative memory become evaluated. For example, in naming a color, no more than a dozen or so condition-action rules test their conditions rather than all rules in a speaking lexicon of some 30,000 words. Moreover, because condition-action rules may be triggered by the activation of elements outside the current focus of attention, the system remains open to what happens in the background.

The idea of goal-referenced control originated with Selz [7] and it flourished in the work of Newell and Simon, Anderson and colleagues [17], and others, on higher-level cognitive processes like problem solving (e.g., playing chess, proving logic theorems, and solving puzzles such as the Tower of Hanoi), where associative models generally failed. However, due to the traditional partitioning of experimental psychology into cognition, perception, and action, with little communication across the boundaries, goal-referenced control models have had little impact in the perception-action literature, because they generally did not aim at fine-grained modeling of the temporal structure of human information processing in the attention and performance tradition. Only recently, goal-referenced control made successful strides into the attention and performance literature. For example, there are now successful models for fine-grained aspects of visual attention, dual-task performance [17], and Stroop [16]. It seems that we are on the verge of a unified account of the control of cognition, perception, and action.

## 4 What Eye Movements Say About Speaking

In the second half of the nineteenth century, well before the modern era of cognitive and brain sciences, Donders and Wundt studied eye movements and constructed mechanical models for them. Whereas before those days the eyes used to be poetically called a window to the soul, Wundt took gazes to be a window into the operation of the attention system. As Wundt [14] reasoned in his *Outlines of psychology*, visual acuity is best at the point of fixation. Therefore,