EDITORS Gokhan Tur • Renato De Mori

SPOKEN LANGUAGE Understanding

Systems for Extracting Semantic Information from Speech

WILEY

Contents

Index

SPOKEN LANGUAGE UNDERSTANDING

SYSTEMS FOR EXTRACTING SEMANTIC INFORMATION FROM SPEECH

Gokhan Tur

Microsoft Speech Labs, Microsoft Research, USA

Renato De Mori

McGill University, Montreal, Canada and University of Avignon, France

WILEY

A John Wiley and Sons, Ltd, Publication

This edition first published 2011

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Spoken language understanding : systems for extracting semantic information from speech / edited by Gokhan Tur, Renato De Mori.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-68824-3 (hardback)

1. Speech processing systems. 2. Semantics. 3. Discourse analysis. 4. Corpora (Linguistics)

I. Tur, Gokhan. II. De Mori, Renato. III. Title.

P95.3.S665 2010

006.4'54–dc22

2010051228

A catalogue record for this book is available from the British Library.

Print ISBN: 9780470688243

E-PDF ISBN: 9781119992707

O-book ISBN: 9781119992691

E-Pub ISBN: 9781119993940

Mobi ISBN: 9781119993957

In memory of Fred Jelinek (1932-2010)

List of Conntibutors

Alex Acero received the degrees of MS from the Polytechnic University of Madrid, Spain, in 1985, MS from Rice University, Houston, TX, in 1987, and PhD from Carnegie Mellon University, Pittsburgh, PA, in 1990, all in electrical engineering. He worked in Apple Computers Advanced Technology Group from 1990 to 1991. In 1992, he joined Telefonica I+D, Madrid, as Manager of the Speech Technology Group. Since 1994, he has been with Microsoft Research, Redmond, WA, where he is currently a Research Area Manager directing an organization with 70 engineers conducting research in audio, speech, multimedia, communication, natural language, and information retrieval. He is also an affiliate Professor of Electrical Engineering at the University of Washington, Seattle. Dr. Acero is author of the books *Acoustical and Environmental Robustness in Automatic Speech Recognition* (Kluwer, 1993) and *Spoken Language Processing* (Prentice-Hall, 2001), has written invited chapters in four edited books and 200 technical papers. He holds 53 US patents.

Dr. Acero has served the IEEE Signal Processing Society as Vice President Technical Directions (2007–2009), 2006 Distinguished Lecturer, as a member of the Board of Governors (2004–2005), as an Associate Editor for the IEEE Signal Processing Letters (2003–2005) and the IEEE Transactions on Audio, Speech and Language Processing (2005–2007), and as a member of the editorial board of the IEEE Journal of Selected Topics in Signal Processing (2006–2008) and the IEEE Signal Processing Magazine (2008–2010). He also served as member (1996–2000) and Chair (2000–2002) of the Speech Technical Committee of the IEEE Signal Processing Society. He was

Publications Chair of ICASSP'98, Sponsorship Chair of the 1999 IEEE Workshop on Automatic Speech Recognition and Understanding, and General Co-chair of the 2001 IEEE Workshop on Automatic Speech Recognition and Understanding. Since 2004, Dr. Acero, along with co-authors Dr. Huang and Dr. Hon, has been using proceeds from their textbook *Spoken Language Processing* to fund the IEEE Spoken Language Processing Student Travel Grant for the best ICASSP student papers in the speech area. Dr. Acero is a member of the editorial board of Computer Speech and Language and he served as a member of Carnegie Mellon University Deans Leadership Council for College of Engineering.

Frédéric Béchet is a researcher in the field of Speech and Natural Language Processing. His research activities are mainly focused on Spoken Language Understanding for both Spoken Dialogue Systems and Speech Mining applications.

After studying Computer Science at the University of Marseille, he obtained his PhD in Computer Science in 1994 from the University of Avignon, France. Since then he worked at the Ludwig Maximilian University in Munich, Germany, as a Professor Assistant at the University of Avignon, France, as an invited professor at AT&T Research Shannon Lab in Florham Park, New Jersey, USA, and he is currently a full Professor of Computer Science at the Aix Marseille Université in France. Frédéric Béchet is the author/co-author of over 60 refereed papers in journals and international conferences.

Ciprian Chelba received his Diploma Engineer degree in 1993 from the Faculty of Electronics and Telecommunications at Politehnica University, Bucuresti, Romania, and the degrees of MS in 1996 and PhD in 2000 from the Electrical and Computer Engineering Department at the Johns Hopkins University. He is a research scientist with Google and has previously worked at Microsoft Research. His research interests are in statistical modeling of natural language and speech, as well as related areas such as machine learning. Recent projects include large scale language modeling for Google Search by Voice, and~indexing, ranking and snippeting of speech content. He is a member of the IEEE, and has served one full term on the IEEE Signal Processing Society Speech and Language Technical Committee (2006–2008), among other community activities.

Renato De Mori received a doctorate degree in Electronic Engineering from Politecnico di Torino (Italy). He is a Fellow of the IEEE Computer Society and has been distinguished lecturer of the IEEE Signal Processing Society.

He has been Professor and Chairman at the University of Turin (Italy) and at McGill University, School of Computer Science (Monteral, Canada), professor at the University of Avignon (France). He is now emeritus professor at McGill University and at the University of Avignon. His major contributions have been in Automatic Speech Recognition and Understanding, Signal Processing, Computer Arithmetic, Software Engineering and Human/Machine Interfaces.

He is Associated Editor of the IEEE Transactions on Audio Speech and Language Processing, has been Chief Editor of SPEECH COMMUNICATION (2003–2005), Associate Editor of the IEEE Transactions on Pattern Analysis and Machine Intelligence (1998–1992). He has been a member of the editorial board of Computer Speech and Language since 1988.

Professor De Mori has been a member of the Executive Advisory Board at the IBM Toronto Lab, Scientific Advisor at France Télécom R&D, Chairman of the Computer and

Information Systems Committee, Natural Sciences and Engineering Council of Canada, Vice-President R&D, Centre de Recherche en Informatique de Montral.

He has been a member of the IEEE Speech Technical Committee (1984–1987, 2003–2006), the Interdisciplinary Board, Canadian Foundation for Innovation, Interdisciplinary Committee for Canadian chairs. He has been involved in many Canadian and European projects and has been scientific leader of the LUNA European project on spoken language understanding (2006–2009).

Li Deng received his Bachelor degree from the University of Science and Technology of China (with the Guo Mo-Ruo Award), and received the degree of PhD from the University of Wisconsin, Madison (with the Jerzy E. Rose Award). In 1989, he joined the Department of Electrical and Computer Engineering, University of Waterloo, Ontario, Canada as an Assistant Professor, where he became a Full Professor in 1996.

From 1992 to 1993, he conducted sabbatical research at Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, Mass, and from 1997 to 1998, at ATR Interpreting Telecommunications Research Laboratories, Kyoto, Japan. In 1999, he joined Microsoft Research, Redmond, WA as a Senior Researcher, where he is currently a Principal Researcher. He is also an Affiliate Professor in the Department of Electrical Engineering at University of Washington, Seattle. His past and current research activities include automatic speech and speaker recognition, spoken language identification and understanding, speech-to-speech translation, machine translation, statistical methods and machine learning, neural information processing, deep-structured learning, machine intelligence, audio and acoustic signal processing, statistical signal processing and digital communication, human speech production and perception, acoustic phonetics, auditory speech processing, auditory physiology and modeling, noise robust speech processing, speech synthesis and enhancement, multimedia signal processing, and multimodal human–computer interaction. In these areas, he has published over 300 refereed papers in leading international conferences and journals, 12 book chapters, and has given keynotes, tutorials, and lectures worldwide. He has been granted over 30 US or international patents in acoustics, speech/language technology, and signal processing.

He is a Fellow of the Acoustical Society of America, and a Fellow of the IEEE. He has authored or co-authored three books in speech processing and learning. He serves on the Board of Governors of the IEEE Signal Processing Society (2008–2010), and as Editor-in-Chief for the IEEE Signal Processing Magazine (2009–2012), which ranks consistently among the top journals with the highest citation impact. According to the Thomson Reuters Journal Citation Report, released June 2010, the SPM has ranked first among all IEEE publications (125 in total) and among all publications within the Electrical and Electronics Engineering Category (245 in total) in terms of its impact factor.

Olivier Galibert is an engineer in the Information Systems Evaluation group at LNE which he joined in 2009. He recieved his engineering degree in 1994 from the Ecole Nationale des Mines de Nancy, France and his PhD in 2009 from the University Paris – Sud 11, France. Previously to his joining LNE, he participated at NIST in the Smartspace effort to help create a standard infrastructure for pervasive computing in intelligent rooms. He then went to the Spoken Language Processing group at LIMSI where he participated in system development for speech recognition and has been a prime contributor in speech understanding, named entity detection, question answering and dialogue systems.

Now at LNE, he is a co-leader of varied evaluations in the domain of speech recognition, speaker diariation, named entity detection and question answering. His current activities focus on annotation visualization and edition tools, evaluation tools and advanced metrics development. He is the author/co-author of over 30 refereed papers in journals and national and international conferences.

Mazin Gilbert (http://www.research.att.com/~{}mazin/) is the Executive Director of Speech and Language Technologies at AT&T Labs-Research. He has a Ph.D. in Electrical and Electronic Engineering, and an MBA for Executives from the Wharton Business School. Dr. Gilbert has over 20 years of research experience working in industry at Bell Labs and AT&T Labs and in academia at Rutgers University, Liverpool University, and Princeton University.

Dr. Gilbert is responsible for the advancement of AT&T's technologies in areas of interactive speech and multimodal user interfaces. This includes fundamental and forward looking research in automatic speech recognition, spoken language understanding, mobile voice search, multimodal user interfaces, and speech and web analytics.

He has over 100 publications in speech, language and signal processing and is the author of the book entitled, Artificial Neural Networks for Speech Analysis/Synthesis (Chapman & Hall, 1994). He holds 40 US patents and is a recipient of several national and international awards including the Most Innovative Award from SpeechTek 2003 and the AT&T Science and Technology Award, 2006.

He is a Senior Member of the IEEE; Board Member, LifeBoat Foundation (2010); Member, Editorial Board for

Signal Processing Magazine (2009–present); Member, ISCA Advisory Council (2007–present); Chair, IEEE/ACL workshop on Spoken Language Technology (2006); Chair, SPS Speech and Language Technical Committee (2004–2006); Teaching Professor, Rutgers University (1998–2001) and Princeton University (2004–2005); Chair, Rutgers University CAIP Industrial Board (2003–2006); Associate Editor, IEEE Transaction on Speech and Audio Processing (1995–1999); Chair, 1999 Workshop on Automatic Speech Recognition and Understanding; Member, SPS Speech Technical Committee (2000–2004); Technical Chair and Speaker for several international conferences including ICASSP, SpeechTek, AVIOS, and Interspeech.

Dilek Hakkani-Tür is a senior researcher at ICSI speech group. Prior to joining ICSI, she was a senior technical staff member in the Voice Enabled Services Research Department at AT&T Labs – Research at Florham Park, NJ. She received her BSc degree from Middle East Technical University, in 1994, and MSc and PhD degrees from Bilkent University, Department of Computer Engineering, in 1996 and 2000, respectively. Her PhD thesis is on statistical language modeling for agglutinative languages. She worked on machine translation during her visit to Carnegie Mellon University, Language Technologies Institute in 1997, and her visit to Johns Hopkins University, Computer Science Department, in 1998. In 1998 and 1999, she visited SRI International, Speech Technology and Research Labs, and worked on using lexical and prosodic information for information extraction from speech. In 2000, she worked in Natural Sciences and Engineering Faculty of Sabanci University, Turkey.

Her research interests include natural language and speech processing, spoken dialogue systems, and active and unsupervised learning for language processing. She has 10 patents and has co-authored more than 100 papers in natural language and speech processing. She is the receipent of three best paper awards for her work on active learning, from IEEE Signal Processing Society (with Giuseppe Riccardi), ISCA (with Gokhan Tur and Robert Schapire) and EURASIP (with Gokhan Tur and Robert Schapire). She is a member of ISCA, IEEE, Association for Computational Linguistics. She was an associate editor of IEEE Transactions on Audio, Speech and Language Processing between 2005 and 2008 and is an elected member of the IEEE Speech and Language Technical Committee (2009–2012) and a member of the HLT advisory board.

Timothy J. Hazen received the degrees of SB (1991), SM (1993), and PhD (1998) from the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology (MIT). From 1998 until 2007, Dr. Hazen was a Research Scientist in the Spoken Language Systems Group at the MIT Computer Science and Artificial Intelligence Laboratory. Since 2007, he has been a member of the Human Language Technology Group at MIT Lincoln Laboratory.

Dr. Hazen is a Senior Member of the IEEE and has served as an Associate Editor for the IEEE Transactions on Speech and Audio Processing (2004–2009) and as a member of the IEEE Signal Processing Society's Speech and Language Technical Committee (2008–2010). His research interests are in the areas of speech recognition and understanding, audio indexing, speaker identification, language identification, multi-lingual speech processing, and multi-modal speech processing.

Yun-Cheng Ju received a BS in electrical engineering from National Taiwan University in 1984 and a Master's and PhD in computer science from the University of

Illinois at Urbana-Champaign in 1990 and 1992, respectively. He joined Microsoft in 1994. His research interests include spoken dialogue systems, natural language processing, language modeling, and voice search. Prior to joining Microsoft, he worked at Bell Labs for two years. He is the author/co-authorof over 30 journal and conference papers and has filed over 40 US and international patents.

Lori Lamel is a senior CNRS research scientist in the Spoken Language Processing group at LIMSI which she joined in October 1991. She received her PhD degree in EECS in May 1988 from the Massachusetts Institute of Technology. Her principal research activities are in speech recognition; lexical and phonological modeling; spoken language systems and speaker and language identification. She has been a prime contributor to the LIMSI participations in DARPA benchmark evaluations and developed the LIMSI American English pronunciation lexicon.

She has been involved in many European projects and is currently leading the speech processing activities in the Quaero program. Dr. Lamel is a member of the Speech Communication Editorial Board and the Interspeech International Advisory Council. She was a member of the IEEE Signal Processing Society's Speech Technical Committee from 1994 to 1998, and the Advisory Committee of the AFCP, the IEEE James L. Flanagan Speech and Audio Processing Award Committee (2006–2009) and the EU-NSF Working Group for Spoken-word Digital Audio Collections. She has over 230 reviewed publications and is co-recipient of the 2004 ISCA Best Paper Award for a paper in the Speech Communication Journal.

Yang Liu received the degrees of BS and MS degrees from Tsinghua University, Beijing, China, in 1997 and

2000, respectively, and the PhD degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 2004.

She was a Researcher at the International Computer Science Institute, Berkeley, CA, from 2002 to 2005. She has been an Assistant Professor in Computer Science at the University of Texas at Dallas, Richardson, since 2005. Her research interests are in the area of speech and language processing.

I. Dan Melamed is a Principal Member of Technical Staff at AT&T Labs – Research. He holds a PhD in Computer and Information Science from the University of Pennsylvania (1998). He has over 40 publications in the areas of machine learning and natural language processing, including the book *Empirical Methods for Exploiting Parallel Texts* (MIT Press, 2001). Prior to joining AT&T, Dr. Melamed was a member of the computer science faculty at New York University.

Roberto Pieraccini has been at the leading edge of spoken dialogue technology for more than 25 years, both in research as well as in the development of commercial applications. He worked at CSELT, Bell Laboratories, AT&T Labs, SpeechWorks, IBM Research and he is currently the CTO of SpeechCycle. He has authored more than 120 publications in different areas of human–machine communication. Dr. Pieraccini is a Fellow of ISCA and IEEE.

Matthew Purver is a lecturer in Human Interaction in the School of Electronic Engineering and Computer Science at Queen Mary, University of London. His research interests lie in the computational semantics and pragmatics of dialogue, both for human/computer interaction and for the automatic understanding of natural human/human dialogue. From 2004 to 2008 he was a researcher at CSLI, Stanford University, where he worked on various dialogue system projects including the in-car CHAT system and the CALO meeting assistant.

Bhuvana Ramabhadran is the Manager of the Speech Transcription and Synthesis Research Group at the IBM T. J. Watson Center, Yorktown Heights, NY. Upon joining IBM in 1995, she made significant contributions to the ViaVoice line of products focusing on acoustic modeling including acoustics- based baseform determination, factor analysis applied to covariance modeling, and regression models for Gaussian likelihood computation.

She has served as the Principal Investigator of two major international projects: the NSF-sponsored MALACH Project, developing algorithms for transcription of elderly, accented speech from Holocaust survivors, and the EU-sponsored TC-STAR Project, developing algorithms for recognition of EU parliamentary speeches. She was the Publications Chair of the 2000 ICME Conference, organized the HLT-NAACL 2004 Workshop on Interdisciplinary Approaches to Speech Indexing and Retrieval, and a 2007 Special Session on Speech Transcription and Machine Translation at the 2007 ICASSP in Honolulu, HI. Her research interests include speech recognition algorithms, statistical signal processing, pattern recognition, and biomedical engineering.

Giuseppe Riccardi heads the Signal and Interactive Systems Lab at University of Trento, Italy. He received his Laurea degree in Electrical Engineering and Master in Information Technology, in 1991, from the University of Padua and CEFRIEL/Politechnic of Milan (Italy), respectively. From 1990 to 1993 he collaborated with Alcatel-Telettra Research Laboratories (Milan, Italy). In 1995 he received his PhD in Electrical Engineering from the Department of Electrical Engineering at the University of Padua, Italy. From 1993 to 2005, he was at AT&T Bell Laboratories and then AT&T Labs-Research where he worked in the Speech and Language Processing Lab. In 2005 joined the faculty of University of Trento (Italy). He is affiliated with Engineering School, the Department of Information Engineering and Computer Science and Center for Mind/Brain Sciences.

He has co-authored more than 100 papers and 30 patents in the field of speech processing, speech recognition, understanding and machine translation. His current research interests are language modeling and acquisition, language understanding, spoken/multimodal dialogue, affective computing, machine learning and machine translation.

Prof. Riccardi has been on the scientific and organizing committee of Eurospeech, Interspeech, ICASSP, NAACL, EMNLP, ACL an EACL. He has co-organized the IEEE ASRU Workshop in 1993, 1999, 2001 and was its General Chair in 2009. He has been the Guest Editor of the IEEE Special Issue on Speech-to-Speech Machine Translation. He has been a founder and Editorial Board member of the ACM Transactions of Speech and Language Processing. He has been elected member of the IEEE SPS Speech Technical Committee (2005–2008). He is a member of ACL, ISCA, ACM and Fellow of IEEE. He has received many national and international awards and more recently the Marie Curie Excellence Grant by the European Commission, 2009 IEEE SPS Best Paper Award and IBM Faculty Award.

Sophie Rosset is a senior CNRS researcher in the Spoken Language Processing group at LIMSI which she joined in May 1994. She received her PhD degree in Computer Science from the University Paris – Sud 11, France, in 2000. Her research activities focus mainly on interactive and spoken question-answering systems, including dialogue management and named entities detection.

She has been prime contributor to the LIMSI participations in QAST evaluations (QA©CLEF) and she is the leader for the spoken language processing group participation in the Quaero program evaluations for question-answering system on Web data and named entity detection. She is responsible of the Named Entity activities within the Quaero program and the French Edylex project. She has been involved in different European projects, most recently the Chil and Vital projects. She is author/co-author of over 60 refereed papers in journals and international conferences.

Murat Saraclar received his BS in 1994 from the Electrical and Electronics Engineering Department at Bilkent University and the degrees of MS in 1997 and PhD in 2001 from the Electrical and Computer Engineering Department at the Johns Hopkins University. He is an associate professor at the Electrical and Electronic Engineering Department of Bogazici University. From 2000 to 2005, he was with AT&T Labs – Research. His main research interests include all aspects of speech recognition, its applications, as well as related fields such as speech and language processing, human/computer interaction and machine learning. He was a member of the IEEE Signal Processing Society Speech and Language Technical Committee (2007–2009). He is currently serving as an associate editor for IEEE Signal Processing Letters and he is on the editorial boards of Computer Speech and Language, and Language Resources and Evaluation. He is a Member of the IEEE.

David Suendermann has been working on various fields of speech technology research over the last 10 years. He has worked at multiple industrial and academic institutions including Siemens (Munich), Columbia University (New York), USC (Los Angeles), UPC (Barcelona), RWTH (Aachen), and is currently the

Principal Speech Scientist of SpeechCycle. He has authored more than 60 publications and patents and holds a PhD from the Bundeswehr University in Munich.

Gokhan Tur was born in Ankara, Turkey in 1972. He received his BS, MS, and PhD degrees from the Department of Computer Science, Bilkent University, Turkey in 1994, 1996, and 2000, respectively. Between 1997 and 1999, he visited the Center for Machine Translation of CMU, then the Department of Computer Science of Johns Hopkins University, and then the Speech Technology and Research Lab of SRI International. He worked at AT&T Labs – Research from 2001 to 2006 and at the Speech Technology and Research (STAR) Lab of SRI International from 2006 to June 2010. He is currently with Microsoft working as a principal scientist. His research interests include spoken language understanding (SLU), speech and language processing, machine learning, and information retrieval and extraction. He has co-authored more than 75 papers published in refereed journals and presented at international conferences.

Dr. Tur is also the recipient of the Speech Communication Journal Best Paper awards by ISCA for 2004–2006 and by EURASIP for 2005–2006. Dr. Tur is the organizer of the HLT-NAACL 2007 Workshop on Spoken Dialog Technologies, and the HLT-NAACL 2004 and AAAI 2005 Workshops on SLU, and the editor of the Speech Communication Special Issue on SLU in 2006. He is also the Spoken Language Processing Area Chair for IEEE ICASSP 2007, 2008, and 2009 conferences, Spoken Dialog Area Chair for HLT-NAACL 2007 conference, Finance Chair for IEEE/ACL SLT 2006 and SLT 2010 workshops, and SLU Area Chair for IEEE ASRU 2005 workshop. Dr. Tur is a senior member of IEEE, ACL, and ISCA, and is currently an associate editor for the IEEE Transactions on Audio, Speech, and Language Processing journal, and was a member of IEEE Signal Processing Society (SPS), Speech and Language Technical Committee (SLTC) for 2006–2008.

Ye-Yi Wang received a BS in 1985 and an MS in 1988, both in in computer science from Shanghai Jiao Tong University, as well as an MS in computational linguistics in 1992 and a PhD in human language technology in 1998, both from Carnegie Mellon University. He joined Microsoft Research in 1998.

His research interests include spoken dialogue systems, natural language processing, language modeling, statistical machine translation, and machine learning. He served on the editorial board of the Chinese Contemporary Linguistic Theory series. He is a coauthor of Introduction to Computational Linguistics (China Social Sciences Publishing House, 1997), and he has published over 40 journal and conference papers. He is a Senior Member of IEEE.

Dong Yu joined Microsoft Corporation in 1998 and Microsoft Speech Research Group in 2002, where he is a researcher. He holds a PhD degree in computer science from the University of Idaho, an MS degree in computer science from Indiana University at Bloomington, an MS degree in electrical engineering from Chinese Academy of Sciences, and a BS degree (with honors) in electrical engineering from Zhejiang University (China). His current research interests include speech processing, robust speech recognition, discriminative training, spoken dialogue systems, voice search technology, machine learning, and pattern recognition. He has published more than 70 papers in these areas and is the inventor/coinventor of more than 40 granted/pending patents.

Dr. Dong Yu is a senior member of IEEE, a member of ACM, and a member of ISCA. He is currently serving as an associate editor of IEEE signal processing magazine and the lead guest editor of IEEE Transactions on Audio, Speech, and Language Processing – Special Issue on Deep Learning for Speech and Language Processing. He is also serving as a guest professor at the University of Science and Technology of China.

Foreword

Speech processing has been an active field of research and development for more than a half-century. While including technologies such as coding, recognition and synthesis, a long-term dream has been to create machines which are capable of interacting with humans by voice. This implies the capability of not merely recognizing what is said, but of understanding the meaning of spoken language. Many of us believe such a capability would fundamentally change the manner in which people use machines.

The subject of understanding and meaning has received much attention from philosophers over the centuries. When one person speaks with another, how can we know whether the intended message was understood? One approach is via a form of the Turing Test: evaluate whether the communication was correctly understood on the basis of whether the recipient responded in an expected and appropriate manner. For example, if one requested, from a cashier, change of a dollar in quarters, then one evaluates whether the message was understood by examining the returned coins. This has been distinguished as linguistic performance, i.e. the actual use of language in concrete actions.

This new book, compiled and edited by Tur and De Mori, describes and organizes the latest advances in spoken language understanding (SLU). They address SLU for human/machine interaction and for exploiting large databases of spoken human/human conversations.

While there are many textbooks on speech or natural language processing, there are no previous books devoted wholly to SLU. Methods have been described piece meal in other books and in many scientific publications, but never gathered together in one place with this singular focus. This