

Computational Linguistics

Articles

Extracting the Lowest-Frequency Words: Pitfalls and Possibilities	Marc Weeber, Rein Vos, and R. Harald Baayen	301
Incremental Processing and Acceptability	Glyn Morrill	319
Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech	Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer.	339
A Compression-based Algorithm for Chinese Word Segmentation	W. J. Teahan, Yingying Wen, Rodger McNab, Ian H. Witten	375
The Rhetorical Parsing of Unrestricted Texts: A Surface-based Approach	Daniel Marcu	395

Book Reviews

Foundations of Computational Linguistics: Man-Machine Communication in Natural Language by Roland Häusser	<i>Reviewed by</i> Alexander F. Gelbukh	449
Syntactic Wordclass Tagging edited by Hans van Halteren	<i>Reviewed by</i> Adwait Ratnaparkhi	456
Natural Language Information Retrieval edited by Tomek Strzalkowski	<i>Reviewed by</i> Simon Corston-Oliver	460
The MIT Encyclopedia of the Cognitive Sciences edited by Robert A. Wilson and Frank C. Keil	<i>Reviewed by</i> John Nerbonne	463
Publications Received		468
Erratum		470



Editor 1993–

Julia Hirschberg
AT&T Labs-Research
180 Park Avenue, A257
Florham Park, NJ 07932-0971
(+1 973) 360-8330
acl@research.att.com

Editorial Assistant

Ana Pereira
AT&T Labs-Research
Room A240
180 Park Avenue, P.O. Box 971
Florham Park, NJ 07932-0971
acl@research.att.com
fax: (973) 360-8970

Book Review Editor

Graeme Hirst
University of Toronto
Department of Computer Science
Toronto, CANADA M5S 3G4
(+1 416) 978-8747
gh@cs.toronto.edu

Squibs and Discussions Editor

Pierre Isabelle
Xerox Research Centre Europe
6, chemin de Maupertuis
38240 Meylan, France
+33 4 76 61 51 98
Pierre.Isabelle@xrce.xerox.com

Editorial Board Members

1998–2000:

Steven Abney, *AT&T Labs-Research*
Chris Brew, *University of Edinburgh*
Jean Carletta, *University of Edinburgh*
Genichiro Kikui, *NTT Information and Communication Systems Labs*
Shalom Lappin, *King's College London*
Christopher Manning, *University of Sydney*
Christine Nakatani, *AT&T Labs-Research*
Philip Resnik, *University of Maryland*
Giorgio Satta, *University of Padua*
Karen Sparck-Jones, *University of Cambridge*
Dekai Wu, *Hong Kong University of Science and Technology*

1999–2001:

Michael Elhadad, *Ben Gurion University*
Kôiti Hasida, *Electrotechnical Laboratory*
Marti Hearst, *University of California at Berkeley*
Andrew Kehler, *SRI International, Menlo Park*
Daniel Jurafsky, *University of Colorado*
Kevin Knight, *USC Information Sciences Institute*
Elizabeth Shriberg, *SRI International, Menlo Park*

2000–2002:

Gosse Bouma, *University of Groningen*
Michael Collins, *AT&T Labs-Research*
Jennifer Chu-Carroll, *Lucent Technologies Bell Labs*
Stanley F. Chen, *IBM/TJ Watson Research Center*
Dan Roth, *University of Illinois, Urbana/Champaign*
Lillian Lee, *Cornell University*
Mark-Jan Nederhof, *DFKI, Saarbrücken*

Computational Linguistics

Articles

Extracting the Lowest-Frequency Words: Pitfalls and Possibilities	Marc Weeber, Rein Vos, and R. Harald Baayen	301
Incremental Processing and Acceptability	Glyn Morrill	319
Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech	Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer	339
A Compression-based Algorithm for Chinese Word Segmentation	W. J. Teahan, Yingying Wen, Rodger McNab, Ian H. Witten	375
The Rhetorical Parsing of Unrestricted Texts: A Surface-based Approach	Daniel Marcu	395

Book Reviews

Foundations of Computational Linguistics: Man-Machine Communication in Natural Language by Roland Hausser	<i>Reviewed by</i> Alexander F. Gelbukh	449
Syntactic Wordclass Tagging edited by Hans van Halteren	<i>Reviewed by</i> Adwait Ratnaparkhi	456
Natural Language Information Retrieval edited by Tomek Strzalkowski	<i>Reviewed by</i> Simon Corston-Oliver	460
The MIT Encyclopedia of the Cognitive Sciences edited by Robert A. Wilson and Frank C. Keil	<i>Reviewed by</i> John Nerbonne	463
Publications Received		468
Erratum		470

Copyright ©2000 by the Association for Computational Linguistics. Permission to photocopy articles for internal or personal use, or the internal or personal use of specific clients, is granted by the copyright owner for users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the fee of \$3.00 per article-copy is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The fee code for users of the Transactional Reporting Service is 0891-2017/00 \$3.00. For those organizations that have been granted a photocopy license with CCC, a separate system of payment has been arranged. Such permission requests and other permission inquiries should be addressed to the Managing Editor at the address given on the inside front cover of this issue.

Computational Linguistics is abstracted and/or indexed in the following publications: *Annual Review of Information Science and Technology*, *CompuMath Citation Index*, *Computer Abstracts*, *Computing Reviews*, *Current Contents: Social and Behavioral Sciences*, *Engineering Index*, *Knowledge Engineering Review*, *Language and Language Behavior Abstracts*, *Linguistic Abstracts*, *Linguistic Bibliography*, *Research Alert*, *Science Abstracts* (INSPEC Section C), *SciSearch*, *Social Sciences Citation Index*, and *Social SciSearch*.

Volume 14 and subsequent volumes of *Computational Linguistics* are available on microfilm through University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA.

Send advertising and mailing list inquiries to Marketing Dept., MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142-1407; Phone: (617) 253-2866; fax: (617) 258-5028; e-mail: journals-info@mit.edu.

Individual Subscriptions to *Computational Linguistics* are available only with membership in the Association for Computational Linguistics (ACL). Annual Membership Dues are as follows: *regular*, \$50.00; *student* — taking a full-time program of courses and not earning a regular income, \$30.00; *retired*, \$30.00; *joint* — for couples receiving one subscription, \$60.00. A membership application together with an order form for back issues and conference proceedings is included at the end of this issue. Send membership applications and address changes to Priscilla Rasmussen, Association for Computational Linguistics (ACL), P.O. Box 6090, Somerset, NJ 08875, USA. However, note that it is also possible to pay by credit card (both by e-mail and by postal mail) and through banks in Switzerland and Japan.

Institutional Subscriptions are \$128.00. Outside the U.S. and Canada add \$18.00 postage/handling. Canadian subscribers add 7% GST. Single copies of current and back issues are \$33.00. Outside U.S. and Canada add \$5.00 per issue for postage and handling. Canadians add 7% GST. Send institutional subscription orders and address changes to MIT Press Journals, Five Cambridge Center, Cambridge, MA 02142-1407; phone: (617) 253-2889; fax: (617) 577-1545; e-mail: journals-orders@mit.edu. Claims for missing issues will be honored free of charge if made within three months after the publication date of the issue. Claims may be emailed to: journals-claims@mit.edu. Prices subject to change without notice. Institutions should order back issues before 1988 and all proceedings from the ACL at the address above.

<http://mitpress.mit.edu/COLI>