

# Frontiers in Corpus Annotations II



## Proceedings of the Workshop

ACL 2005  
Ann Arbor, Michigan  
June 29, 2005

Production and Manufacturing by  
*Omnipress Inc.*  
*Post Office Box 7214*  
*Madison, WI 53707-7214*

©2005 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
75 Paterson Street, Suite 9  
New Brunswick, NJ 08901  
USA  
Tel: +1-732-342-9100  
Fax: +1-732-342-9339  
[acl@aclweb.org](mailto:acl@aclweb.org)

## Preface

This volume contains the papers prepared for and presented at the *Frontiers in Corpus Annotation II: Pie in the Sky* workshop held on June 29, 2005 at the University of Michigan, as part of the 2005 annual meeting of the Association of Computational Linguistics.

This workshop is the second in a series of workshops about innovation in corpus annotation and its effect on the future of Computational Linguistics.

I wish to thank and acknowledge the program committee, participants in the workshop, those who contributed and submitted papers, those who participated in the “Pie in the Sky” email list and my wife Jenny for the cover illustration.

Adam Meyers  
Computer Science Dept.  
New York University

**Chair:** Adam Meyers, New York University

**Program Committee:**

Charles J. Fillmore, International Computer Science Institute, Berkeley  
Eva Hajicova, Center for Computational Linguistics, Charles University, Prague  
Boyan A. Onyshkevych, U.S. Dept. of Defense  
Martha Palmer, University of Pennsylvania, Philadelphia  
David Farwell, Computing Research Laboratory, New Mexico State University, Las Cruces, NM  
Sergei Nirenburg, University of Maryland, Baltimore County  
Owen Rambow, Columbia University, NYC  
Beth Sundheim, SPAWAR Systems Center, San Diego  
Gerald Penn, University of Toronto, Toronto  
James Pustejovsky, Brandeis University, Waltham, Mass.

**Additional Reviewer:** Chadwick McHenry, MITRE

**Cover Illustration:** Jennifer Eve Meyers

**Conference Website:** <http://nlp.cs.nyu.edu/meyers/frontiers/2005.html>

**Pie in the Sky Website:** <http://nlp.cs.nyu.edu/meyers/pie-in-the-sky.html>

## Table of Contents

<i>Introduction to Frontiers in Corpus Annotation II: Pie in the Sky</i> Adam Meyers .....	1
<i>Merging PropBank, NomBank, TimeBank, Penn Discourse Treebank and Coreference</i> James Pustejovsky, Adam Meyers, Martha Palmer and Massimo Poesio .....	5
<i>A Unified Representation for Morphological, Syntactic, Semantic, and Referential Annotations</i> Erhard W. Hinrichs, Sandra Kübler and Karin Naumann .....	13
<i>Parallel Entity and Treebank Annotation</i> Ann Bies, Seth Kulick and Mark Mandel .....	21
<i>Attribution and the (Non-)Alignment of Syntactic and Discourse Arguments of Connectives</i> Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi and Bonnie Webber	29
<i>Investigating the Characteristics of Causal Relations in Japanese Text</i> Takashi Inui and Manabu Okumura .....	37
<i>A Framework for Annotating Information Structure in Discourse</i> Sasha Calhoun, Malvina Nissim, Mark Steedman and Jason Brenier .....	45
<i>Annotating Attributions and Private States</i> Theresa Wilson and Janyce Wiebe .....	53
<i>A Parallel Proposition Bank II for Chinese and English</i> Martha Palmer, Nianwen Xue, Olga Babko-Malaya, Jinying Chen and Benjamin Snyder .....	61
<i>Semantically Rich Human-Aided Machine Annotation</i> Marjorie McShane, Sergei Nirenburg, Stephen Beale and Thomas O’Hara .....	68
<i>The Reliability of Anaphoric Annotation, Reconsidered: Taking Ambiguity into Account</i> Massimo Poesio and Ron Artstein .....	76
<i>Annotating Discourse Connectives in the Chinese Treebank</i> Nianwen Xue .....	84



## Conference Program

**Wednesday, June 29, 2005, 9:00–18:00**

- 9:00–9:30     *Introduction to Frontiers in Corpus Annotation II: Pie in the Sky*  
Adam Meyers
- 9:30–10:00    *Merging PropBank, NomBank, TimeBank, Penn Discourse Treebank and Coreference*  
James Pustejovsky, Adam Meyers, Martha Palmer and Massimo Poesio
- 10:00–10:30    *A Unified Representation for Morphological, Syntactic, Semantic, and Referential Annotations*  
Erhard W. Hinrichs, Sandra Kübler and Karin Naumann
- 10:30–11:00    Break
- 11–11:30       *Parallel Entity and Treebank Annotation*  
Ann Bies, Seth Kulick and Mark Mandel
- 11:30–12:00    *Attribution and the (Non-)Alignment of Syntactic and Discourse Arguments of Connectives*  
Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi and Bonnie Webber
- 12–12:30       Discussion
- 12:30–14:00    Lunch
- 14:00–14:30    *Investigating the Characteristics of Causal Relations in Japanese Text*  
Takashi Inui and Manabu Okumura
- 14:30–15:00    *A Framework for Annotating Information Structure in Discourse*  
Sasha Calhoun, Malvina Nissim, Mark Steedman and Jason Brenier
- 15:00–15:30    *Annotating Attributions and Private States*  
Theresa Wilson and Janyce Wiebe
- 15:30–16:00    Break
- 16:00–16:30    *A Parallel Proposition Bank II for Chinese and English*  
Martha Palmer, Nianwen Xue, Olga Babko-Malaya, Jinying Chen and Benjamin Snyder

**Wednesday, June 29, 2005, 9:00–18:00 (continued)**

- 16:30–17:00 *Semantically Rich Human-Aided Machine Annotation*  
Marjorie McShane, Sergei Nirenburg, Stephen Beale and Thomas O’Hara
- 17:00–17:30 *The Reliability of Anaphoric Annotation, Reconsidered: Taking Ambiguity into Account*  
Massimo Poesio and Ron Artstein
- 17:30–18:00 Discussion