

# The MIT Press

## Journals

Sign In / Register

Books

Journals

Digital

Resources

About

Contact

Home | Computational Linguistics | List Article navigation of Issues | Volume 40 , No. 4 | Reflections on the Penn Discourse TreeBank, Comparable Corpora, and Complementary Annotation



# Reflections on the Penn Discourse TreeBank, Comparable Corpora, and Complementary Annotation

Quarterly (March, June, September, December)

160pp. per issue

6 3/4 x 10

Founded: 1974

2018 Impact Factor: 1.319

2018 Google Scholar h5-index: 32

ISSN: 0891-2017

E-ISSN: 1530-9312

### Journal Resources

- Editorial Info
- Abstracting and Indexing
- Release Schedule
- Advertising Info

### Author Resources

- Submission Guidelines

Rashmi Prasad, Bonnie Webber and Aravind Joshi

Posted Online November 12, 2014  
[https://doi.org/10.1162/COLI\\_a\\_00204](https://doi.org/10.1162/COLI_a_00204)

© 2014 Association for Computational Linguistics

Computational Linguistics  
Volume 40 | Issue 4 | December 2014  
p.921-950

**Download Options** >


**Abstract** Full Text Authors

The Penn Discourse Treebank (PDTB) was released to the public in 2008. It remains the

Publication  
Agreement  
Author Reprints

## Reader Resources

Rights and Permissions  
Most Read  
Most Cited

More About Computational Linguistics 

Metrics 



7 Total citations

7 Recent citations

4.09 Field Citation Ratio  
n/a Relative Citation Ratio

Open Access 



Computational Linguistics Computational Linguistics is Open Access. All content is freely available in electronic format (Full text HTML, PDF, and PDF Plus) to


largest manually annotated corpus of discourse relations to date. Its focus on discourse relations that are either lexically-grounded in explicit discourse connectives or associated with sentential adjacency has not only facilitated its use in language technology and psycholinguistics but also has spawned the annotation of comparable corpora in other languages and genres.


Given this situation, this paper has four aims: (1) to provide a comprehensive introduction to the PDTB for those who are unfamiliar with it; (2) to correct some wrong (or perhaps inadvertent) assumptions about the PDTB and its annotation that may have weakened previous results or the performance of decision procedures induced from the data; (3) to explain variations seen in the annotation of comparable resources in other languages and genres, which should allow developers of future comparable resources to recognize whether the variations are relevant to them; and (4) to enumerate and explain relationships between PDTB annotation and complementary annotation of other linguistic phenomena. The paper draws on work done by ourselves and others since the corpus was released.


## Forthcoming

### Most Read

[See More](#)

 **Lexicon-Based Methods for Sentiment Analysis** (13965 times)  
Maite Taboada et al.  
Computational Linguistics  
Volume: 37, Issue: 2, pp. 267-307

 **Computational Linguistics and Deep Learning** (10500 times)  
Christopher D. Manning  
Computational Linguistics  
Volume: 41, Issue: 4, pp. 701-707


 **Near-Synonymy and Lexical Choice** (3653 times)  
Philip Edmonds et al.  
Computational Linguistics  
Volume: 28, Issue: 2, pp. 105-144


(Note that the Most Read numbers are based on the number of full text downloads over the last 12 months.)


### Most Cited

[See More](#)

readers across the globe. All articles are published under a [CC BY-NC-ND 4.0 license](#). For more information on allowed uses, please view the [CC license](#). [Support OA at MITP](#)

 **Lexicon-Based Methods for Sentiment Analysis** (436 times)  
Maite Taboada et al.  
Computational Linguistics  
Volume: 37, Issue: 2, pp. 267-307



 **A Systematic Comparison of Various Statistical Alignment Models** (174 times)  
Franz Josef Och et al.  
Computational Linguistics  
Volume: 29, Issue: 1, pp. 19-51

 **Opinion Word Expansion and Target Extraction through Double Propagation** (147 times)  
Guang Qiu et al.  
Computational Linguistics  
Volume: 37, Issue: 1, pp. 9-27

(Note that the Most Cited numbers are based on Crossref's [Cited-by service](#) and reflect citation information for the past 24 months. )

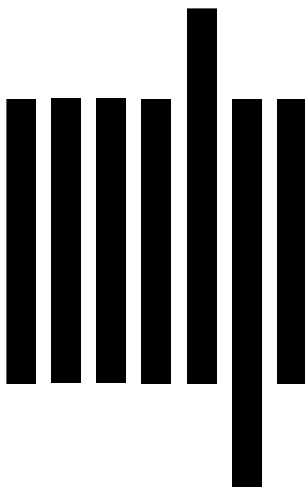
### Download > Options

Favorite  Sign up for Alerts 

Download Citation  RSS TOC 

RSS Citation  Submit your article

[Support OA at MITP](#) 



Journals

Terms & Conditions

Privacy Statement

Contact Us

Books

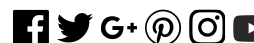
US

UK

Connect

One Rogers Street  
Cambridge MA  
02142-1209

Suite 2, 1 Duchess  
Street London,  
W1W 6AN, UK



© 2018 The MIT Press  
Technology Partner:  
[Atypon Systems, Inc.](#)  
[CrossRef Member](#)  
[COUNTER Member](#)  
The MIT Press colophon is registered in the

U.S. Patent and Trademark Office

