

**Contact Details**

Email(s): yogo@cs.biu.ac.il yoav.goldberg@gmail.com  
Homepage: http://www.cs.biu.ac.il/~yogo/

**Current Position**

Since February 2013 Senior Lecturer, Computer Science, Bar Ilan University.

2011 – February 2013 Research Scientist (post-doc), Google Research New York.

**Education**

2007-2011 **PhD** in Computer Science (Natural Language Processing)  
Thesis: “*Automatic Syntactic Processing of Modern Hebrew.*”  
Department of Computer Science  
Ben Gurion University, Beer Sheva, Israel  
Thesis advisor: Prof. Michael Elhadad

2005-2007 **MSc** (*Magna cum laude*) in Computer Science, Ben Gurion University

2002-2005 **BSc** in Computer Science, Computer Science department  
Ben Gurion University of the Negev, Beer Sheva, Israel

**Scientific Interests (highlights)**

- Natural Language Processing, including Hebrew, Syntax (Parsing and Machine Translation), Coordination, Morphology, Efficient Computation, Linguistic Creativity.
- Machine Learning, including Structured Prediction, Feature Representation, Unsupervised and Semi-Supervised learning, Confidence Estimation, Imitation Learning.
- Word-representation, word embeddings, neural networks for language processing.
- Sentence Processing and Syntactic-representation in humans.

**Grants**

My work is funded by:

- **DIP (German-Israeli Project Cooperation)**
- **Google Research Award (Award)**
- **AI2 (Sponsored Research Agreement)**
- **Intel ICRI-CI (Sponsored Research Agreement)**

**Honors and Awards**

- **Best Paper Award** (CoNLL 2014). Work with Omer Levy.
- **Best Student-Paper Award** (ACL-2012 Student-research workshop). Work with Raphael Cohen and Michael Elhadad.
- **Best Paper Award** (RECOMB-2011). Work with Shay Zakov, Michal Ziv-Ukelson and Michael Elhadad.

- **Fulbright post-doctoral fellowship** (2011). The United States-Israel Educational Foundation (declined, as I ended up taking an offer from Google Research).
- **Outstanding Teaching Assistant Award**. Computer Science Department, Ben Gurion University (2011).
- **NEGEV Scholarship** for Outstanding PhD students. Ben Gurion University.
- **Wolf Prize** for Outstanding MSc students. The Wolf Foundation.
- **Yossi Friedman Award** for incoming MSc students, based on outstanding undergraduate project. Department of Computer Science, Ben Gurion University.

### Academic Professional Activities

- **Co-Organizer** of the 5<sup>th</sup> Workshop of Parsing Morphologically-rich Languages (with Ines Rehbein, Yanick Versley and Yuval Marton). Workshop was held in conjunction with COLING 2014.
- **Co-Organizer** of the 4<sup>th</sup> Workshop of Parsing Morphologically-rich Languages (with Ines Rehbein, Yanick Versley and Yuval Marton). Workshop was held in conjunction with EMNLP 2013.
- **Area Chair** for EMNLP 2013 Syntax and parsing (co-chaired with Jennifer Foster).
- **Co-Organizer** of NAACL 2010 Workshop on Statistical Parsing of Morphologically Rich Languages.
- **Reviewer** for EACL (2011, 2012, 2013, 2014), CoNLL (2010, 2011, 2012, 2013, 2014), ACL (2011, 2012, 2013, 2014, 2015), IJCAI (2011, 2013), EMNLP (2011, 2012, 2013, 2015), IWPT (2011, 2013), NAACL (2012, 2013, 2014), COLING (2012, 2014).
- **Reviewer** for Computational Linguistics Journal and for AI Journal.
- **Guest Editorial Board**, Computational Linguistics Journal special issue on parsing morphologically rich languages.
- **Program Committee member**, ESSLLI 2011 Workshop on Lexical Resources.
- **Co-Organizer** of ISCOL 2008, 2013, 2015 (Israeli Seminar on Computational Linguistics).
- **Co-Organizer** of BGU Computer-science weekly Graduate Seminar (2008-2010).
- I was a **contributor** to the NLTK open-source project.

### Publications

My h-index is 19, and my i10-index is 26.

For an up-to-date list of publications, see my Google Scholar page.

#### Journal Papers

*“Improving distributional similarity with lessons learned from word embeddings”*

Omer Levy, **Yoav Goldberg** and Ido Dagan.

Transactions of the Association for Computational Linguistics 3, 211-225.

*“Constrained Arc-Eager Dependency Parsing”*

Joakim Nivre, **Yoav Goldberg** and Ryan McDonald.

Computational Linguistics. June 2014.

*“A Tabular Method for Dynamic Oracles in Transition-based Parsing”.*

**Yoav Goldberg**, Giorgio Satta and Francesco Sartorio

Transactions of the Association for Computational Linguistics.

April 2014.

*“Training Deterministic Parsers with Non-Deterministic Oracles.”*

**Yoav Goldberg** and Joakim Nivre.

Transactions of the Association for Computational Linguistics.

October 2013.

*“Word-segmentation, Unknown Word Resolution and Morphological Agreement in a Hebrew Parsing System.”*

**Yoav Goldberg** and Michael Elhadad.

Computational Linguistics 39.1 (2013) 121-160

*“Rich Parameterization improves RNA Structure Prediction.”*

Shay Zakov, **Yoav Goldberg**, Michael Elhadad and Michal Ziv-Ukelson.

Journal of Computational Biology 18.11 (2011) 1525-1542

### Conferences

#### **2015**

*“Template Kernels for Dependency Parsing”*

Hillel Taub-Tabib, **Yoav Goldberg** and Amir Globerson

NAACL 2015

#### **2014**

*“Neural Word Embeddings as Implicit Matrix Factorization”*

Omer Levy and Yoav Goldberg

NIPS 2014

*“Dependency Based Word Embeddings”*

Omer Levy and Yoav Goldberg

ACL 2014

*“Automatic Detection of Machine Translated Text and Translation Quality Estimation “*

Roei Aharoni, Moshe Koppel and **Yoav Goldberg**

ACL 2014

*“Linguistic Regularities in Sparse and Explicit Word Representations”*

Omer Levy and **Yoav Goldberg**

CONLL 2014. **Best Paper Award**

#### **2013**

*“Dynamic-oracle Transition-based Parsing with Calibrated Probabilistic Output.”*

**Yoav Goldberg**

IWPT 2013

*“A Non-Monotonic Arc-Eager Transition System for Dependency Parsing.”*

Matthew Honnibal, **Yoav Goldberg** and Mark Johnson.

CONLL 2013

*“Efficient Implementation of Beam-Search Incremental Parsers.”*

**Yoav Goldberg**, Kai Zhao and Liang Huang.

ACL 2013.

*"A Dataset of Syntactic-Ngrams over Time from a Very Large Corpus of English Books."*

**Yoav Goldberg** and Jon Orwant.

\*SEM 2013.

*"Universal Dependency Annotation for Multilingual Parsing."*

Ryan McDonald, Joakim Nivre, Yvonne Quirnbach-Brundage, **Yoav Goldberg**, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló, Jungmee Lee

ACL 2013.

## 2012

*"A Dynamic Oracle for Arc-Eager Dependency Parsing."*

**Yoav Goldberg and Joakim Nivre**

Coling 2012.

## 2011

*"Rich Parameterization improves RNA Structure Prediction."*

Shay Zakov, **Yoav Goldberg**, Michael Elhadad and Michal Ziv-Ukelson.

RECOMB 2011. **Best Paper Award**

*"Language Independent Parsing with Empty Elements."*

Shu Cai, David Chiang and **Yoav Goldberg**.

ACL 2011

*"Joint Hebrew Segmentation and Parsing using a PCFG/PCFG-Lattice Parser."*

**Yoav Goldberg** and Michael Elhadad.

ACL 2011

## 2010

*"Inspecting the Structural Biases of Dependency Parsing Algorithms."*

**Yoav Goldberg** and Michael Elhadad.

CONLL 2010

*"An Efficient Algorithm for Easy-First Non-Directional Dependency Parsing."*

**Yoav Goldberg** and Michael Elhadad.

NACCL 2010

## 2009

*"Hebrew Dependency Parsing: Initial Results".*

**Yoav Goldberg** and Michael Elhadad.

IWPT 2009 (Short Paper)

*"On the Role of Lexical Features in Sequence Labeling".*

**Yoav Goldberg** and Michael Elhadad.

EMNLP 2009

*“Enhancing Unlexicalized Parsing Performance using a Wide Coverage Lexicon, Fuzzy Tag-set Mapping, and EM-HMM-based Lexical Probabilities”.*

**Yoav Goldberg**, Reut Tsarfaty, Meni Adler and Michael Elhadad.  
EACL 2009.

## 2008

*“EM Can Find Pretty Good HMM POS-Taggers (When Given a Good Start)”.*

**Yoav Goldberg**, Meni Adler and Michael Elhadad.  
ACL 2008

*A Single Generative Model for Joint Morphological Segmentation and Syntactic Parsing”.*

**Yoav Goldberg** and Reut Tsarfaty.  
ACL 2008.

*“Unsupervised Lexicon-Based Resolution of Unknown Words for Full Morphological Analysis”.*

Meni Adler, **Yoav Goldberg**, David Gabay and Michael Elhadad.  
ACL 2008.

*“splitSVM: Fast, Space-Efficient, non-Heuristic, Polynomial Kernel Computation for NLP Applications”.*

**Yoav Goldberg** and Michael Elhadad.  
ACL 2008 (Short Paper)

*“Tagging a Hebrew Corpus: The Case of Participles”.*

Meni Adler, Yael Netzer, **Yoav Goldberg**, David Gabay and Michael Elhadad.  
LREC 2008.

*“Word-Based or Morpheme-Based? Annotation Strategies fro Modern Hebrew Clitics”.*

Reut Tsarfaty and **Yoav Goldberg**.  
LREC 2008

*“Identification of Transliterated Foreign Words in Hebrew Script”.*

**Yoav Goldberg** and Michael Elhadad.  
CICLING 2008

## 2007

*“SVM Model Tampering and Anchored Learning: A Case Study in Hebrew NP Chunking”.*

**Yoav Goldberg** and Michael Elhadad.  
ACL 2007

## 2006

*“Noun Phrase Chunking in Hebrew: Influence of Lexical and Morphological Features”.*

**Yoav Goldberg**, Meni Adler and Michael Elhadad.  
ACL 2006

Workshop Papers (peer reviewed)

*“Intermediary Semantic Representation Through Proposition Structures”*

Gabriel Stanovsky, Jessica Ficler, Ido Dagan and **Yoav Goldberg**

ACL 2014 Semantic Parsing Workshop

*“Domain adaptation of a dependency parser with a class-class selectional preference model ”*

Raphael Cohen, **Yoav Goldberg** and Michael Elhadad

ACL 2012 Student-Research Workshop. **Best Student Paper Award.**

*“Statistical Parsing of Morphologically Rich Languages: What, How and Whither.”*

Reut Tsarfaty, Djame Seddah, **Yoav Goldberg**, Sandra Kubler, Marie Candito, Jenniffer Foster,

Yannick Versley, Ines Rehbein, and Lamia Tounsi.

NAACL 2010 Workshop on Parsing Morphologically Rich Languages.

*“Easy First Dependency Parsing of Modern Hebrew.”*

**Yoav Goldberg** and Michael Elhadad.

NAACL 2010 Workshop on Parsing Morphologically Rich Languages.

*“Gaiku: Generating Haiku with Word Association Norms”.*

Yael Netzer, David Gabay, **Yoav Goldberg** and Michael Elhadad.

*CALC-2009 (NAACL Workshop on Computational Approaches to Linguistic Creativity)*

*“Ontology Evaluation Through Text Classification”.*

Yael Netzer, David Gabay, Meni Adler, **Yoav Goldberg** and Michael Elhadad.

ENQOIR 2009

*“Generating Language from BlissSymbols Using Semantic Authoring”.*

Yael Netzer, Ofer Biller, Michael Elhadad and **Yoav Goldberg.**

*ISSAC 2006*

**Academic Internships / Research Assistantship**

Summer 2010:

**Internship at USC/ISI**, working with Kevin Knight, David Chiang and Liang Huang on Minimized Models for Syntax-based Machine Translation.

2008 - 2011:

**Research Assistant**, Deutsche Telekom laboratories at Ben Gurion University.

Work on various projects related to Semantic Search, Faceted Search, Sentiment Analysis and Ontology-based reasoning.

**Teaching**

As Senior Lecturer in Bar Ilan University

- Developed and teaching a graduate level elective course on Machine Translation.
- Developed and teaching the Introduction to Natural Language Processing course together with Prof. Ido Dagan.
- Teaching Introduction to Object Oriented Programming course (the course went through a major re-design in order to better fit current industry needs).

### During Msc and Phd studies at Ben Gurion University

Spring 2011 - Lab Instructor for System Programming Lab, BGU  
- Design and TA a course on scalable computing in cloud environments, BGU  
Fall 2011 - Head Teaching Assistant for System Programming, BGU  
Spring 2010 - Lab Instructor for System Programming Lab, BGU  
Fall 2010 - Teaching Assistant for System Programming, BGU  
Spring 2009 - Lab Instructor for System Programming Lab, BGU  
- Grader for Machine Learning, BGU  
Fall 2009 - Teaching Assistant for System Programming, BGU  
Spring 2008 - Lab Instructor for System Programming Lab, BGU  
- Teaching Assistant for System Programming for IAF, BGU/IAF  
Fall 2008 - Teaching Assistant for System Programming, BGU  
Spring 2007 - Lab Instructor for System Programming Lab, BGU  
Fall 2007 - Teaching Assistant for System Programming, BGU  
Spring 2006 - Teaching Assistant for Object Oriented Software Design, BGU  
Fall 2006 - Grader for Compiler Construction, BGU

Teaching assistantship include giving weekly frontal recitation sessions, as well as preparation of teaching materials and assignments. Head TA responsibilities include management, coordination and guidance of a group of 6 TAs and several graders.

I constantly get high teaching evaluations from students.

I supervised 4 undergraduate students on 3 final projects. Two of the works resulted in publishable results (in submission).

### **Non-Academic/Non-Research Professional Experience**

#### Since 2013

I consult to various companies on topics in Machine Learning, Natural Language Processing, and Data Analysis.

#### 2004-2006:

**Consultant**, Picscout inc., Israel.

My main role there was the design and creation of a web crawler which can handle dynamic web pages and flash content. I also advised them on many Internet / security related issues.

Later, I worked on the foundations of the *picapp* engine, as well as on an NLP pilot of attaching relevant pictures to SMS messages.

#### 2002-2005:

**Consultant**, IDF Intel. Corp.

#### 1999-2002:

Military service in the IDF Intel. Corp, technology unit.

I worked as a programmer and an analyst on various teams and projects.

During my last year I was a **team-leader of 5 people**.

### **Some Past Projects**

- Hebrew Dialect of the Scheme Programming Language (joint work with Dr. Mayer Goldberg). This is intended as a foundation for a course teaching computer-science concepts to Israeli school-kids.

- Design and Implementation of a Blissymbolics lexicon software (part of a PhD thesis by Yael Netzer). Blissymbolics is a symbol based Augmentative Communication language. The lexicon (and the special queries it provides) is used for building intelligent communication boards for Bliss users.

### **Technical Skills**

- Programming Languages: Python, Java, C#, Perl, Scheme, 8086-Assembly, C++, C, Scala.
- Various "web" languages: Javascript, CSS, DHTML, PHP etc, with a security perspective.
- Expertise in computer security, esp. of web-based applications and web related technologies.

### **Languages**

Hebrew	native
English	fluent