

# Yonatan Bisk

(Grounded) Natural Language Processing  
Computer Science

✉ ybisk@cs.cmu.edu

## Positions

- 2020 **Assistant Professor**, *Carnegie Mellon University*  
Language Technologies Institute (primary) & Robotics Institute (courtesy)
- ... **Visiting Researcher**, *Microsoft Research* w/ Jianfeng Gao
- 2019 **Postdoctoral Scholar**  
University of Washington & Allen Institute for Artificial Intelligence w/ Yejin Choi  
Information Sciences Institute, University of Southern California w/ Daniel Marcu

## Education

- 2015 **PhD in Computer Science**, *The University of Illinois at Urbana-Champaign*,  
Thesis: **Unsupervised Grammar Induction with Combinatory Categorical Grammars**  
Advisor: Julia Hockenmaier  
Committee: Jason Eisner, Dan Roth, ChengXiang Zhai
- 2009 **BS in Computer Science**, *The University of Texas at Austin*, Turing Scholar,  
Thesis: The Necessity of Separating Control and Logic When Grounding Language using Neuroevolution  
Advisor: Risto Miikkulainen Second Reader: Raymond Mooney

## Students

- |     |   |     |        |
|-----|---|-----|--------|
| PhD | <b>Hao Zhu</b> , <i>Unsupervised structure induction and Theory of Mind</i> , | LTI | 2019 – |
| PhD | <b>Jared Fernandez</b> , <i>Efficient Multimodal Models</i> ,                 | LTI | 2020 – |
| MLT | <b>Josh Zhanson</b> , <i>Visual Representations for Embodied Agents</i> ,     | LTI | 2020 – |
| MLT | <b>Yingshan Chang</b> , <i>Multimodal Visual Reasoning</i> ,                  | LTI | 2020 – |
| PhD | <b>Vidhi Jain</b> , <i>Language to Control</i> ,                              | RI  | 2021 – |
| PhD | <b>Tiffany Min</b> , <i>Knowledge for Embodied agents</i> ,                   | MLD | 2021 – |

## Conference and Journal Publications

- 2021 **Dependency Induction Through the Lens of Visual Perception**  
Ruisi Su, Shruti Rijhwani, Hao Zhu, Junxian He, Xinyu Wang, Yonatan Bisk, Graham Neubig,  
*The SIGNLL Conference on Computational Natural Language Learning*
- 2021 **Language Grounding with 3D Objects**  
Jesse Thomason, Mohit Shridhar, Yonatan Bisk, Chris Paxton, Luke Zettlemoyer,  
*Conference on Robot Learning (CoRL)*
- 2021 **TACo: Token-aware Cascade Contrastive Learning for Video-Text Alignment**  
Jianwei Yang, Yonatan Bisk, Jianfeng Gao,  
*International Conference on Computer Vision*
- 2021 **Few-shot Language Coordination by Modeling Theory of Mind**  
Hao Zhu, Graham Neubig, Yonatan Bisk,  
*The Thirty-eighth International Conference on Machine Learning*
- 2021 **Grounding ‘Grounding’ in NLP**  
Khyathi Raghavi Chandu, Yonatan Bisk, Alan W Black,  
*Findings of The 2021 Conference of the Association for Computational Linguistics*
- 2021 **An Empirical Study on the Generalization Power of Neural Representations Learned via Visual Guessing Games**  
Alessandro Suglia, Yonatan Bisk, Ioannis Konstas, Antonio Vergari, Emanuele Bastianelli, Andrea Vanzo, Oliver Lemon,  
*The 2021 Conference of the European Chapter of the Association for Computational Linguistics, (Oral)*

- 2021 **Knowledge-driven Data Construction for Zero-shot Evaluation in Commonsense Question Answering**  
Kaixin Ma, Filip Ilievski, Jonathan Francis, [Yonatan Bisk](#), Eric Nyberg, Alessandro Oltramari,  
*35th AAAI Conference on Artificial Intelligence*
- 2021 **ALFWorld: Aligning Text and Embodied Environments for Interactive Learning**  
Mohit Shridhar, Xingdi Yuan, Marc-Alexandre Côté, [Yonatan Bisk](#), Adam Trischler, Matthew Hausknecht,  
*Proceedings of the International Conference on Learning Representations*
- 2020 **Imagining Grounded Conceptual Representations from Perceptual Information in Situated Guessing Games**  
Alessandro Suglia, Antonio Vergari, Ioannis Konstas, [Yonatan Bisk](#), Emanuele Bastianelli, Andrea Vanzo, Oliver Lemon,  
*The 28th International Conference on Computational Linguistics*, (Oral)
- 2020 **The Return of Lexical Dependencies: Neural Lexicalized PCFGs**  
Hao Zhu, [Yonatan Bisk](#), Graham Neubig,  
*Transactions of the Association for Computational Linguistics*
- 2020 **Experience Grounds Language**  
[Yonatan Bisk](#), Ari Holtzman, Jesse Thomason, Jacob Andreas, Yoshua Bengio, Joyce Chai, Mirella Lapata, Angeliki Lazaridou, Jonathan May, Aleksandr Nisnevich, Nicolas Pinto, Joseph Turian,  
*Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (Oral)
- 2020 **RMM: A Recursive Mental Model for Dialog Navigation**  
Homero Roman Roman, [Yonatan Bisk](#), Jesse Thomason, Asli Celikyilmaz, Jianfeng Gao,  
*Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings*
- 2020 **ALFRED: A Benchmark for Interpreting Grounded Instructions for Everyday Tasks**  
Mohit Shridhar, Jesse Thomason, Daniel Gordon, [Yonatan Bisk](#), Winson Han, Roozbeh Mottaghi, Luke Zettlemoyer, Dieter Fox,  
*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
- 2020 **PIQA: Reasoning about Physical Commonsense in Natural Language**  
[Yonatan Bisk](#), Rowan Zellers, Ronan Le Bras, Jianfeng Gao, Yejin Choi,  
*Thirty-Fourth AAAI Conference on Artificial Intelligence*, (Oral)
- 2019 **Robust Navigation with Language Pretraining and Stochastic Sampling**  
Xiujuan Li, Chunyuan Li, Qiaolin Xia, [Yonatan Bisk](#), Asli Celikyilmaz, Jianfeng Gao, Noah A. Smith, Yejin Choi,  
*Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing*
- 2019 **Defending Against Neural Fake News**  
Rowan Zellers, Ari Holtzman, Hannah Rashkin, [Yonatan Bisk](#), Ali Farhadi, Franziska Roesner, Yejin Choi,  
*Thirty-third Conference on Neural Information Processing Systems*
- 2019 **Improving Robot Success Detection using Static Object Data**  
Rosario Scalise, Jesse Thomason, [Yonatan Bisk](#), Siddhartha Srinivasa,  
*Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems*
- 2019 **Early Fusion for Goal Directed Robotic Vision**  
Aaron Walsman, [Yonatan Bisk](#), Saadia Gabriel, Dipendra Misra, Yoav Artzi, Yejin Choi, Dieter Fox,  
*Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems*
- 2019 **HellaSwag: Can a Machine Really Finish Your Sentence?**  
Rowan Zellers, Ari Holtzman, [Yonatan Bisk](#), Ali Farhadi, Yejin Choi,  
*Proceedings of the Association for Computational Linguistics*
- 2019 **From Recognition to Cognition: Visual Commonsense Reasoning**  
Rowan Zellers, [Yonatan Bisk](#), Ali Farhadi, Yejin Choi,  
*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Oral)
- 2019 **Tactical Rewind: Self-Correction via Backtracking in Vision-and-Language Navigation**  
Liyiming Ke, Xiujuan Li, [Yonatan Bisk](#), Ari Holtzman, Zhe Gan, Jingjing Liu, Jianfeng Gao, Yejin Choi, Siddhartha Srinivasa,  
*The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Oral)
- 2019 **Prospection: Interpretable Plans From Language By Predicting the Future**  
Chris Paxton, [Yonatan Bisk](#), Jesse Thomason, Arunkumar Byravan, Dieter Fox,  
*Proceedings of the International Conference on Robotics and Automation (ICRA)*
- 2019 **Shifting the Baseline: Single Modality Performance on Visual Navigation & QA**  
Jesse Thomason, Daniel Gordon, [Yonatan Bisk](#),  
*Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics*

- 2019 **Benchmarking Hierarchical Script Knowledge**  
Yonatan Bisk, Jan Buys, Karl Pichotta, Yejin Choi,  
*Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics*, (Oral)
- 2019 **Character-based Surprisal as a Model of Reading Difficulty in the Presence of Errors**  
Michael Hahn, Frank Keller, Yonatan Bisk, Yonatan Belinkov,  
*The 41st Annual Meeting of the Cognitive Science Society*, (Oral)
- 2018 **SWAG: A Large-Scale Adversarial Dataset for Grounded Commonsense Inference**  
Rowan Zellers, Yonatan Bisk, Roy Schwartz, Yejin Choi,  
*Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (Oral)
- 2018 **Synthetic and Natural Noise Both Break Neural Machine Translation**  
Yonatan Belinkov, Yonatan Bisk,  
*Proceedings of the 6th International Conference on Learning Representations*, (Oral)
- 2018 **Learning Interpretable Spatial Operations in a Rich 3D Blocks World**  
Yonatan Bisk, Kevin Shih, Yejin Choi, Daniel Marcu,  
*Proceedings of the Thirty-Second Conference on Artificial Intelligence (AAAI-18)*
- 2017 **Natural Language Inference from Multiple Premises**  
Alice Lai, Yonatan Bisk, Julia Hockenmaier,  
*Proceedings of the Eighth International Joint Conference on Natural Language Processing*, (Oral)
- 2016 **Natural Language Communication with Robots**  
Yonatan Bisk, Deniz Yuret, Daniel Marcu,  
*Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics*, (Oral)
- 2016 **Supertagging with LSTMs**  
Ashish Vaswani, Yonatan Bisk, Kenji Sagae, Ryan Musa,  
*Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics (Short Papers)*, (Oral)
- 2016 **Evaluating Induced CCG Parsers on Grounded Semantic Parsing**  
Yonatan Bisk, Siva Reddy, John Blitzer, Julia Hockenmaier, Mark Steedman,  
*Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*
- 2015 **Probing the Linguistic Strengths and Limitations of Unsupervised Grammar Induction**  
Yonatan Bisk, Julia Hockenmaier,  
*Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, (Oral)
- 2015 **Labeled Grammar Induction with Minimal Supervision**  
Yonatan Bisk, Christos Christodoulopoulos, Julia Hockenmaier,  
*Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*
- 2013 **An HDP Model for Inducing Combinatory Categorical Grammars**  
Yonatan Bisk, Julia Hockenmaier,  
*Transactions of the Association for Computational Linguistics*
- 2012 **Simple Robust Grammar Induction with Combinatory Categorical Grammar**  
Yonatan Bisk, Julia Hockenmaier,  
*Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI-12)*
- 2012 **Document-Topic Hierarchies from Document Graphs**  
Tim Weneringer, Yonatan Bisk, Jiawei Han,  
*Proceedings of the 21st ACM international conference on Information and knowledge management (CIKM 2012)*
- 2010 **Normal-form parsing for CCGs with generalized composition and type-raising**  
Julia Hockenmaier, Yonatan Bisk,  
*Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*

## Workshop Publications & Tech Reports

- 2021 **WebQA: Multihop and Multimodal QA**  
Yingshan Chang, Mridu Narang, Hisami Suzuki, Guihong Cao, Jianfeng Gao, Yonatan Bisk,  
*ArXiv*
- 2021 **KB-VLP: Knowledge Based Vision and Language Pretraining**  
Kezhen Chen, Qiuyuan Huang, Yonatan Bisk, Daniel McDuff, Jianfeng Gao,  
*ICML21 Workshop on Self-Supervised Learning*

- 2021 **Worst of Both Worlds: Biases Compound in Pre-trained Vision-and-Language Models**  
Tejas Srinivasan, [Yonatan Bisk](#),  
*ArXiv*
- 2020 **A Benchmark for Structured Procedural Knowledge Extraction from Cooking Videos**  
Frank F. Xu, Lei Ji, Botian Shi, Junyi Du, Graham Neubig, [Yonatan Bisk](#), Nan Duan,  
*Workshop on NLP Beyond Text*
- 2020 **Multi-View Learning for Vision-and-Language Navigation**  
Qiaolin Xia, Xiujun Li, Chunyuan Li, [Yonatan Bisk](#), Zhifang Sui, Jianfeng Gao, Yejin Choi, Noah A. Smith,  
*ArXiv*
- 2019 **FIND: Identifying Functionally and Structurally Important Features in Protein Sequences with Deep Neural Networks**  
Ranjani Murali, James Hemp, Victoria Orphan, [Yonatan Bisk](#),  
*bioRxiv*
- 2018 **Inducing Grammars with and for Neural Machine Translation**  
Ke Tran, [Yonatan Bisk](#),  
*Proceedings of the 2nd Workshop on Neural Machine Translation*
- 2018 **Bridging HMMs and RNNs through Architectural Transformations**  
Jan Buys, [Yonatan Bisk](#), Yejin Choi,  
*32nd Conference on Neural Information Processing Systems (NIPS 2018), IRASL workshop*
- 2018 **CHALET: Cornell House Agent Learning Environment**  
Claudia Yan, Dipendra Misra, Andrew Bennett, Aaron Walsman, [Yonatan Bisk](#), Yoav Artzi,  
*ArXiv Preprint 1801.07357*
- 2016 **Towards a Dataset for Human Computer Communication via Grounded Language Acquisition**  
  
[Yonatan Bisk](#), Daniel Marcu, William Wong,  
*Proceedings of the AAAI 2016 Workshop on Symbiotic Cognitive Systems*
- 2016 **Unsupervised Neural Hidden Markov Models**  
Ke Tran, [Yonatan Bisk](#), Ashish Vaswani, Daniel Marcu, Kevin Knight,  
*Proceedings of the Workshop on Structured Prediction for NLP, (Oral)*
- 2012 **Induction of Linguistic Structure with Combinatory Categorical Grammars**  
[Yonatan Bisk](#), Julia Hockenmaier,  
*NAACL HLT Workshop on Induction of Linguistic Structure*

## Thesis Committees

- **PhD**, [Volkan Cirik](#), Louis-Philippe Morency and Taylor Berg-Kirkpatrick, CMU
- **PhD**, [Shruti Palaskar](#), Alan Black and Florian Metze, CMU
- **PhD**, [Pengcheng Yin](#), Graham Neubig, CMU
- **PhD**, [Dheeraj Rajagopal](#), Ed Hovy, CMU
- **PhD**, [Evangelia Spiliopoulou](#), Ed Hovy, CMU
- 2020 **PhD**, [Zihang Dai](#), Yiming Yang, CMU
- 2021 **MLT**, [Jing Wen](#), Katerina Fragkiadaki, CMU
- **Qual**, [Mohit Sharma](#), Oliver Kroemer, CMU

## Teaching

- 2021 **11-767 On-Device Machine Learning**, *Co-Emma Strubell*, # Students: 31, Fall
- 2021 **11-777 MultiModal Machine Learning**, # Students: 42, Spring

## Diversity and Inclusion

- 2021 **NAACL 2022 DEI Co-Chair**, ongoing
- 2020 **LTI DEI Committee**, ongoing
- 2019 **CrashCourse AI**, Educational YouTube series targeted at high-school age viewers. Hosted by Jabril Ashe. Co-Written with Lana Yarosh and Tim Weninger, 1yr
- 2018 **ACM Council on Diversity & Inclusion**, Investigating and initiating institutional change within ACM on behalf of underrepresented groups., 1yr

- **Grace Hopper PC member**, Review submissions to AI Technical Track, 1yr
- 2016 **NCWIT Collegiate Award**, Review applications for the National Center for Women & Information Technology's Collegiate Award honoring college women, ISI 3yrs
- 2009 **NCWIT Aspirations in Computing**, Review applications for the National Center for Women & Information Technology's Award for Aspirations in Computing which supports high-school girls interested in CS, UIUC, ISI 6yrs
- **Tech Junket/Women in CS**, Outreach for local Urbana-Champaign pre-college students, UIUC 2yrs
- 2008 **Freshman Research Mentor**, Mentored students to perform research, UTexas 1yr
- **First Bytes**, Helped organize a summer camp for girls interested in CS, UTexas
- **Co-Founded Breakfast Bytes**, Professor talks for pre-college students, UTexas 1yr
- **CS Roadshow**, CS presentations at Austin middle and high schools, UTexas 1yr

## Funding and Awards

- 2021 **NSF EAGER**,  
Co-PI Graham Neubig, Dan Yurovsky, Aug 2021 – Dec 2022, \$300K
- 2020 **Facebook RFP**, *Learning Dynamic Multimodal Inference for On-Device VQA*,  
Co-PI Emma Strubell, Sept 2020 – May 2021, \$75K
- 2020 **Google exploreCSR**, *Broad & Inclusive Language Technology (BILT) @ CMU LTI*,  
Co-PI Emma Strubell, Graham Neubig, Summer 2021, \$15k

## Organization (Workshops and Conference)

- 2021 **Workshop**, *Embodied AI Workshop – Challenge Organizer*,  
Computer Vision and Pattern Recognition
- 2020 **Workshop**, *Embodied Vision, Actions & Language Workshop*,  
European Conference on Computer Vision
- **Workshop**, *Workshop on Dataset Curation and Security*,  
NeurIPS
- 2019 **Workshop**, *Spatial Language Understanding and Grounded communication for Robotics*,  
North American Chapter of the Association for Computational Linguistics
- 2018 **Conference**, *Social Media Chair*,  
North American Chapter of the Association for Computational Linguistics
- **Workshop**, *Generalization in the Age of Deep Learning*,  
North American Chapter of the Association for Computational Linguistics
- **Workshop**, *Wordplay: Reinforcement and Language Learning in Text-Based Games*, NeurIPS
- 2017 **Conference**, *Student Scholarship Chair and Student Volunteer Coordinator*,  
Conference on Empirical Methods in Natural Language Processing
- **Workshop**, *Language Grounding for Robotics*,  
Conference of the Association for Computational Linguistics

## Area Chair

- 2022 **International Conference on Learning Representations**
- **Association for the Advancement of Artificial Intelligence**, *Senior Meta-Reviewer*
- 2021 **International Conference on Learning Representations**
- **Association for Computational Linguistics**, *Machine Learning*
- **Empirical Methods in Natural Language Processing**, *Speech, Vision, Robotics, Multimodal Grounding*
- **Neural Information Processing Systems (NeurIPS)**
- **ACL Rolling Review**
- 2020 **Association for Computational Linguistics**, *Language Grounding to Vision, Robotics and Beyond*
- **Neural Information Processing Systems (NeurIPS)**
- **Asia-Pacific chapter of the Association for Computational Linguistics**

- 2019 **Association for Computational Linguistics, Vision, Robotics, Multimodal Grounding, and Speech**  
 2018 **International Conference on Computational Linguistics**  
 2016 **Conference on Lexical and Computational Semantics (\*SEM)**, *Sem. Parsing and Role Labeling*

## Reviewer

NLP	<b>ACL</b>	Association for Computational Linguistics Rolling Review,	2021
	<b>Rolling</b>		
	<b>ACL</b>	Association for Computational Linguistics,	2017, 2018
	<b>CL</b>	Computational Linguistics Journal,	2017, <i>standing</i>
	<b>CoLing</b>	International Conference on Computational Linguistics,	2016
	<b>CoNLL</b>	Conference on Computational Natural Language Learning,	2016 – 2018
	<b>EACL</b>	European Chapter of the Association for Computational Linguistics,	2014, 2017, 2020
	<b>EMNLP</b>	Empirical Methods in Natural Language Processing,	2013 – 2020
	<b>IJCNLP</b>	International Joint Conference on Natural Language Processing,	2017
	<b>NAACL</b>	N. American Chapter of the Association for Computational Ling,	2016 – 2019
	<b>NLE</b>	Natural Language Engineering Journal,	2016
	<b>*SEM</b>	Joint Conference on Lexical and Computational Semantics,	2020
Robo	<b>TACL</b>	Transactions of the Assoc. for Computational Linguistics,	2019
	<b>CoRL</b>	Conference on Robot Learning,	2019 – 2021
	<b>HRI</b>	Human Robot Interaction,	2020 – 2021
	<b>ICRA</b>	International Conference on Robotics and Automation,	2019
	<b>IJRR</b>	International Journal of Robotics Research,	2017
	<b>IROS</b>	IEEE/RSJ International Conference on Intelligent Robots and Systems,	2019
AI/ML	<b>RA-L</b>	IEEE Robotics and Automation Letters,	2019
	<b>AAAI</b>	Association for the Advancement of Artificial Intelligence,	2018, 2021
	<b>ICLR</b>	International Conference on Learning Representations,	2019 – 2020
	<b>ICML</b>	International Conference on Machine Learning,	2019
	<b>IJCAI</b>	International Joint Conference on Artificial Intelligence,	2019
CV	<b>NeurIPS</b>	Neural Information Processing Systems,	2018 – 2019
	<b>CVPR</b>	Computer Vision and Pattern Recognition,	2020

## Invitations and Talks

- 2021 **2nd Workshop on Advances in Language and Vision Research (ALVR)**, *NAACL Instructions, Abstraction, and Theory-of-Mind, KoUniversity AI center*  
**Understanding the Potential of Scaling Artificial Intelligence for Science Applications Panel**, *PNNL*  
**Guest Lecture: Embodiment for NLP**, *UPenn*  
**Guest Lecture: Multimodality in NLP**, *UWashington Linguistics*
- 2020 **The 2nd Workshop Beyond Vision and Language: Integrating Real-world Knowledge**, *CoLing*  
**The University of Texas at Austin – Forum for AI**  
**Cornell Learning Machines seminar**  
**Stanford NLP seminar**  
**SRI Center for Vision Technology Lab Panel**, *Embodied AI*
- 2018 **Colloquium Talk**, *Northwest NLP*
- 2017 **Invited Talk**, *University of Notre Dame NLP Seminar*  
**Invited Talk**, *University of California, Irvine AI/ML Seminar*  
**Invited Talk**, *Pacific Northwest National Laboratory*
- 2016 **Invited Poster**, *The Impact of Computing in Our Physical World session of the CRA's Computing Community Consortium (CCC) Symposium on Computing Research*
- 2015 **Colloquium Talk**, *Toyota Tech. Inst. of Chicago's Midwest Speech & Language Days*

2013 **Colloquium Talk**, *Toyota Tech. Inst. of Chicago's Midwest Speech & Language Days*