

# HILA GONEN

(+650)5869403 ✧ hilagnn@gmail.com  
https://gonenhila.github.io ✧ https://github.com/gonenhila

## PROFESSIONAL EXPERIENCE

---

<b>Meta AI and University of Washington</b> – Postdoctoral Researcher	<i>2021 - now</i>
<b>Amazon</b> – Research Scientist, Alexa shopping – Postdoc position	<i>2020 - 2021</i>
<b>Google NY</b> – summer internship	<i>2019</i>
<b>Google Israel</b> – research student Automatically Identifying Gender Issues in Machine Translation using Perturbations.	
<b>“Beyond Minds”</b> – providing expert knowledge to AI companies: Consultant.	<i>2019</i>
<b>Intel</b> – WiFi core and features team: developing WiFi driver in C Language.	<i>2012 - 2013</i>

## EDUCATION

---

<b>Bar-Ilan University</b> PhD student in computer science. Research field: <b>Natural Language Processing and deep learning</b> , under the supervision of Prof. Yoav Goldberg. Thesis: <b>Multilinguality and Bias in Language Modeling</b> . Grades: average – <b>97.3</b>	<i>2015 - 2021</i>
<b>The Hebrew University of Jerusalem</b> M.Sc. in computer science (with honor). Research field: formal verification, under the supervision of Prof. Orna Kupferman. Thesis: <b>Inherent Vacuity in Lattice Automata</b> . Grades: average – <b>95.37</b> , final exam – <b>96</b> , thesis – <b>95</b> .	<i>2012 - 2014</i>
<b>The Hebrew University of Jerusalem</b> B.Sc. in computer science (major) and Physics (minor). Studied during autumn 2011 in the University of Helsinki as an exchange student. Grades: average – <b>91.19</b> , CS average – <b>92.31</b> .	<i>2009 - 2012</i>

## AWARDS

---

Best Doctoral Dissertation Award Runner Up, Israeli Association for Artificial Intelligence (IAAI).	<i>2022</i>
EECS Rising Starts Award.	<i>2022</i>
Best paper award at the RepL4NLP workshop at ACL.	<i>2022</i>
Fulbright Postdoctoral Fellowship – declined.	<i>2021</i>
Rothschild Postdoctoral Fellowship, Yad Hanadiv.	<i>2021</i>
The Eric and Wendy Schmidt Postdoctoral Award for Women in Mathematical and Computing Sciences.	<i>2019</i>
Best paper award at CoNLL (Conference on Computational Natural Language Learning).	<i>2019</i>
Excellence scholarship for PhD students from Bar-Ilan University (the President’s Scholarship).	<i>2016</i>
M.Sc. in computer science with honor from the Hebrew University of Jerusalem.	<i>2014</i>

Excellence scholarship for M.Sc. students from the school of Computer Science and Engineering at the Hebrew University of Jerusalem.

2012

Dean's list, second year of B.Sc.

2011

## PUBLICATIONS

---

J. Lee, A. Liu, O. Ahia, H. Gonen, N. A. Smith. That was the last straw, we need more: Are Translation Systems Sensitive to Disambiguating Context? In *findings* of EMNLP, 2023.

H. Gonen, S. Iyer, T. Blevins, N. A. Smith, L. Zettlemoyer. Demystifying prompts in language models via perplexity estimation. In *findings* of EMNLP, 2023.

W. Shi, X. Han, H. Gonen, A. Holtzman, Y. Tsvetkov, L. Zettlemoyer. Toward Human Readable Prompt Tuning: Kubrick's The Shining is a good movie, and a good prompt too? In *findings* of EMNLP, 2023.

M. Ghazvininejad, H. Gonen, L. Zettlemoyer. Dictionary-based Phrase-level Prompting of Large Language Models for Machine Translation. arXiv preprint, arXiv:2302.07856.

D. Liang, H. Gonen, Y. Mao, R. Hou, N. Goyal, M. Ghazvininejad, L. Zettlemoyer, M. Khabsa. XLM-V: Overcoming the Vocabulary Bottleneck in Multilingual Masked Language Models. In proceedings of EMNLP, 2023.

A. Asai, S. Kudugunta, X. Yu, T. Blevins, H. Gonen, M. Reid, Y. Tsvetkov, S. Ruder, H. Hajishirzi. BUFFET: Benchmarking Large Language Models for Few-shot Cross-lingual Transfer. arXiv preprint, arXiv:2305.14857.

O. Ahia, S. Kumar, H. Gonen, J. Kasai, D. R. Mortensen, N. A. Smith, Y. Tsvetkov. Do All Languages Cost the Same? Tokenization in the Era of Commercial Language Models. In proceedings of EMNLP, 2023.

O. Ahia, H. Gonen, V. Balachandran, N. A. Smith, Y. Tsvetkov. LEXPLAIN: Improving Model Explanations via Lexicon Supervision. \*SEM conference 2023.

T. Blevins, H. Gonen, L. Zettlemoyer. Prompting Language Models for Linguistic Structure. In proceedings of ACL, 2023.

T. Blevins, H. Gonen, L. Zettlemoyer. Analyzing the Mono-and Cross-Lingual Pretraining Dynamics of Multilingual Language Models. In proceedings of EMNLP, 2022.

A. Cohen, H. Gonen, O. Shapira, R. Levy, Y. Goldberg. McPhraSy: Multi context phrase similarity and clustering. In *findings* of EMNLP, 2022.

H. Gonen, S. Ravfogel and Y. Goldberg. Analyzing Gender Representation in Multilingual Models. In Proceedings of the ACL Workshop RepL4NLP: Representation Learning for NLP, 2022, **Best Paper**.

D. Cirillo, H. Gonen, E. Santus, A. Valencia, M. R. Costa-juss, and M. Villegas. "Sex and Gender Bias in Natural Language Processing". Sex and Gender Bias in Technology and Artificial Intelligence, edited by Davide Cirillo, Silvina Catuara Solarz and Emre Guney, Elsevier, May 2022, pages 113–132.

I. Gamzu, H. Gonen, G. Kutiel, R. Levy and E. Agichtein. Identifying Helpful Sentences in Product Reviews. In proceedings of NAACL, 2021.

E. Rabinovich, H. Gonen and S. Stevenson. Pick a Fight or Bite your Tongue: Investigation of Gender Differences in Figurative Language Usage. In proceedings of COLING, 2020.

H. Gonen, S. Ravfogel, Y. Elazar, and Y. Goldberg. It's not Greek to mBERT: Inducing Word-Level Translations from Multilingual BERT. In Proceedings of the EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP, 2020.

- H. Gonen and K. Webster. Automatically Identifying Gender Issues in Machine Translation using Perturbations. In *findings of EMNLP*, 2020.
- S. Ravfogel, Y. Elazar, H. Gonen, M. Twiton and Y. Goldberg. Null it out: Guarding protected attributes by iterative nullspace projection. In proceedings of ACL, 2020.
- H. Gonen, G. Jawahar, Y. Goldberg and D. Seddah. Simple, Interpretable and Stable Method for Detecting Words with Usage Change across Corpora. In proceedings of ACL, 2020.
- H. Gonen, Y. kementchedjhieva and Y. Goldberg. How does Grammatical Gender Affect Noun Representations in Gender-Marking Languages? In proceedings of CoNLL 2019, **Best Paper**.
- R. H. Maudslay, H. Gonen, R. Cotterell and S. Teufel. It's All in the Name: Mitigating Gender Bias with Name-Based Counterfactual Data Substitution. In proceedings of EMNLP, 2019.
- H. Gonen and Y. Goldberg. Lipstick on a Pig: Debiasing Methods Cover up Systematic Gender Biases in Word Embeddings But do not Remove Them. In proceedings of NAACL, 2019.
- H. Gonen and Y. Goldberg. Language Modeling for Code-Switching: Evaluation, Integration of Monolingual Data, and Discriminative Training. In proceedings of EMNLP, 2019.
- H. Gonen and Y. Goldberg. Semi Supervised Preposition-Sense Disambiguation using Multilingual Data. In proceedings of COLING, 2016.
- H. Gonen and O. Kupferman. Inherent vacuity in lattice automata. In Fields of Logic and Computation II, volume 9300 of Lecture Notes in Computer Science, pages 174-192. Springer, 2015.

## ACADEMIC ACTIVITY

---

### REVIEWING

- |   |      |
|---|------|
| Area Chair (AC), track of large language models, ACL 2023.  | 2023 |
| Reviewer for the 2nd Multilingual Representations Learning Workshop (MRL), EMNLP 2022.                          | 2022 |
| Reviewer for Northern European Journal of Language Technology (NEJLT).  | 2022 |
| Reviewer for EACL 2021, NAACL 2021, TrustNLP 2021, Multilingual Representation Learning Workshop 2021 and TACL. | 2021 |
| Reviewer for ACL 2020.  | 2020 |
| Reviewer for ACL 2019, *Sem 2019, repEval 2019, EurNLP 2019 and TACL.   | 2019 |
| Reviewer for NAACL 2018, COLING 2018, SRW 2018.   | 2018 |

### ORGANIZATION

- |  |      |
|--|------|
| Co-organizer of the 3rd Multilingual Representations Learning Workshop (MRL), EMNLP.   | 2023 |
| Co-organizer of the shared task on Clause-level Morphology, the 2nd Multilingual Representations Learning Workshop (MRL), EMNLP. | 2022 |
| Co-organizer of the 2nd Multilingual Representations Learning Workshop (MRL), EMNLP.   | 2022 |
| Co-organizer of the 4th Workshop on Research in Computational Typology and Multilingual NLP (SIGTYP), NAACL.                     | 2022 |
| Co-organizer of the 4th Gender Bias in NLP Workshop (GeBNLP), NAACL.   | 2022 |
| Co-organizer of the 3rd Gender Bias in NLP Workshop (GeBNLP), ACL.   | 2021 |

## INVITED TALKS

### **How negative results fuel our research:**

#### **Insights from multilinguality and gender bias**

Invited Keynote, Workshop on Insights from Negative Results in NLP

2023

### **Demystifying Prompts in Language Models via Perplexity Estimation:**

The University of British Columbia, NLP group

2023

Stanford University, Prof. Dan Jurafsky's group

2023

Hebrew University of Jerusalem, NLP group

2022

New-York University, NLP group

2022

UT Austin, NLP group

2022

### **Gender Bias in Word Embeddings:**

Guest lecture at the NLP course at UW (Prof. Noah A. Smith)

2022

### **Subspaces in multilingual models:**

University of Washington, Prof. Emily Bender's group

2021

### **What's in a Representation:**

Weizmann Institute

2021

Technion

2021

Hebrew University of Jerusalem

2021

University of Washington, Prof. Noah A. Smith's group

2021

### **Panelist at a Gender Equality workshop at the WSIS forum (UN/UNESCO)**

2021

### **Panelist at Widening NLP Workshop at AACL-IJCNLP**

2020

### **Representation and Bias in NLP:**

Seminar, Ben Gurion University

2019

### **Gender Bias in Word Embeddings:**

FastAI Women course, Tel-Aviv

2020

AI Week, Tel-Aviv

2019

WiDS, Tel-Aviv – selected to appear at **Best of WiDS Global 2021**

2019

Inria lab, Paris

2019

Basis, Israel

2019

### **Language Modeling for Code-switching:**

AI Data Science Summit, Israel

2019

Microsoft Research, Israel

2018

ONLP lab, Israel

2018

### **Tutorial on word embeddings:**

WiDS, Tel-Aviv

2018

### **Semi Supervised Preposition-Sense Disambiguation using Multilingual Data:**

The Data Science Summit Europe, Jerusalem, Israel

2017

## EVENTS

Participated in the EECS Rising Stars Workshop at UT Austin

2022

Participated in the Logic, Language and Information summer school (NASSLLI) at CMU

2018

Participated in Lisbon Machine Learning School (LXMLS)

2016

Participated in Marktoberdorf Summer School (Dependable Software Systems Engineering)

2014

Participated in Women in Theory workshop (New York)

*2014*

## **COURSES AND SERVICE TO HOME INSTITUTION**

---

University of Washington: organizer and leader of a reading group on Multilinguality in NLP for graduate students.

*2022*

The Hebrew University: TA in the course “Data Structures” for first-year students (frontal teaching).

*2014*

The Hebrew University: TA in the course “Computational Models, Computability and Complexity” for second-year students.

*2013*

The Hebrew University: TA in the course “Data Bases” for second-year students.

*2012*

## **SKILLS**

---

Programming languages: Python, Matlab, C, C++, Java

Tools and Technologies: Deep Learning (DyNet, PyTorch), NLP (spaCy, NLTK), git

## **LANGUAGES**

---

Hebrew: native tongue.

English: high level.

Arabic: high level. Rich experience in translating Arabic texts.

French: basic level.