

# Leshem Choshen - Curriculum Vitae

MIT-IBM



## Honors and Awards

- 2024 - Culture Catalyst Award
- 2024 - IBM Group Research Accomplishment for Scientific and Business **Impact**
- 2024 - ISW Best Poster, ZipNN – A Lossless Compression library tailored for AI Models
- 2023 - ACL Area Chair **Best Paper** Award, Disentqa: Disentangling parametric and contextual knowledge with counterfactual question answering
- 2023 - IBM Outstanding Technical Achievement Award, for [label-sleuth](#) contributions
- 2023 - **Rothschild** postdoctoral Fellowship (50K\$)
- 2023 - **Fulbright** postdoctoral Fellowship nominee (50K\$)
- 2023 - **Best Dissertation** award IAAI -The Israeli Association for Artificial Intelligence
- 2022 - Blavatnik Award for the most promising PhD
- 2022 - KLA scholarship for academic **achievement** and **excellence**
- 2021 - ColD-Fusion chosen as IBM's Next-Up for projects with great potential
- 2021 - IBM Group Research Accomplishment for Scientific and Business **Impact**
- 2021 - KLA scholarship for academic **achievement** and **excellence**
- 2020 - IBM Group Research Accomplishment for Scientific and Business **Impact**
- 2020 - Kamin lab grant promoting applied PhD research (~100K\$)
- 2020 - KLA scholarship for academic **achievement** and **excellence**
- 2019 - **Clore** Scholars Programme scholarship for **outstanding** researchers
- 2019 - Selected to National scholarship for **outstanding** Data Science PhD students
- 2018 - Held the first Human-Machine debate, acclaimed by public and social media
- 2018 - IBM's **top** research papers of the year: Will it Blend? Blending Weak and Strong Labeled Data in a Neural Network for Argumentation Mining
- 2018 - **Best** thesis **award**, Department of cognitive sciences
- 2018 - **Valedictorian** (M.A.)
- 2018 - **Dean's List** of Excellence, Faculty of Natural Sciences
- 2018 - **Best paper** award, Faculty of Natural Sciences Huji. Propagating Directed Exploration in Model-Free Reinforcement Learning
- 2017 - A certificate of **excellence** and a scholarship for **outstanding achievement** in M.A. Studies, Department of cognitive sciences
- 2017 - **Outstanding** Poster, Computer science department prize Improved Methodology for reference-based Evaluation Grammatical Error Correction
- 2017 - The **Rector's scholarship** for **uniquely outstanding** students in interdisciplinary fields
- 2017 - **Dean's list** of **excellence**, Faculty of Natural Sciences
- 2016 - A certificate of **excellence** and a scholarship for **outstanding achievement**, Department of cognitive sciences
- 2016 - **Dean's list** of **excellence** and **Dean's Scholarship**, Faculty of Natural Sciences
- 2015 - Selected to Microsoft's national **excellence program** and **scholarship**
- 2015 - A certificate of **excellence** and a scholarship for **outstanding achievement**, Department of cognitive sciences
- 2015 - **Dean's list** of **excellence** and **Dean's Scholarship**, Faculty of Natural Sciences
- 2013 - The Hebrew University of Jerusalem's **excellence scholarship**
- 2013 - A project I initiated at the army and managed won **best project award**
- 2012 - Certificate of **excellence** from a secret army unit
- 2009 - Graduated with **honours** from "Hebrew University Secondary School"

## **Education and Research Lead**

2023-(ongoing) - **MIT-IBM** Postdoc

2021-2023 - Project **lead**, IBM Research (~5 researchers + supporters)

2017-2023 - Ph.D., the Hebrew University of Jerusalem (HUJI), Computer Sciences

2015-2017 - **Summa cum laude (valedictorian)**, M.A. at the Hebrew University of Jerusalem: Cognition Sciences. Final grade: **99.12**, Thesis: **100**, Master's test: **100**

2013-2016 - **Summa cum laude**, B.Sc. at the Hebrew University of Jerusalem:

Cognition and Computer Sciences. Final grade: **97.54**

Israeli SAT exam score: 736 (amongst the **top 1%** of the examinee)

## **Related Professional experience**

2023-(ongoing) - MIT-IBM Postdoc

2021-2023 - Project **lead**, IBM Research (~5 researchers + supporters)

2017-2021 - Ph.D. - Managing up to **17** parallel **projects** (guiding ~**12 students**).

2019-(ongoing) - **Collaboration** with Google Research

2016-2021 - IBM research, A.I., NLP researcher in [Project Debater](#)

Among the many **honors** and **awards**:

- Gold innovation medal by 2019 Edison Awards

- A reason for IBM's recognition as "**world-changing** company of the year" by FastCompany

2014-2016 - **Translator** from English and Arabic to Hebrew, [onehourtranslation.com](#)

2013-2016 - **Translator** from English and Arabic to Hebrew, Israeli TV Channel 1

## **Internal voluntary services**

2021-(ongoing) - Founder and Coordinator of IBM Media Group

2020-2022 - Founder and Coordinator, HUJI Computer Science Media Group

2021-2022 - Manager HUJI NLP lab Twitter account

## **External voluntary services**

2023-(ongoing) - **TACL Journal Reviewer**

2023-(ongoing) **ARR Area Chair**

2021-2024 ARR reviewer

2025 - Organizing Theory of mind competition Neurips

2025 - Organizing BabyLM Challenge and Workshop organizer EMNLP

2025 - Organizing Interplay Workshop Neurips

2025 - Organizing Actionable Interpretability Workshop CoLM

2025 - Area Chair for ICLR, CoLM

2025 - Reviewer Neurips, ICML, CVPR

2024 - CoNLL Emergency **Area Chair**

2024 - [BTTAI](#) mentor, advancing **diversity** by supporting women's integration into tech

2024 - TMLR evaluation **expert opinion** for awards

2024 - BabyLM challenge **Organizer**, CoNLL2024 shared task

2024 - Model merging Competition Neurips **Organizer**

2024 - HILD workshop ICML **Organizer**

2024 - NAACL, ACL, ICLR **Area chair**

2024 - CoNLL Reviewer

2024 - Neurips LLM Efficiency Challenge - committee advisor

2023 - EACL Area Chair

2023 - **Ph.D. committee**, National Research University Higher School of Economics

2023 - Reviewer TMLR

2022-2023 **Initiated** and **organized**: BabyLM challenge, CoNLL2023 shared task

2022 - EACL Publicity **Chair** - organizing committee

2022 - Reviewer CogSci, BEA, COLING, EMNLP, NeurIPS, TMLR

2021 - CoNLL Social Media **Chair** - organizing committee

2021 - Emergency reviewer: NAACL EACL. Reviewer: BEA, ICML, ACL, ACL rolling, EMNLP, NeurIPS, NeurIPS datasets

I appreciate the time you took to read my CV

2020 - Reviewer: CoNLL, BEA-workshop, EMNLP, COLING, AACL, NeurIPS, SIGIR, ACL Student Workshop, ACL  
 2019 - Reviewer: CoNLL secondary, EMNLP, ACL, NAACL, SEM\*  
 2019 - **Organizing** committee: SemEval shared task - cross-lingual semantic parsing  
 2018 - Secondary reviewer: COLING, CoNLL. Emergency reviewer: NAACL Reviewer: EMNLP.  
 2018 - **Organizing** committee: 5<sup>th</sup> workshop on argumentation mining  
 2017 - **Organizing** committee: ISCOL, The Annual Meeting of the Israeli Seminar on Computational Linguistics

### **Invited talks**

2025 - MIT-IA Summer Broad Research Seminar “Shared AI”  
 2025 - @ACL, Generalization Evaluation and Metrics Workshop  
 2025 - IBM Research Evaluation group “Please Sir, I want some evaluations”  
 2024 - JHU, Computer Science Seminar “on loss spaces, fusing and evolving pretraining”  
 2024 - UMD-Adobe Research joint seminar “Recycling finetuned models to pretrain”  
 2024 - Paper Club, Human Feedback Foundation & AI Tinkerers “open human feedback”  
 2024 - Modern evaluation tutorial LREC-COLING  
 2024 - KULeuven PSI seminar, “Collaborative Pretraining: on merging models”  
 2024 - Microsoft Research, “MicroBabyLM”  
 2024 - IBM Research, Large Ideas talk - Inaugural  
 2023 - ICML, Collaborative, Communal, Continual ML workshop (not accepted)  
 2023 - Cambridge University, NLIP Seminar, “Collaborative Pretraining”  
 2023 - IBM Research MNLP seminar, “Efficient Benchmarking”  
 2023 - IBM Research fireside, “On Efficient Benchmarking”  
 2023 - Reichman University NLP course  
 2023 - Climate-friendly AI: Model recycling and efficient training (conference canceled)  
 2023 - Neurips workshop on Collaborative, Communal, and Continual Machine Learning  
 2023 - Lectures at Technion, BGU, TAU, BIU  
 2023 - Munich@NLP the joint Munich seminar “Model recycling”  
 2023 - Lecture Tour in the US (MIT, UW, UNC, Google, IBM, AI2, NYU, Cornell)  
 2022 - University of North Carolina, “Recycling finetuned models to pretrain”  
 2022 - Research **talents** round table with the IBM senior VP, Director of Research  
 2022 - NLP seminar of Yandex and HSE University in St. Petersburg, Russia, “Do language models learn the same things in the same order?”  
 2022 - The Faculty of Industrial Engineering & Management Technion, “Utilizing linguistic knowledge for text generation”  
 2022 - Guest Lecture Hebrew University of Jerusalem “the science behind the buzz”  
 2021 - University of Edinburgh NLP Group, “Enhancing Machine Translation with syntax” cancelled due to COVID-19  
 2021 - Technion, Roi Reichart’s lab, a meeting with a researcher  
 2021 - Guest Lecture Hebrew University of Jerusalem “the science behind the buzz”  
 2020 - Technion Data Seminar, "On the evaluation of Grammatical Error Correction"  
 2020 - Guest Lecture Hebrew University of Jerusalem “the science behind the buzz”

### **Teaching**

2023-2024 - **Lecturer** in Reichman University, Reinforcement Learning, undergraduate  
 2022-2024 - Introduction to LLMs, a full-day course given to multiple audiences  
 2018-2021 - **Supervisor** final projects for engineers (equivalent to Meng)  
 2015-2019 - T.A. MLA College

### **Professional Social Visibility**

(ongoing) - **4K-20K** engaging actively monthly through Twitter (x**10M** reached)  
 (ongoing) - +**2M** people benefited from contributions to **Stack overflow**  
 (ongoing) - top 2% among researchers with first publication in 2018 (ResearchGate)

I appreciate the time you took to read my CV

## **Patents**

- 2024 (under submission) - NumeroLogic: Number Encoding for Enhanced LLMs' Numerical Reasoning. Eli Schwartz, Leshem Choshen, Joseph Shtok, Sivan Doveh, Leonid Karlinsky, Assaf Arbelle - P202401777US01
- 2024 (under submission) - Label-Efficient Model Selection for Text Generation. Shir Ashury-Tahan, Benjamin Sznajder, Leshem Choshen, Liat Ein-Dor, Eyal Shnarch, Ariel Gera
- 2024 - Efficient exponent compression for large AI models - P202403184US01
- 2023 - Automatic Training on massive multitask. Leshem Choshen, Elad Venezian, Shacar Don-Yechia - P202300586US01
- 2022 - System for choosing a base model for fine-tuning. Leshem Choshen, Elad Venezian, Shacar Don-Yechia, Noam Slonim, Yoav Katz P202300586US01
- 2021 - (Provisional) - US Provisional Application No.: 63/141,360; Transition-based Graph Decoder for Neural Machine Translation. Leshem Choshen, Omri Abend
- 2021 - P202103159US01 INTER-TRAINING OF PRE-TRAINED TRANSFORMER-BASED LANGUAGE MODELS USING PARTITIONING AND CLASSIFICATION. Eyal Shnarch, Ariel Gera, Alon Halfon, Lena Dankin, Leshem Choshen, Ranit Aharonov, Noam Slonim

## **Publications**

\* represents an equal contributor as the main author, or co-main author

^ represents the supervising author

- ICLR (under review)** - "Sloth: scaling laws for LLM skills to predict multi-benchmark performance across families" Felipe Maia Polo, Seamus Somerstep, Leshem Choshen, Yuekai Sun, Mikhail Yurochkin
- ICML\*** - "A Hitchhiker's Guide to Scaling Law Estimation" Leshem Choshen, Yang Zhang, Jacob Andreas
- ICML (under review)** - "Compress then Serve: Serving Thousands of LoRA Adapters with Little Overhead" Rickard Br  el-Gabrielsson, Jiacheng Zhu, Onkar Bhardwaj, Leshem Choshen, Kristjan Greenewald, Mikhail Yurochkin, Justin Solomon
- ARR^ (under review)** - "Benchmark Agreement Testing Done Right: A Guide for LLM Benchmark Evaluation" Yotam Perlitz, Ariel Gera, Ofir Arviv, Asaf Yehudai, Elron Bandel, Eyal Shnarch, Michal Shmueli-Scheuer, Leshem Choshen
- ARR (under review)** - "Learning from Naturally Occurring Feedback" Shachar Don-Yehiya, Leshem Choshen, Omri Abend
- ACL** - "The ShareLM Collection and Plugin: Contributing Human-Model Chats for the Benefit of the Community" Shachar Don-Yehiya, Leshem Choshen, Omri Abend
- Arxiv** - "Do these LLM benchmarks agree? Fixing benchmark evaluation with BenchBench" Yotam Perlitz, Ariel Gera, Ofir Arviv, Asaf Yehudai, Elron Bandel, Eyal Shnarch, Michal Shmueli-Scheuer, Leshem Choshen
- Arxiv** - "Can You Trust Your Metric? Automatic Concatenation-Based Tests for Metric Validity" Ora Nova Fandina, Leshem Choshen, Eitan Farchi, George Kour, Yotam Perlitz, Orna Raz
- Nature Machine Intelligence^** - "The Future of Open Human Feedback" - Led by me (20 authors, 15 affiliations)
- TMLR** - "ComPEFT: Compression for Communicating Parameter Efficient Updates via Sparsification and Quantization" Prateek Yadav, Leshem Choshen, Colin Raffel, Mohit Bansal
- CoNLL2023** - "Findings of the BabyLM Challenge: Sample-Efficient Pretraining on Developmentally Plausible Corpora" Michael Y Hu, Aaron Mueller, Candace Ross, Adina Williams, Tal Linzen, Chengxu Zhuang, Ryan Cotterell, Leshem Choshen, Alex Warstadt, Ethan Gotlieb Wilcox

**NAACL** - “Beneath the Surface of Consistency: Exploring Cross-lingual Knowledge Representation Sharing in LLMs” M Ifergan, L Choshen, R Aharoni, I Szpektor, O Abend

**MLSYS (under review)** - “Lossless Compression for AI Models” Moshik Hershcovitch, Leshem Choshen, Andrew Wood, Ilias Enmouri, Peter Chin, Swaminathan Sundararaman, Danny Harnik

**TACL (under review)** - “Elements of World Knowledge (EWOK): A cognition-inspired framework for evaluating basic world knowledge in language models” 5th from end, numerous authors.

**TMLR** - “A Survey on Model MoErging: Recycling and Routing Among Specialized Experts for Collaborative Learning” Prateek Yadav, Colin Raffel, Mohammed Muqeeth, Lucas Caccia, Haokun Liu, Tianlong Chen, Mohit Bansal, Leshem Choshen, Alessandro Sordani

**ICLR** - “Model merging with SVD to tie the Knots” George Stoica, Pratik Ramesh, Boglarka Ecsedi, Leshem Choshen, Judy Hoffman

**ICLR** - “LiveXiv - A Multi-Modal live benchmark based on Arxiv papers content” Nimrod Shabtay, Felipe Maia Polo, Sivan Doveh, Wei Lin, Muhammad Jehanzeb Mirza, Leshem Choshen, Mikhail Yurochkin, Yuekai Sun, Assaf Arbelle, Leonid Karlinsky, Raja Giryes

**Arxiv** - “Global MMLU: Understanding and Addressing Cultural and Linguistic Biases in Multilingual Evaluation” Many authors

**Neurips2025** - “Efficient multi-prompt evaluation of LLMs” Felipe Maia Polo, Ronald Xu, Lucas Weber, Mirian Silva, Onkar Bhardwaj, Leshem Choshen, Allysson Flavio Melo de Oliveira, Yuekai Sun, Mikhail Yurochkin

**EMNLP2024** - “Fuse to Forget: Bias Reduction and Selective Memorization through Model Fusion” Kerem Zaman, Leshem Choshen, Shashank Srivastava

**EMNLP2024** - “NumeroLogic: Number Encoding for Enhanced LLMs' Numerical Reasoning” Eli Schwartz, Leshem Choshen, Joseph Shtok, Sivan Doveh, Leonid Karlinsky, Assaf Arbelle

**JML2024** (conditionally accepted) - “Bigger is not always better: The importance of human-scale language modeling for psycholinguistics” Ethan Gotlieb Wilcox, Michael Hu, Aaron Mueller, Tal Linzen, Alex Warstadt, Leshem Choshen, Chengxu Zhuang, Ryan Cotterell, Adina Williams

**ACL 2024** - “Label-Efficient Model Selection for Text Generation” Shir Ashury-Tahan, Benjamin Sznajder, Leshem Choshen, Liat Ein-Dor, Eyal Shnarch, Ariel Gera

**TACL2024** - “Holmes: Benchmark the Linguistic Competence of Language Models” Andreas Waldis, Yotam Perlitz, Leshem Choshen, Yufang Hou, Iryna Gurevych

**ICML2024** - “TinyBenchmarks: evaluating LLMs with fewer examples” Felipe Maia Polo, Lucas Weber, Leshem Choshen, Yuekai Sun, Gongjun Xu, Mikhail Yurochkin

**ICML2024** - “Asymmetry in low-rank adapters of foundation models” Jiacheng Zhu, Kristjan Greenewald, Kimia Nadjahi, Haitz Sáez de Ocariz Borde, Rickard Br  el Gabrielsson, Leshem Choshen, Marzyeh Ghassemi, Mikhail Yurochkin, Justin Solomon

**ICLR2024^** - “Genie: Achieving human parity in content-grounded datasets generation” Asaf Yehudai, Boaz Carmeli, Yosi Mass, Ofir Arviv, Nathaniel Mills, Assaf Toledo, Eyal Shnarch, Leshem Choshen

**CogSci2024** - “Insights from the first BabyLM Challenge: Training sample-efficient language models on a developmentally plausible corpus” Alex Warstadt, Aaron Mueller, Leshem Choshen, Ethan Gotlieb Wilcox, Chengxu Zhuang, Adina Williams, Ryan Cotterell, Tal Linzen

**ACL2024** - “Deductive Closure Training of Language Models for Coherence, Accuracy, and Updatability” Afra Feyza Aky  rek, Ekin Aky  rek, Leshem Choshen, Derry Wijaya, Jacob Andreas

I appreciate the time you took to read my CV



**NAACL2024<sup>^</sup>** - “Efficient Benchmarking (of Language Models)” Yotam Perlitz, Elron Bandel, Ariel Gera, Ofir Arviv, Liat Ein-Dor, Eyal Shnarch, Noam Slonim, Michal Shmueli-Scheuer, Leshem Choshen

**LREC-Coling 2024** - “Jump to Conclusions: Short-Cutting Transformers With Linear Transformations” Alexander Yom Din, Taelin Karidi, Leshem Choshen, Mor Geva

**EMNLP 2023 Demo** - “Unitxt: Flexible, shareable and reusable data preparation and evaluation for generative ai” Elron Bandel, Yotam Perlitz, Elad Venezian, Roni Friedman-Melamed, Ofir Arviv, Matan Orbach, Shachar Don-Yehia, Dafna Sheinwald, Ariel Gera, Leshem Choshen, Michal Shmueli-Scheuer, Yoav Katz

**CONDA@ACL** - “Data Contamination Report from the 2024 CONDA Shared Task” Numerous.

**EMNLP2023** - “Human Learning by Model Feedback: The Dynamics of Iterative Prompting with Midjourney” Shachar Don-Yechia, Leshem Choshen, Omri Abend

**Neurips2023** - “TIES-Merging: Resolving Interference When Merging Models” Prateek Yadav, Derek Tam, Leshem Choshen, Colin Raffel, Mohit Bansal

**CoNLL2023** - “Findings of the BabyLM Challenge: Sample-Efficient Pretraining on Developmentally Plausible Corpora” Alex Warstadt, Aaron Mueller, Leshem Choshen, Ethan Gotlieb Wilcox, Chengxu Zhuang, Adina Williams, Ryan Cotterell, Tal Linzen

**CoNLL2023** - “MuLER: Multi-Level Evaluation with Reference” Tailin Karidi, Leshem Choshen, Gal Patel, Omri Abend

**EMNLP2023<sup>^</sup>** - “Knowledge is a Region in Weight Space for Fine-tuned Language Models” Almog Gueta, Elad Venezian, Colin Raffel, Noam Slonim, Yoav Katz, Leshem Choshen

**EMNLP2023\*** - “Where to start? Analyzing the potential value of intermediate models” Leshem Choshen, Elad Venezian, Shachar Don-Yechia, Yoav Katz, Noam Slonim

**ACL2023<sup>^</sup>** - “CoLD Fusion: Collaborative Descent for Distributed Multitask Finetuning” Shachar Don-Yehia, Elad Venezian, Colin Raffel, Noam Slonim, Yoav Katz, Leshem Choshen

**ACL2023** - “DisentQA: Disentangling Parametric and Contextual Knowledge with Counterfactual Question Answering” Ella Neeman, Roei Aharoni, Or Honovich, Leshem Choshen, Idan Szpektor, Omri Abend

**EMNLP2022** - “PreQuEL: Quality Estimation of Machine Translation Outputs in Advance” Shachar Don-Yechia, Leshem Choshen, Omri Abend

**CoNLL2022 (also Blackbox NLP)** - “On the Representation of Sentence Structural Invariances in Neural Machine Translation” Gal Patel, Leshem Choshen, Omri Abend

**CoNLL2022\*** - “Transition based Graph Decoder for Neural Machine Translation” Leshem Choshen, Omri Abend

**COLING2022**- “Tackling large action spaces in machine translation” Assaf Yehudai, Leshem Choshen, Omri Abend

**EMNLP2022** - “Label Sleuth: From Unlabeled Text to a Classifier in a Few Hours” Eyal Shnarch, Alon Halfon, Ariel Gera, Marina Danilevsky, Yannis Katsis, Leshem Choshen, Martin Santillan Cooper, Dina Epelboim, Zheng Zhang, Dakuo Wang, Lucy Yip, Liat Ein-Dor, Lena Dankin, Ilya Shnayderman, Ranit Aharonov, Yunyao Li, Naftali Liberman, Philip Levin Slesarev, Gwilym Newton, Shila Ofek-Koifman, Noam Slonim, Yoav Katz

**ACL2022\*** - “The Grammar-Learning Trajectories of Neural Language Models” Leshem Choshen, Guy Hacohen, Daphna Weinshall, Omri Abend

**ACL2022** - “Cluster & Tune: Boost BERT Cold Start Performance in Topical Text Classification” Eyal Shnarch, Ariel Gera, Alon Halfon, Lena Dankin, Leshem Choshen, Ranit Aharonov, Noam Slonim

**Arxiv\*** - "Some Grammatical Errors are Frequent, Others are Important" Leshem Choshen, Ofir Shifman and Omri Abend

**SEM** - "Semantics-aware Attention Improves Neural Machine Translation" Aviv Slobodkin, Leshem Choshen, Omri Abend

**LREC2022** - "GrASP: A Library for Extracting and Exploring Human-Interpretable Textual Patterns" Piyawat Lertvittayakumjorn, Leshem Choshen, Eyal Shnarch, Francesca Toni

**ArXiv\*** - "Fusing finetuned models for better pretraining" Leshem Choshen, Elad Venezian, Noam Slonim, Yoav Katz

**ArXiv** - "ComSum: Commit Messages Summarization and Meaning Preservation" Leshem Choshen & Idan Amit

**ArXiv** - "Part of speech and universal dependency effects on english arabic machine translation" Ofek Rafaeli, Omri Abend, Leshem Choshen, Dmitry Nikolaev

**EMNLP2021** - "Q<sup>2</sup>: Evaluating Factual Consistency in Knowledge-Grounded Dialogues via Question Generation and Question Answering" Or Honnovich, Leshem Choshen, Roei Aharoni, Idan Szpektor, Omri Abend †

**ArXiv\* (Technical report)** - "SERRANT: a syntactic classifier for English Grammatical Error Types" Leshem Choshen, Matanel Oren, Dmitry Nikolaev, Omri Abend

**NAACL2021** - "Mediators in Determining what Processing BERT Performs First" Aviv Solobodkin, Leshem Choshen, Omri Abend

**Nature 2021 (cover)** - "Project Debater - an Autonomous Debating System" 5th author, full list of authors omitted due to length restrictions (55 authors)

**CONLL 2020\*** - "Classifying Syntactic Errors in Learner Language" Leshem Choshen, Dmitry Nikolaev, Yevgeni Berzak, Omri Abend

**EMNLP 2020\*** findings- "Unsupervised Expressive Rules Provide Explainability and Assist Human Experts Grasping New Domains" Eyal Shnarch, Leshem Choshen, Guy Moshkovich, Ranit Aharonov, Noam Slonim

**EMNLP 2020** - "Active Learning for BERT: An Empirical Study" Liat Ein-Dor, Alon Halfon, Ariel Gera, Eyal Shnarch, Lena Dankin, Leshem Choshen, Marina Danilevsky, Ranit Aharonov, Yoav Katz and Noam Slonim

**ICML 2020** - "Let's Agree to Agree: Neural Networks Share Classification Order on Real Datasets" Guy Hacohen, Leshem Choshen, Dafna Weinshall

**ICLR 2020\*** - "On the Weaknesses of Reinforcement Learning for Neural Machine Translation" Leshem Choshen, Lior Fox, Zohar Aizenbud, Omri Abend

**AAAI 2020** - "Corpus wide argument mining - a working solution" Liat Ein-Dor, Eyal Shnarch, Lena Dankin, Alon Halfon, Benjamin Sznajder, Ariel Gera, Carlos Alzate, Martin Gleize, Leshem Choshen, Yufang Hou, Yonatan Bilu, Ranit Aharonov, Noam Slonim

**CoNLL 2019\*** - "Automatically Extracting Challenge Sets for Non-local Phenomena in Neural Machine Translation" Leshem Choshen & Omri Abend

**BEA@ACL 2019\*** - "Learning to combine Grammatical Error Corrections" Yoav Kantor, Yoav Katz, Leshem Choshen, Edo Cohen-Karlik, Naftali Liberman, Assaf Toledo, Amir Menczel, Noam Slonim

**ACL 2019\*** - "The Language of Legal and Illegal Activity on the Darknet" Leshem Choshen, Dan Eldad, Daniel Hershcovich, Elior Sulem, Omri Abend

**ACL 2019** - "Are You Convinced? Choosing the More Convincing Evidence with a Siamese Network" Martin Gleize, Eyal Shnarch, Leshem Choshen, Lena Dankin, Guy Moshkovich, Ranit Aharonov, Noam Slonim

**SemEval@EMNLP 2019** - "SemEval 2019 Task 1: Cross-lingual Semantic Parsing with UCCA" Daniel Hershcovich, Zohar Aizenbud, Leshem Choshen, Elior Sulem, Ari Rappoport, Omri Abend

**ACL 2018\*** - "Inherent Biases in Reference-based Evaluation for Grammatical Error Correction and Text Simplification" Leshem Choshen & Omri Abend

**ACL 2018\*** - "Automatic Metric Validation for Grammatical Error Correction" Leshem Choshen & Omri Abend



**ACL 2018** - "Will it Blend? Weak and Manual Labeled Data in a Neural Network for Argumentation Mining" Eyal Shnarch, Carlos Alzate, Lena Dankin, Martin Gleize, Yufang Hou, Leshem Choshen, Ranit Aharonov and Noam Slonim (short)  
**NAACL-HLT 2018\*** - "Reference-less Measure of Faithfulness for Grammatical Error Correction" Leshem Choshen & Omri Abend  
**ICLR 2018\*** - "DORA The Explorer: Directed Outreaching Reinforcement Action-Selection" Leshem Choshen, Lior Fox, Yonatan Loewenstein

**Popular Media** (That came to my attention - Partial list)

- 2025 MIT news - [How to build AI scaling laws for efficient LLM training and budget maximization](#)
- 2025 lavez - [Universal Guide estimates how LLMs will perform based on smaller models in the same family](#)
- 2025 Tipranks - [IBM's ZipNN Can Shrink AI Models by 33% with Zero Performance Loss](#)
- 2025 Transmitter - [The BabyLM Challenge: In search of more efficient learning algorithms, researchers look to infants](#)
- 2025 WIRED - [Why are researchers turning to small LLMs](#) (Translated to various languages e.g. [1](#), [2](#))
- 2025 [Quant Magazine](#) - [Small LLMs](#)
- 2025 - TextArena IBM Blog [Can LLMs learn social skills by playing games?](#)
- 2025 - ZipNN IBM Blog [Lossless compression tailored for AI](#)
- 2025 - Text Arena reaches millions of people in social media including support from Elon Musk and Ai celebrities
- 2024 - New York Times - [The Race to Make A.I. Smaller \(and Smarter\)](#)
- 2024 - Stanford [Blog](#) on Unitxt
- 2024 - TinyBenchmark IBM Blog [Tiny benchmarks for large language models](#)