

# SUCHIR SALHAN

Suchir Salhan is a PhD Candidate in Computer Science at Gonville and Caius College, the University of Cambridge, working on Cognitively-Inspired Multilingual AI and Small Language Models.

[suchirsalhan.com](http://suchirsalhan.com) | [sas245@cam.ac.uk](mailto:sas245@cam.ac.uk) | [linkedin.com/in/ssalhan](https://www.linkedin.com/in/ssalhan) | [ResearchGate](https://www.researchgate.net/profile/SSalhan)

## EDUCATION

---

Oct 2024 - Jul. 2028

### PhD in Computer Science

UNIVERSITY OF CAMBRIDGE

- Fully-funded PhD Studentship from the Department of Computer Science & Technology to research Small and Interpretable Language Models for multi-agent alignment. Awarded **two outstanding paper awards** (Oct 2025) for research on Small Language Models.
- Academic Collaborators from ETH, EPFL, Bocconi University, KAIST AI, Helmholtz Munich and BabyLM Community. Industry Collaborators from Naver Cloud and Amazon Rufus and UK AI Security Institute (AISI).

Oct 2020 - Jul. 2024

### Computer Science & Linguistics Triposes

UNIVERSITY OF CAMBRIDGE

Bachelor of Arts, Class I with Distinction ( Class 1\*/"Starred First") | Part III (MEng), Distinction

- **Part III Computer Science Tripos (Distinction):** Admitted to the integrated MEng/Part III of the Computer Science Tripos: High Distinctions in Machine Learning Modules include **Advanced Topics in Machine Learning (Explainable AI, Imitation Learning)** and **Machine Learning for Language Processing**. High Distinctions in Advanced Modules in Natural Language Processing, including **Natural Language Syntax and Parsing** (91%, leading to subsequent appointment as guest lecturer) and **Introduction to Computational Semantics**. Part III thesis, **Less is More: Augmenting Small-Scale Language Models with Syntactic Inductive Biases**, published at BabyLM @ EMNLP 2024 (Miami, FL, USA).
- **Bachelor of Arts (Class I with Distinction, "Starred First"):** Class I with Distinction aggregate mark in my undergraduate studies. High Class 1 in every Tripos Paper in Finals, including 75+ in Linguistic Theory, Syntax/Semantics, Typology, Morphology and Phonological Theory, Computational Linguistics (83, 98% in practical assessment) and Part IIB Dissertation Published.
- **Examiner's Feedback on Dissertation:** "This is an **outstanding dissertation** which presents a new approach to the topic of computational language modelling: incorporating theoretical principles from the work of Theresa Biberauer (from the model of 'Maximise Minimal Means') into the framework of modern-day 'Transformer' neural networks. The argumentation and discussion are **sophisticated and accomplished**, the point of view is confident. This dissertation could serve as the basis for a future graduate research project which would empirically test the predictions made in this work. The conjunction of linguistic theory and computational modelling is a relatively neglected area which has great potential for exploration, as proposed and demonstrated in this dissertation. There is evidence of extensive background reading and research, supplemented with thorough critical analyses in which the relevant literature is handled with dexterity and acumen. Presentation and referencing are immaculate. The addition of a case study also demonstrates strong practical skills, the ability to handle large datasets, and perform analyses at the intersection of computational method and linguistic theory."

## PUBLICATIONS

---

**Research Interests:** INTERPRETABILITY, MULTILINGUALITY, TOKENIZATION, COMPUTATIONAL COGNITIVE SCIENCE & LINGUISTICS, ALIGNMENT

**Key Themes:** (i) Cognitively-Inspired Design, Interpretability and Evaluation (♣), (ii) Language Model Pretraining and Interpretability (★), (iii) Multilinguality (★), (iv) Tokenization (✿), (v) Alignment and Interaction (¤), and (vi) Cognitive Science and Linguistics (♦).

### CONFERENCE AND JOURNAL PUBLICATIONS

---

PEER-REVIEWED MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING CONFERENCE PUBLICATIONS

1. **Pico: A Modular Framework for Hypothesis-Driven Small Language Model Research.** Richard Diehl-Martinez, David Demitri Africa, Yuval Weiss, **Suchir Salhan**, Ryan Daniels, Paula Buttery. *EMNLP Systems Demonstration 2025*. Presentation in Suzhou, China. ★
2. **BabyBabelLM: A Multilingual Benchmark of Developmentally Plausible Training Data.** Jaap Jumelet, Abdellah Fourtassi, Akari Haga, Bastian Bunzeck, Bhargav Shandilya, Diana Galvan-Sosa, Faiz Ghifari Haznitrama, Francesca Padovani, François Meyer, Hai Hu, Julen Etxaniz, Laurent Prévôt, Linyang He, María Grandury, Mila Marcheva, Negar Foroutan, Nikitas Theodoropoulos, Pouya Sadeghi, Siyuan Song, **Suchir Salhan**, Susana Zhou, Yurii Paniv, Ziyin Zhang, Arianna Bisazza, Alex Warstadt, Leshem Choshen. *EACL 2026 (Main Conference)*. Presentation in Rabat, Morocco. ♣ ★ Co-Organising Multilingual Track Shared Task with Jaap Jumelet based on this paper for BabyLM Workshop @ EMNLP 2026.

### WORKSHOP PUBLICATIONS

---

1. **†Salhan, S.A.**, Diehl-Martinez, Richard, Goriely, Zebulon & Buttery, Paula (2024). **Less is More: Pre-Training Cross-Lingual Small-Scale Language Models with Cognitively-Plausible Curriculum Learning Strategies.** *CoNLL 2024 BabyLM Challenge (Paper Track)*. Presentation in Miami, Florida, USA (Nov 2024). ♣ ★
2. **ByteSpan: Information-Driven Subword Tokenisation.** Zebulon Goriely, **Suchir Salhan**, Pietro Lesci, Julius Cheng, Paula Buttery. *ICML 2025 Tokenization Workshop (TokShop)*. Delivered ByteSpan Poster Presentation in Vancouver, Canada (August 2025). ✿
3. **Measuring Grammatical Diversity from Small Corpora: Derivational Entropy Rates, Mean Length of Utterances, and Annotation Invariance.** Fermin Moscoso del Prado Martin, **Suchir Salhan**. *ACL 2025 Main Conference (Poster)*. Presented in Vienna, Austria (August 2025). ♦
4. **Meta-Pretraining for Zero-Shot Cross-Lingual Named Entity Recognition in Low-Resource Philippine Languages.** David Demitri Africa, **Suchir Salhan**, Yuval Weiss, Paula Buttery, Richard Diehl-Martinez. *5th Workshop on Multilingual Representation Learning (MRL), EMNLP 2025*. Presentation in Suzhou, China. ★ ★
5. **Teacher Demonstrations in a BabyLM's Zone of Proximal Development for Contingent Multi-Turn Interaction.** **Suchir Salhan**, Hongyi Gu, Donya Rooein, Diana Galvan-Sosa, Gabrielle Gaudeau, Andrew Caines, Zheng Yuan, Paula Buttery. *BabyLM Workshop, EMNLP 2025*. Presentation in Suzhou, China. ♣ ¤ Outstanding Paper Award
6. **Looking to Learn: Token-wise Dynamic Gating for Low-Resource Vision-Language Modelling.** Bianca-Mihaela Ganescu, **Suchir Salhan**, Andrew Caines, Paula Buttery. (Supervised MPhil Advanced Computer Science Thesis). *BabyLM Workshop, EMNLP 2025*. Presentation in Suzhou, China. ♣ Outstanding Paper Award
7. **BLiSS: Evaluating Bilingual Learner Competence in Second Language Small Language Models.** Yuan Gao, **Suchir Salhan**, Andrew Caines, Paula Buttery, Weiwei Sun. *BabyLM Workshop, EMNLP 2025*. Presentation in Suzhou, China. ♣

8. **What's the Best Sequence Length for BabyLM?** Suchir Salhan, Richard Diehl-Martinez, Zebulon Goriely, Paula Buttery. *BabyLM Workshop, EMNLP 2025*. Presentation in Suzhou, China. ★
9. **Pedagogical Alignment of LLMs Requires Diverse Cognitively-Inspired Student Proxies.** Suchir Salhan, Andrew Caines, Paula Buttery. *NeurIPS First Workshop on CogInterp: Interpreting Cognition in Deep Learning Models*, 2025. Presentation in San Diego, California, USA. ↗
10. **Theoretical Linguistics Constrains Hypothesis-Driven Causal Abstraction in Mechanistic Interpretability.** Suchir Salhan, Konstantinos Voudouris. *NeurIPS First Workshop on CogInterp: Interpreting Cognition in Deep Learning Models*, 2025. Presentation in San Diego, California, USA. ♠
11. **Glints of Gold or Troubling Waters? Can a School of Merged Monolingual Goldfish Models Swim in Bilingual Seas?** Suchir Salhan, EJ Zhou, Laura Barbenel, Aoife O'Driscoll, Lily Goulder, Lucas Resck, Catherine Arnett & Paula Buttery *EACL Workshop on Multilingual Multicultural Evaluation (MME), Non-Archival Full Paper*, 2025. Presentation in Rabat, Morocco. ★

#### WORKING PAPERS

1. **†Salhan, S.A. (2023) On the Potential for Maximising Minimal Means in Transformer Language Models: A Dynamical Systems Perspective.** In Cambridge Occasional Papers in Linguistics. Volume 15, Article 3: pp. 55–110. ISSN: 2050-5949. Available online [here](#).
2. **Salhan, Suchir**, Liu, Fangyu & Collier, Nigel. (2022/preprint) **Multimodal Language Modelling across Languages and Cultures: Grounding Strategies for Nouns and Verbs.** (Research Project, Language Technology Lab, Department of Theoretical and Applied Linguistics, University of Cambridge). [partially supported by a research award from Gonville & Caius College, Cambridge]
3. **The Distribution of Phonemes across Languages: Chance, costs, and integration across linguistic tiers.** 2026. Fermin Moscoso del Prado Martin, Suchir Salhan. 23rd Old-World Conference in Phonology (OCP23), Gonville & Caius College. (Accepted Oral). ♦
4. **Convergent Equilibria in Cross-Lingual Phoneme Surprisal Distributions: Statistical and Simulation-Based Analysis.** 2026. Suchir Salhan, Fermin Moscoso del Prado Martin. 23rd Old-World Conference in Phonology (OCP23), Gonville & Caius College. (Accepted Oral). Abstract ♦
5. **Extended Abstract for “Linguistic Universals”: Emergent Shared Features in Independent Monolingual Language Models via Sparse Autoencoders.** Ej Zhou, Suchir Salhan. 5th Workshop on Multilingual Representation Learning (MRL), EMNLP 2025. Presentation in Suzhou, China. ★
6. **SALHAN, S.A (2025) Linguistics in the Age of Language Models: What can Cognitively-Inspired Language Models offer to Linguistic Theory?** (POSITION PAPER in Cambridge Occasional Papers in Linguistics (CoPiL), Accepted, Volume 17) ♠

#### INVITED TALKS, SEMINARS & PRESENTATIONS

1. **Less is More: Pre-Training Cross-Lingual Small-Scale Language Models with Cognitively-Plausible Curriculum Learning Strategies** Presentations at Cambridge Language Sciences Symposium (November 2024), Poster at HumanCLAIM Workshop organised by Prof Lisa Beinborn in Gottingen Germany in March 2025. Accepted Poster and Demonstration at Cambridge CHIA (Centre for Human-Inspired AI) Annual Conference in June 2025.
2. **Human-Validated Grammar Profiles for Language Models.** Tübingen, Germany; March 2025 in a workshop organised by Prof Detmar Meurers

3. **The Distribution of Phonemes across Languages: Chance, Costs, and Integration across Linguistic Tiers.** Fermin Moscoso del Prado Martin, **Suchir Salhan**. *13th Conference on the Mental Lexicon*, 2025. Invited Keynote delivered by Fermin Moscoso del Prado Martin at McGill University, Montreal (June 2025). Slides. ♦
4. **Salhan, S.A. (2024)**, Diehl-Martinez, Richard, Goriely, Zebulon, Caines, Andrew & Butterly, Paula LLMs “off-the-shelf” or Pretrain-from-Scratch? Recalibrating Biases and Improving Transparency using Small-Scale Language Models. Learning and Human Intelligence Group, Department of Computer Science & Technology, University of Cambridge.
5. **Salhan, S.A. (2023)** **On the Potential for Maximising Minimal Means in Transformer Language Models**. Talk presented at SyntaxLab (14th February 2023), Department of Theoretical and Applied Linguistics, a departmental seminar organised by Dr Theresa Biberauer.
6. **Salhan, S.A. (2021)**. **UROP Project Report 2021: Providing Automatic Feedback on Argumentation Quality to Learners of English** (UROP Research Internship, Natural Language and Information Processing Group, Department of Computer Science and Technology, University of Cambridge). [supervised by Prof Paula Butterly, Dr Andrew Caines, Dr Russell Moore (NLIP Group); Dr Thiemo Wambsganss (University of St Gallen/EPFL) and funded by Cambridge Assessment]

---

## RESEARCH EXPERIENCE

---

2024-

### PicoLM

RESEARCHER

Supported by a generous research grant from the Accelerate Programme for Scientific Discovery at the University of Cambridge, I help develop Pico, an open-source toolkit for pretraining small language models with systematic checkpointing of model states, activations, and gradients for tracking learning dynamics over time for developmental interpretability of Language Models. Launched in March 2025, with over 100K+ views of launch video. Managing a team of Cambridge students working on PicoLM.

2022

### Language Technology Lab

RESEARCHER

I undertook an extended research project with Prof Nigel Collier and Fangyu Liu (Google DeepMind) in the Language Technology Lab (University of Cambridge). We worked on a creative probing of state-of-the-art multimodal language models. This research has, in part, been supported by Gonville and Caius College.

2021

### Department of Computer Science & Technology

RESEARCH ASSISTANT

I supported my supervisor Dr Li Nguyen (University of Cambridge) as a research assistant in the drafting of her computational sociolinguistics paper on variation in overt/null pronouns by Vietnamese-English bilingual speakers.

2021

### Department of Computer Science & Technology

UNDERGRADUATE RESEARCHER

I worked on a project in collaboration with Thiemo Wambsganss on developing the back-end machine learning architecture for an application that supports the argumentation skills of English language learners.

---

## TEACHING

---

### GUEST LECTURER

*Delivered first guest lecture at the University of Cambridge aged 22, with subsequent heavy involvement in University lecturing and teaching for Cambridge undergraduates and postgraduates.*

### L95 (ACS/PART III) INTRODUCTION TO NATURAL LANGUAGE SYNTAX AND PARSING. 2025-26.

- Led a substantial update and refactoring in course content, introducing weekly paper presentations, selecting and maintaining a course reading list. Course taught jointly with Prof Paula Butterly and Dr Fermin Moscoso del Prado Martin.

- Developed and delivered **two weeks of lecture content on language model tokenization and interpretability**.
- Developed a new practical assignment for the course (90% of the module assessment).

**Guest Lecturer for LI18 COMPUTATIONAL LINGUISTICS, 2025-26** (Part II Linguistics Tripos).

- Delivered *two guest lectures on Computational Linguistics* for second/third year undergraduates Course convened by Dr Guy Emerson.

**L95 (ACS/PART III) INTRODUCTION TO NATURAL LANGUAGE SYNTAX AND PARSING. 2024-25.**

- Delivered a lecture on **Language Model Evaluation and Mechanistic Interpretability** (Nov 2024). Course taught with Prof Paula Butterly and Dr Fermin Moscoso del Prado Martin.

#### **CAMBRIDGE UNIVERSITY UNDERGRADUATE ADMISSIONS INTERVIEWER**

Served as an Admissions Interviewer for Gonville & Caius College for the Computer Science and Linguistics Triposes (*December 2025*)

#### **TEACHING ASSISTANT**

L95 (ACS/Part III) Introduction to Natural Language Syntax and Parsing. Teaching Assistant 2024 - now

Machine Learning & Real World Data (Part IA, Computer Science Tripos). Teaching Assistant (2024-25)

#### **CAMBRIDGE UNDERGRADUATE SUPERVISION**

Machine Learning and Bayesian Inference (Part II, Computer Science Tripos)

Data Science (Part IB, Computer Science Tripos)

Formal Models of Language (Part IB, Computer Science Tripos)

Artificial Intelligence (Part IB, Computer Science Tripos)

Probability (Part IA, Computer Science Tripos)

L118 Computational Linguistics (Part IIA/IIB Linguistics Tripos)

College Supervisor for Linguistics Tripos (Gonville & Caius College) – Linguistic Theory (Part IIB, Linguistics Tripos), Part I Linguistics Tripos.

College Examiner for Computer Science Tripos Mock Examinations (Gonville & Caius College)

#### **RESEARCH SUPERVISION**

---

#### **MPhil ADVANCED COMPUTER SCIENCE THESIS SUPERVISION**

Bianca-Mihaela Ganescu (2024)

Yeqi Heo (2025)

#### **UNDERGRADUATE RESEARCH INTERNS**

Ellie Polyakova Reed (2025)

Shivan Arora (2025)

#### **RESEARCH ASSISTANTS**

ALTA Institute 2025-26 Research Assistants (Laura Barbenel, Aoife O'Driscoll, Lily Goulder), line-managed by Prof Andrew Caines and Prof Paula Butterly.

#### **VISITING STUDENTS**

Andrzej Szalewski (working jointly with Dr Tiago Pimentel and Dr Clara Meister)

#### **RESEARCH MENTOR**

Ali Kheirkhah (Summer 2025)

*I was a PicoLM Research Mentor as part of the Google DeepMind Research Ready Programme*

## OTHER ACTIVITIES

---

2023-

### Co-Founder and Editor-in-Chief of Per Capita Media

MEDIA & JOURNALISM

Per Capita is a progressive national student publication founded in Cambridge University in collaboration with University of Art's London, covering News, Features and Culture. President of Cambridge University Per Capita Media (CU PCM). Affiliated College Societies include The Newnham Political ("The Poly", First established in 1884, now revived in 2025) and the Caius Media & Journalism Society (CMJS).

2026-2027

### Gonville & Caius College MCR

COMMUNITY

I was elected Vice President of Caius MCR in 2026

2025-2026

### Gonville & Caius College MCR

COMMUNITY

I was elected Treasurer of Caius MCR in 2025

2024-

### Founder and President of Gonville & Caius Media and Journalism Society (CMJS)

MEDIA &

JOURNALISM

2023-

### The Wilberforce Society Head of Policy, Editor & Head of Events

POLICY

Head of Policy (2024-2026) for The Wilberforce Society (TWS), the UK's oldest student think tank in the UK based at the University of Cambridge. Editor of Cybersecurity Policy Paper for [policy thinktank](#), published in Nov 2023 and presented in St John's College, available online [here](#). Organised **Freedom of Press Panel Event in St John's College** and a panel event with the Foreign Minister of Sri Lanka.

2024

### The Cambridge Student

MEDIA & JOURNALISM

Deputy Editor— the student paper of Cambridge University.

2024

### Gonville & Caius College MCR

COMMUNITY

I was Returning Officer of Caius MCR in 2024, running college committee elections.

2023

### The Cambridge Student

MEDIA & JOURNALISM

News & Investigations Editor and Interviews writer— the student paper of Cambridge University.

2023

### BBC The One Show

MEDIA & JOURNALISM

Invited by producers at The One Show to help produce a documentary about higher education post-COVID. Responsible for organising student contributors around the country. Also, invited by producers at BBC Radio 5Live to discuss the Marking and Assessment Boycott.

2023

### Radio Presenter

MEDIA & JOURNALISM

Qualified Radio Presenter at CamFM.

2022-23

### Varsity Newspaper

MEDIA & JOURNALISM

Member of the Editorial Team (Investigations Editor, Senior Science Editor, News Correspondent, 'Chief Sub-editor') of Varsity— the student paper of Cambridge University.

2022

### Gonville & Caius College JCR

COMMUNITY

I was elected Secretary of Caius JCR in 2022-23

2021-2024

### University of Cambridge

SOCIETIES

CamFM, Caius Boat Club (Novice), Member of Cambridge Law Society (CULS), Cambridge Investment Banking Society (CIBS), Cambridge Union, Sub-editor and Author of Polyglossia Magazine 2020-2021, Member of Gonville and Caius College Access Team (helping in open days, access events).

## AWARDS & SCHOLARSHIPS

---

2025

EMNLP 1ST BABYLM WORKSHOP

**TWO Outstanding Paper Awards** from 1st BabyLM Workshop at EMNLP

2024

DEPARTMENT OF COMPUTER SCIENCE & TECHNOLOGY

**Fully-Funded PhD Studentship Award** from the Department of Computer Science and Technology for research on mitigating risks of Large Language Models.

2023

GONVILLE & CAIUS COLLEGE

**Memorial Prize**, one of two academic accolades awarded to Gonville & Caius graduating cohort 2023

2023

GONVILLE & CAIUS COLLEGE

**Examination Prize**

2023

GONVILLE & CAIUS COLLEGE

**Re-election to Senior Scholarship**

2022

GONVILLE & CAIUS COLLEGE

**Senior Scholarship**

2022

GONVILLE & CAIUS COLLEGE

**Examination Prize** for First Class Tripos Performance

2022

GONVILLE & CAIUS COLLEGE

**Research Award** awarded by the Senior Tutor of Gonville and Caius (Dr Andrew Spencer), for my research project with Prof Nigel Collier.

2022

COMPUTER SCIENCE DEPT.

Funding of £175 for **Research Assistant** work

2021

COMPUTER SCIENCE DEPT.

Funding by Cambridge Assessment for a UROP in the Department of Computer Science & Technology, worth £3000. I was the only first-year student in the entire University to participate in the two/three month Natural Language Processing UROP program.

## PROFESSIONAL ACTIVITY

---

- **Organiser and Host of the Natural Language & Information Processing Seminars, 2024 -.** **Natural Language & Information Processing Group (Department of Computer Science & Technology).** Organising 30+ departmental seminars with leading academics and industry researchers on Language Models, Computational Linguistics and Natural Language Processing.
- **University-Wide & Interdisciplinary Initiatives Language Sciences Annual Symposium 2025: Ambitions for language science in 2050.** Poster Session Organiser for 2025 Cambridge Language Sciences Symposium with Sammy Weiss (MRC Cognition and Brain Sciences Unit) and Shanshan Hu (TAL).
- **23rd Old-World Conference in Phonology (OCP23). Member of Organising Committee.** Gonville & Caius College (January 2026). Phonetics Laboratory, Department of Theoretical & Applied Linguistics. *Managed and co-ordinated a large conference budget (£10,000+) for a well-attended linguistics conference (over 100 attendees from across the world).*

## ACADEMIC REVIEWING AND SERVICE

---

Reviewer for BabyLM 2024. ACL 2025 Emergency Reviewer. Reviewer for The First Workshop on Large Language Model Memorization – L2M2 Proceedings @ ACL 2025. Reviewer for NEURIPS CogInterp Workshop. Reviewer for NEURIPS What Can't Transformers Do (WCTD) Workshop.

## SKILLS

---

- **Programming Languages:** Python (proficient); basic experience in MATLAB.
- **Machine Learning Frameworks:** NumPy, SciPy, matplotlib, keras, scikit-learn, tensorflow
- **Natural Language Processing Frameworks/Libraries:** NLTK, SpaCy, HuggingFace libraries