Swarnendu **Moitra**

Data Scientist · Linguist · Neuroscientist · Electronics Engineer

✉ s.moitra@qmul.ac.uk | 🏠 swarmoi.github.io/ | swarmoi | linkedin.com/in/swarnendumoitra | 📍 London UK

Work Experience

Pre Doctoral Research Assistant on the SaVANT Project *London, United Kingdom*

Queen Mary University of London | **New York University, Abu Dhabi** *Aug. 2021 – present*

1. Spearheaded groundbreaking research in neurolinguistics, leading projects from conception to publication (**1 peer-reviewed paper**) and dissemination (**10 conference presentations** & **2 invited talks**).
2. Recruited over 200 participants for diverse neuroimaging (MEG/EEG) and behavioural experiments conducted across two universities.
3. Leveraged **machine learning algorithms** (in **Python** & **R**) to **analyze and build statistical model** of complex experimental data, extracting key insights.
4. Established an EEG/MEG data processing pipeline that streamlines neural data acquisition, pre-processing, analysis, and visualization (using open source python packages) for cognitive neuroscience experiments.
5. Empowered over 50+ students (including undergraduates) from diverse backgrounds with a foundational understanding of linguistics & neuroscience through summer school programs.
6. <u>Publications</u>: ① **Moitra, S. et al. (2024).** `How long is long? Word length effects in reading correspond to minimal graphemic units: An MEG study in Bangla.` **PLOS ONE, 19(4), e0292979** [Link].

Research Data Scientist on the Democratizing AI Project *London, United Kingdom*

Queen Mary University of London | **The Alan Turing Institute, UK** *Feb. 2024 – present*

1. Designed and developed an innovative Word Sense Disambiguation dataset addressing polysemy and contextual representation issues.
2. Enhanced LLM performance for under-represented languages, boosting accuracy by 40% through Indian language-specific fine-tuning
3. Conducted a comprehensive evaluation of pre-trained contextualised models (e.g., mBERT, XLM-R) to benchmark effectiveness
4. Collaborated with interdisciplinary teams to enable scalable multilingual data processing and improve operational efficiency in text analysis workflows

Education

M.A in Linguistics – Jadavpur University (Distinction, GPA: 9.53/10) *2017 – 2019*

- **Dissertation:** Neurocognition and Neurophysiology of Scope Ambiguity in Head Final Languages.
- Co-founded **LingClub**: a student-led body to organize colloquiums and reading groups.

B.Tech in Electronics and Communications Engineering – M.A.K.A.U.T (GPA: 7.67/10) *2013 - 2017*

- **Dissertation:** Study and analysis of data encryption techniques and implementation using MATLAB.

Skills

Programming	**python** [Scikit-learn, tensorflow, spaCy, eelbrain, mne-python], **R** [tidyverse, brms, lme4], **JavaScript**
Mark-up	LaTeX, (R)Markdown, quarto, HTML
Experiments	jsPsych, presentation, Linger, OpenSesame, PsychoPy, Ibex Farm
Cloud Platforms	High Performance Computing: Apocrita, AWS SageMaker
Data Science & Misc.	Data science pipeline (cleansing, wrangling, visualisation, modelling, interpretation), Statistics, A/B testing, ETL, Time series, Experimental design, Hypothesis testing, OOP, Excel, Git, latex

Research Projects

Effects of scope interactions : An experimental study on Scrambling | [Project Website](#)
Dr. Paul Marty (L-Università ta' Malta), Dr. Ishani Guha (Indian Institute of Technology - Delhi)

Feb 2020 - Ongoing

Collaborator(s)

1. **Co-designed and executed** behavioural psycholinguistic experiments to investigate the influence of *syntactic ambiguity* in sentence processing on speakers of Bangla from India and English from the UK.
2. Designed 100+ experimental stimuli, utilizing the jsPsych framework to build the online experiment.
3. Deployed the experiment on a local Linux server with recruitment platforms (Amazon Mechanical Turk/ Prolific) and implemented Node.js backend for real-time data collection and storage.
4. Presented findings at **3** international conferences (2 proceedings published). The research is currently under peer-review in a high-impact journal.

lexiVAULT - A Lexical database for understudied languages
Dr. Samantha Wray (New York University, Abu Dhabi & Dartmouth College)

Hind Saddiki (New York University, Abu Dhabi)

Aug 2021 - Ongoing

Collaborator(s)

1. Enriched datasets by developing web scraping tools that integrated with existing data collections, enabling comprehensive text analysis.
2. Facilitated cross-lingual text analysis by constructing language-specific stemmers, improving efficiency by 82%
3. Automated feature extraction and advanced lexical statistics calculation through the design and implementation of machine learning models.
4. Promoted collaboration and wider use by developing a user-friendly, open-source interface for public access and utilization of these tools.

Stock Market Volatility of the G20 Countries: Effects of the 2007 Financial Crisis and Covid-19
Ms. Eva Mondal (University of York)

Dec 2023 - March 2024

Collaborator

1. Conducted an in-depth investigation of the effects of the 2007-08 Financial Crisis and COVID-19 pandemic on stock market volatility by analyzing 20 years of monthly stock market data from 8 countries from G20.
2. Performed a comprehensive **econometric analysis** of stock returns and volatility to identify trends.
3. Implemented a **fixed-effects panel regression model** to isolate the impact of the crises while controlling for country-specific factors.
4. Tested **hypotheses** on the differential impacts of the crises on stock market volatility, uncovering varied responses between Western and Asian countries.
5. The analysis revealed that during the COVID-19 pandemic, Asian countries experienced smaller increases in volatility compared to Western countries, whereas during the 2007-08 Financial Crisis, Asian countries exhibited significantly larger increases in volatility.

Decoding Visual Stimulus from Visual Cortex Data | [Neuromatch Project](#)
Fatemeh Abdollahi (University of Alberta), Rekha Crawford (Swarthmore College)

Sreekar Reddy (University of Cincinnati), Pranjal Gupta (Indian Institute of Technology, Bombay)

July 2020

Collaborator(s)

1. Leveraged Neuropixel data from the mouse visual cortex (Steinmetz dataset) to decode visual stimuli using machine learning classifiers (LDA, Random Forest) during a Neuromatch Academy group project.
2. Classified stimulus vs. no stimulus conditions and left vs. right visual contrasts with high accuracy using firing rate data. Random Forest achieved the best performance among the employed classifiers.

Publications & Presentations

Invited Talks

Tracing word length in non-alphabetic orthographies
Language and Mind, 30th November 2023

UK Research and Innovation (UKRI)

London, UK

Understudied languages in neuroscience: from Slovenian to Arabic and more
University of Ljubljana, 18th May 2023

V nevroznanosti šibkeje zastopani jeziki: od slovenščine do arabščine in še naprej

Ljubljana, Slovenia

Publications & Proceedings

Beyond the Left Hemisphere: MEG Evidence for Right Temporal Lobe recruitment in Processing Morphosyntax

osf.io/3bc69

WITH DUSTIN A. CHACÓN AND LINNAEA STOCKALL

Submitted

Democratizing AI: Widening the WiC Lens, A multilingual Polysemy Dataset Exposes Cross-Lingual Gaps

HAIM DUBOSSARSKY, ROKSANA GOWOREK, SWARNENDU MOITRA AND 12 OTHERS

2025

in Proceedings of the `7th Workshop on Research in Computational Linguistic Typology and Multilingual NLP`

Scope preferences of universal quantifiers: a comparison between English and Bengali

GUHA, ISHANI, SWARNENDU MOITRA, & PAUL MARTY.

2025

Glossa, `Under Review 2`

How long is long? Word length effects in reading correspond to minimal graphemic units: An MEG study in Bangla.

[Preprint] [doi]

MOITRA, SWARNENDU, DUSTIN A. CHACÓN, & LINNAEA STOCKALL.

2024

`PLOS ONE` 19(4), e0292979.

Inverse Scope in Scrambling Languages: the case of Bangla [paper] [Book]

ISHANI GUHA, SWARNENDU MOITRA, PAUL MARTY

2024

in Proceedings of the `39th West Coast Conference on Formal Linguistics.`

Interaction between Word Order, Quantifier Scope and Lexical Information: Experimental Evidence from Bangla.

ISHANI GUHA, SWARNENDU MOITRA, PAUL MARTY

Forthcoming

in Proceedings of `the 57th Annual Meeting of the Chicago Linguistic Society.`

Study and analysis of data encryption techniques and implementation using MATLAB programming

[link]

ANANYA BANERJEE, SWARNENDU MOITRA, URMEE KARMAKAR, SAYAN GHOSH, SAHELI DATTA AND SANDIPAN DEBNATH

2017

`International Journal of Advanced Research in Computer Engineering and Technology (IJARCET)`

Volume 06, Issue 05, May 2017, ISSN : 2278 - 1323

Book Chapter

Three's a Crowd: An Experimental Investigation Into Person Features

JANE MIDDLETON & SWARNENDU MOITRA

Forthcoming

Artificial language learning. In D. Harbour (ed.), Oxford Handbook of Experimental Syntax.

Manuscripts in Prep

Experimental Linguistics in Low-Resource Contexts: Building Lexical Tools for Bangla

[Tool]

WITH SAMANTHA WRAY AND HIND SADDIKI

in prep

Impossible Numbers: Experimental Investigations into the Marked Values of Phi Features

osf.io/b7yex

WITH JANE MIDDLETON AND DANIEL HARBOUR

in prep

Conference Presentations (Talks/Posters)

Beyond the Left Hemisphere: MEG Evidence for Right Temporal Lobe Recruitment in Bangla Morphosyntax Processing

Edinburgh, Scotland

SWARNENDU MOITRA, DUSTIN A. CHACÓN AND LINNAEA STOCKALL

`Architectures and Mechanisms for Language Processing [AMLaP 2024],`

University of Edinburgh, Scotland. 2024

Syntactic licensing and semantic interpretation processes occur in parallel in complex word recognition in Abugidas

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

International Word Processing Conference [WoProc 2024],

University of Belgrade, Serbia. 2024

Belgrade, Serbia

Tracing word length in non-alphabetic orthographies

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

International Word Processing Conference [WoProc 2024],

University of Belgrade, Serbia. 2024

Belgrade, Serbia

Towards developing a functional localiser for letter-specific brain responses for Bangla: An MEG Study

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

Annual Conference of Cognitive Science [ACCS - 10],

Indian Institute of Technology - Kanpur, 9 - 11 December, 2023

Kanpur, India

Syntactic vs. semantic computation in posterior temporal lobe: an MEG study on Bangla nominal prefixes.

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

The Society for the Neurobiology of Language [SNL-2023],

The Palais du Pharo, Marseille, 24 - 26 October, 2023

Marseille, France

Morphological decomposition in bilateral fusiform areas in Bangla: An MEG study.

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

The Society for the Neurobiology of Language [SNL-2023],

The Palais du Pharo, Marseille, 24 - 26 October, 2023

Marseille, France

The Effect of Determiner Type on Quantifier Scope Ambiguity in English and Bengali

Ishani Guha, Swarnendu Moitra, and Paul Marty

Human Sentence Processing [HSP - 36],

University of Pittsburgh, Pittsburgh, 8 - 11 March, 2023

Pittsburgh, USA

How long is long? An MEG study of M100 response in tracking word length in Bangla/Bengali

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

9th Annual Conference of Cognitive Science [ACCS-9],

Indian Institute of Technology, Delhi, 8 - 10 December, 2022

New Delhi, INDIA

How long is long? M100 response in Bangla tracks number of phonemes not graphemes of glyphs

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

The Society for the Neurobiology of Language [SNL-2022],

Loews Philadelphia Hotel, Philadelphia, 6 - 8 October, 2022

Philadelphia, USA

Morphological recomposition and the concrete/abstract distinction in Bangla

Swarnendu Moitra, Dustin A. Chacón and Linnaea Stockall

Architectures and Mechanisms for Language Processing [AMLaP-2022],

The Principal, York, 7 - 9 September, 2022

York, UK

Complex word de/recomposition across languages: lexical decision and MEG studies

Linnaea Stockall, Dave Kenneth Cayado, Dustin Alfonso Chacón, Karin Kavcic, Swarnendu Moitra, Christina Manouilidou, Alec Marantz, Suhail Matar, Bojana Ristić, Samantha Wray and Rok Žaucer

The Crosslinguistic Perspectives on Processing and Learning [X-PPL 2022],

University of Zurich, 12 - 13 September, 2022

Zurich, Switzerland

Interaction between Word Order, Quantifier Scope and Lexical Information: Experimental Evidence from Bangla

Illinois, USA

Ishani Guha, Swarnendu Moitra and Paul Marty

Annual Meeting of the Chicago Linguistic Society 57 [CLS - 57],
University of Chicago, 6 - 8 May, 2021

Inverse Scope in Scrambling Languages: the case of Bangla

Tucson, USA

Ishani Guha, Swarnendu Moitra and Paul Marty

West Coast Conference on Formal Linguistics 39 [WCCFL - 39],
University of Arizona, 8 - 11 April, 2021.

In Search of Inverse Scope in Bangla

Minneapolis, USA

Ishani Guha, Swarnendu Moitra and Paul Marty

11th (Formal) Approaches to South Asian Languages, [(F)ASAL - 11]
University of Minnesota 26th - 28th March 2021

Internships & Summer Schools

Internships

2018 - 19	**Intercultural Communication in Multilingual Urban Spaces - Uni of Calcutta**	*Kolkata*
2019	**Neuro Cognition Lab - Sir C V RAMAN Centre for Physics and Music - Jadavpur University**	*Kolkata*
2019	**Indian Statistical Institute - Kolkata**	*Kolkata*

Summer Schools

2022	**SMLP 2022,** Summer School on Statistical Methods for Linguistics and Psychology, **Uni of Potsdam**	*Germany*
2021	**CuttingEEG,** 5th symposium for cutting edge methods on EEG research, **Aix-en-Provence**	*France*
2020	**Neuromatch Academy,** Computational Neuroscience (Org. by : **UPenn, DeepMind and others**)	*Online*
2019	**ESSLLI - 31,** European Summer School in Logic, Language and Information, **Uni of Latvia**	*Riga, Latvia*

Grants and Achievements

United Kingdom Research and Innovation (UKRI) - Global Talent

2021

Endorsed by **UKRI**, sponsored by the Department for Science, Innovation and Technology (DSIT) in support of the Global Talent Migrant Visa for conducting Neuroscience research in the UK.

Rashtriya Uchchatar Shiksha Abhiyan (RUSA) - Travel Grant for Overseas Conference

2019

Awarded **1125 USD** Travel Grant in support for paper presentation at **25th Himalayan Languages Symposium**

Neuromatch Academy Summer School

2019-20

Awarded Full scholarship to attend the Summer School

Swami Vivekananda Merit-cum-Means Scholarship

2019-20

Awarded **850 USD** scholarship per-annum for MPhil.

University Grants Commission (UGC) - (NET) National Eligibility Test in Linguistics

2018

Qualified for Assistant Professor in India

Services

Reviewer

2025	**Curriculum Review,** Reviewed undergraduate and postgraduate Linguistics curricula for **Amity School of English Studies and Research, AUUP NOIDA**, providing detailed feedback on course design, content, and learning outcomes.

Teaching - Queen Mary University of London

2024   **Course Instructor**, 25 pupils | Experimental Linguistics Taster Day - 12th June
2023   **Course Instructor**, 9 pupils | Winter Workshop on Neurolinguistics
2023   **Teaching Assistant**, 4 pupils | London Summer School on Neurolinguistics
2022   **Teaching Assistant**, 9 pupils | London Summer School on Neurolinguistics

Others

2022   **London School Outreach**, London Summer School on Neurolinguistics, **Queen Mary University of London**
2020   **Organizing Committee**, Rhythm in Speech and Music Conference, **Jadavpur University**

Extracurricular Activity

LingClub, Jadavpur University   *Jadavpur University*

Co-founder and Coordinator   *Aug. 2018 - 2021*

- Organised student lead Colloquiums, Reading Groups and Outreach programmes

Hindustani Classical Flautist

Trained in North - Indian Style Classical Music

- Completed 2nd Year Diploma in Hindustani Classical Music : Bamboo Flute (Bansuri)

Referees

Dr. Dustin A. Chacón (PhD, University of Maryland)   *UC Santa Cruz*

Assistant Professor, Department of Linguistics

duchacon@ucsc.edu

Prof. Linnaea Stockall (PhD, Massachusetts Institute of Technology)   *Queen Mary University of London*

Professor, Department of Linguistics

l.stockall@qmul.ac.uk