

Rajaswa Patil

Pre-Doctoral Researcher, TCS Research










🌐 [rajaswa.github.io](https://github.com/rajaswa) @ patilrajaswa@gmail.com 🌐 [rajaswa](#) 🎓 [Google Scholar](#) [in](#) [rajaswa-patil](#)
📍 TCS TRDDC, 54-B, Hadapsar Industrial Estate, Pune, MH, India. 411013

Education





Jun 2021 Aug 2017	Birla Institute of Technology and Science (BITS) Pilani B.E., ELECTRICAL & ELECTRONICS CGPA: 8.26/10.00 Thesis: Modelling Coherence in Spoken Discourse Study in Advanced Topics: [1] Meta Learning; [2] Computational Neuroscience	Goa, India
----------------------	--	------------

Experience

Present Jul 2021	Tata Consultancy Services (TCS) - Research Deep Learning & AI Group 🌐 Pre-Doctoral Research Fellow Advisors: <i>Dr. Gautam Shroff, Dr. Lovekesh Vig, Dr. Shirish Karande</i> ➤ Working on the broad research topics: NLP for Programming, Cross-domain Semantic Parsing, and Program Synthesis.	Pune, India
Jun 2021 Apr 2021	Cognizer AI 🌐 Data Scientist (NLP) - Intern Manager: <i>Vineeth T M</i> ➤ Developed APIs & data resources for Question-generation, Grammar-correction, and Text-simplification for the Cognizer Genius platform.	Bengaluru, India
Jun 2021 Apr 2021	Max Planck Institute for Human Cognitive & Brain Sciences 🌐 Summer Research Intern Advisor: <i>Dr. Lars Meyer</i> ➤ Experimented with Combinatory Categorical Grammar - CCG parsers (left-branching-based incremental parsing) for modelling human-like language processing.	(Remote) Leipzig, Germany
Mar 2021 Aug 2020	IIIT Delhi Multimodal Digital Media Analysis (MIDAS) Lab 🌐 Research Assistant (Bachelor Thesis) Advisors: <i>Prof. Rajiv Ratn Shah, Yaman Kumar Singla</i> ➤ Worked on modelling information structure and coherence in spontaneous spoken discourse. ➤ Developed annotation pipelines & data resources for disfluency and content relevance & sufficiency for ESL speech data.	Delhi, India
Dec 2019	Research Intern Advisors: <i>Prof. Rajiv Ratn Shah, Yaman Kumar Singla</i> ➤ Contributed to the Berlitz SLTI automatic language proficiency - speech scoring system.	
May 2021 Aug 2019	BITS Pilani (Goa Campus) Cognitive Neuroscience Lab 🌐 Research Assistant Advisor: <i>Prof. Veeky Baths</i> ➤ Developed a benchmark for syntactic evaluation of multilingual language models for Indic languages. ➤ Lead projects on Mental Lexicon Development and EEG-based brain imagery. ➤ Organized and volunteered for various events under the ERASMUS iBRAIN project. ➤ Designed and implemented an experiment for EEG data collection (6 subjects with 200 trials each) to inspect the role of low-level visual representations (BA-18 & BA-19) in spoken word processing.	Goa, India
May 2021 Dec 2019	Indian Institute of Technology, Bombay Department of CS&E 🌐 Summer Research Intern Advisor: <i>Prof. Ganesh Ramakrishnan</i> ➤ Implemented a NLP pipeline for open-domain opinion-mining & aspect-based sentiment-analysis in natural language text. ➤ Built various text-based utilities for entity resolution, topic extraction, theme extraction, and lexicon-based sentiment polarity extraction.	Mumbai, India

- [S.2] **Towards Using Diachronic Distributed Word Representations as Models of Lexical Development** 
Arijit Gupta*, Rajaswa Patil*, Veeky Baths
In Submission OpenReview
- [S.1] **Towards Modelling Coherence in Spoken Discourse** 
Rajaswa Patil*, Yaman Kumar Singla*, Rajiv Ratn Shah, Mika Hama, Roger Zimmermann
Working Paper arXiv
- [C.1] **DRIFT: A Toolkit for Diachronic Analysis of Scientific Literature**  
Abheesht Sharma*, Gunjan Chhablani*, Harshit Pandey*, Rajaswa Patil
The 2021 Conference on Empirical Methods in Natural Language Processing (Virtual) EMNLP '21
- [W.4] **Vyākaraṇa: A Colorless Green Benchmark for Syntactic Evaluation in Indic Languages**  
Rajaswa Patil, Jasleen Dhillon, Siddhant Mahurkar, Saumitra Kulkarni, Manav Malhotra, Veeky Baths
The First Workshop on Multilingual Representation Learning (Virtual) MRL '21
- [W.3] **BPGC at SemEval-2020 Task 11: Propaganda Detection in News Articles with Multi-Granularity Knowledge Sharing and Linguistic Features Based Ensemble Learning**  
Rajaswa Patil, Somesh Singh, Swati Agarwal
The Fourteenth Workshop on Semantic Evaluation (Virtual) SemEval '20
- [W.2] **CNRL at SemEval-2020 Task 5: Modelling Causal Reasoning in Language with Multi-Head Self-Attention Weights Based Counterfactual Detection**  
Rajaswa Patil, Veeky Baths
The Fourteenth Workshop on Semantic Evaluation (Virtual) SemEval '20
- [W.1] **LRG at SemEval-2020 Task 7: Assessing the Ability of BERT and Derivative Models to Perform Short-Edits Based Humor Grading**  
Siddhant Mahurkar*, Rajaswa Patil*
The Fourteenth Workshop on Semantic Evaluation (Virtual) SemEval '20
- [P.2] **Word2Brain2Image: Visual Reconstruction from Spoken Word Representations**  
Ajay Subramanian, Rajaswa Patil, Veeky Baths
The Sixth Annual Conference of the Association for Cognitive Science, 2019 ACCS '19
- [P.1] **Citta: A Lite Semantic Recommendation Framework for Digital Libraries**  
Rajaswa Patil, Siddhant Mahurkar
The NDLI-UNESCO International Symposium on Knowledge Engineering for Digital Library Design 2019 KEDL '19

Select Research Projects

- Semantic Parsing: A Living Literature Review**   Sep'21 - Present
- > A website covering modelling and data resources for Semantic Parsing in NLP.
 - > Covers the Text-to-SQL semantic parsing task.
 - > Covers the Abstract Representation parsing task (Ex: UCCA, AMR, DRG, EDS, etc.).
 - > Covers the Shallow Semantic parsing tasks (Ex: semantic role labelling).
- Diachronic Analysis of Scientific Literature (DRIFT)**   Mar'21 - Present
- > DRIFT is a toolkit for Diachronic analysis of scientIFic liTerature. [Demo @ EMNLP'21]
 - > DRIFT offers UI utilities for convenient training of Diachronic Embeddings for Scientific Corpora.
 - > DRIFT covers **nine** analysis methods: *word cloud*, *productivity & frequency plot*, *acceleration plot*, *semantic drift*, *tracking clusters*, *acceleration heatmap*, *similarity trends*, *keyword visualization*, and *LDA topic modeling*.
 - > **Presented at:** EMNLP'21, Aon Research.

Academic Service

(* = sub-reviewer)

Program Committee	ACL-SRW'22, MRL'21
Reviewer	ACL-SRW'22, ICON'21*, MRL'21, EAAI'21*
Volunteer	ACL'21, Neuromatch'21

Talks

“Vyākaraṇa: A Colorless Green Benchmark for Syntactic Evaluation in Indic Languages”

- > MRL EMNLP 2021 Presentation [📺] November 2021 (Remote)
- > Invited Talk - BITS Goa September 2021 (Remote)
- > ACL-IJCNLP SRW 2021 Presentation [📺] August 2021 (Remote)

“DRIFT: A Toolkit for Diachronic Analysis of Scientific Literature”

- > EMNLP 2021 Presentation [📺] November 2021 (Remote)
- > Invited Talk - Aon Research October 2021 (Remote)

“Syntactic Processing and Language Comprehension”

- > Round Table Discussion - CCCP Conference December 2020 (Remote)

“Getting Started with ML/NLP/Undergraduate Research”

- > Getting Started with ML - ACM BITS Goa Chapter [📺] July 2021 (Remote)
- > Thesis Talks - CTE BITS Goa [📺] May 2021 (Remote)
- > Introduction to NLP Research - TIP BITS Goa September 2020 (Goa, India)

Teaching and Leadership Roles

Language Research Group (LRG) - BITS Goa *Founder & Lead* [📺]

Aug'19 - Present

- > Established a student research group for NLP & Psycholinguistics
- > The group has currently published 11 papers & three projects

Birla Institute of Technology and Science (BITS) Pilani, Goa Campus *Teaching Assistant*

Aug'20 - May'21

- > BITS F315: Introduction to Cognitive Neuroscience
- > BITS C382: Neuroscience and AI

QSTP & CTE (student-run clubs at BITS Goa) *Instructor*

May'19 - Dec'19

- > Introduction to Machine Learning & Deep Learning
- > Deep Learning with PyTorch

Software and Open Source Contributions

- > **ML4Code**: Machine Learning for Big Code and Naturalness - [commits](#), [website](#)
- > **Semantic Parsing**: Modeling and Data resources for Semantic Parsing - [commits](#), [website](#)
- > **DecepticoNLP**: Python Library for Robustness Monitoring and Adversarial Debugging of NLP models - [commits](#)

References

- > Dr. Gautam Shroff *Senior Vice President, Chief Scientist, and Head - TCS Research* [📺]
- > Dr. Lovekesh Vig *Principal Scientist - TCS Research* [📺]
- > Dr. Shirish Karande *Senior Scientist - TCS Research* [📺]
- > Vineeth T M *Data Science Architect - Cognizer AI* [📺]
- > Dr. Lars Meyer *Research Group Leader - Max Planck Research Group Language Cycles* [📺]
- > Prof. Veeky Baths *Associate Professor - Dept. of Biological Sciences, BITS Goa* [📺]
- > Prof. Rajiv Ratn Shah *Assistant Professor - Dept. of Computer Science & Engineering, IIIT Delhi* [📺]
- > Yaman Kumar *Research Scientist - Adobe Media Data Science Research* [📺]
- > Prof. Swati Agarwal *Assistant Professor - Dept. of Computer Science & Information Systems, BITS Goa* [📺]
- > Prof. Ganesh Ramakrishnan *Chair Professor - Dept. of Computer Science & Engineering, IIT Bombay* [📺]