

Romi Hill

+1 (617) 659 1235 ◇ romihill@bu.edu ◇ *github*

EDUCATION

Boston University

PhD in Linguistics

September 2024 - 2029 (expected)

Boston, Massachusetts

University of Konstanz

Master of Speech and Language Processing

October 2021 - February 2024

Konstanz, Germany

Specialization: Machine Language Processing

Thesis Title: Syntactic Constraints on the Morphophonology of Reduplication: A Theoretical and Computational Perspective

Thesis supervisors: Junior Professor Colin Davis and Dr. Tina Bögel

Final Grade: 1.0 (4.0 GPA equivalent)

University of Melbourne

Bachelor of Arts (Honours) & Concurrent Diploma in Mathematics

February 2015 - December 2019

Melbourne, Australia

Major: Linguistics & Applied Linguistics

Diploma Specialisation: Discrete Mathematics and Operations Research

Thesis Title: Murrinhpatha Noun Classifiers - Syntax and Discourse Functions

Thesis supervisors: Professor Rachel Nordlinger and Dr. John Mansfield

Final Grade: 93

Ludwig Maximilian University

Exchange Student

March 2017 - July 2017

Munich, Germany

RESEARCH INTERESTS

Interactions at the morphosyntactic and syntax-phonology interface, reduplication, subtractive morphology

RESEARCH ASSISTANT EMPLOYMENT

Prosodic Structure at the Interface Between Language and Speech

Project leads:

Dr. Tina Bögel (University of Konstanz),

Professor Catherine Lai & Professor Alice Turk (University of Edinburgh)

March 2022 - March 2024

University of Konstanz

This joint 3-year DFG-AHRC project seeks to investigate the relationship between prosodic structure and language redundancy effects. Both language redundancy and prosodic structure affect characteristics of speech related to acoustic salience and intelligibility. The idea that prosodic structure is the interface between grammar and speech led to proposals that speakers plan the prosodic structure of an utterance in order to make less predictable parts of speech more intelligible via its effects on acoustic saliency. The goals of this project are 1) to test the hypothesis that prosodic structure is the interface between language and speech, and 2) to implement this interface in a (computational) LFG-grammar in order to test and evaluate our architectural assumptions. My responsibilities include:

- Working with the team to develop an interface between prosody and syntax that includes language redundancy information, in the Lexical Functional Grammar framework.

- Implementing this interface in a computational XLE grammar.
- Writing a Python script to calculate the average duration of syllables in German, which will be used as a measure of acoustic salience.

Information Structure and Questions in Urdu/Hindi

December 2021 - January 2022

Project lead

University of Konstanz

Professor Miriam Butt (University of Konstanz)

This DFG funded project aims to investigate how such non-canonical questions are structured lexically, morphologically, grammatically and prosodically and how they are perceived.

I was employed as a research assistant to write Praat scripts for Urdu phonetic transcriptions. These scripts performed the following tasks:

- Calculate the average duration of phonemes, syllables, words, and utterances.
- Calculate the mean intensity of words.
- Extract F0 values of some number of equidistant points from a word.
- Identify the location of pauses that precede a word.
- General .textgrid file management, including the renaming of tiers, duplication of tiers, and adding labels to tiers, based on particular requirements.

The output of these scripts will be used to analyze the prosody of types of non-canonical questions.

Are super-complex words represented like sentences in speakers' minds? May 2021 - September 2021

Project leads:

University of Melbourne

Associate Professor Brett Baker (University of Melbourne)

Dr. Rikke Bundgaard-Nielsen (Western Sydney University)

This ARC funded project aims to examine speakers' knowledge of 'super-complex' words in the Australian language Wubuy. The results will experimentally investigate what a 'word' is in Wubuy, and thereby inform not only language description/typology but also likely challenge current models of speech processing and language acquisition.

I was responsible for writing Python scripts for the following tasks:

- Cleaning a corpus of Wubuy in XML into easy-to-read HTML.
- Automate interlinear glossing.
- Implementing phonological rules to generate all phonological variants of dictionary headwords.

The output of these scripts will be used to calculate the information density of words, as a means to model the complexity of a 'word' in Wubuy.

Learning to tell a narrative in Murrinhpatha

May 2021 - July 2021

Project leads:

University of Melbourne

Associate Professor Barbara Kelly, Professor Gillian Wigglesworth,

& Dr. Lucy Davidson (University of Melbourne)

This ARC funded project aims to highlight the linguistic skills that these children demonstrate when they engage in a task in their native language, the polysynthetic language Murrinhpatha.

I used ELAN for the following tasks:

- To code information structure in recorded videos of children's narratives spoken in Murrinhpatha.
- To develop a controlled vocabulary to focus on relevant areas in information structure.

Nafsan Corpus

November 2019 - January 2020

Project lead:

Melbourne, Australia

Dr. Rosey Billington (University of Melbourne)

I contributed to a corpus of the Oceanic language Nafsan by manually fixing annotation boundaries of phonetic transcriptions in Praat.

Summer Scholar

November 2018 - January 2019

Supervisors:

Australian National University

Professor Jane Simpson

& Associate Professor Hanna Suominen (Australian National University)

Funded by the Australian Signals Directorate, I was a summer scholar for two projects, using computational methods for linguistic outcomes.

Project 1: Warlpiri Corpus

I developed a workflow using forced alignment techniques, to time align large amounts of data for the central Australian language Warlpiri. I tested different forced alignments for non-technically proficient linguists, prepared guidelines to document the workflow for future use.

Project 2: Dhurga Dictionary

I worked closely with a community leader to develop a revival dictionary for the southern east coast Australian language Dhurga. I used HTML and CSS to create the web dictionary. On this project, I also contributed to the development of WordSpinner, a tool to create online dictionaries for under-resourced languages.

PUBLICATIONS

Language Documentation & Conservation

2023

Accepted

WordSpinner: Developing a tool to convert plain-text lexicon files into dictionary webpages

Authors: Caroline Jones, Jesse Tran, Eleanor Jorgensen, **Romi Hill**, Patricia Ellis, Jane Simpson, Felicity Meakins, Ben Foley, Marcel Reverter-Rambaldi, Lowana Tudor-Smith, Paul Williams, Clair Hill, Mark Richards, Quang Vinh Nguyen

Nuanced Garbling

2019

Melbourne University Linguistics Club Journal

pp. 119 - 133

An Analysis of Motion Verbs in Hindi

[\[Link\]](#)

Authors: **Romi Hill**, Tiger Liu and My Ngoc Tran

CONFERENCE PROCEEDINGS

Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)

2023

R. Skarnitzl and J. Volín (eds.).

pp. 1459 - 1464

The phonetics-phonology-syntax interface: a computational implementation

[\[Link\]](#)

Authors: Tina Bögel, **Romi Hill**, Justin Hofenbitzer, Tianyi Zhao

Also: Poster presentation in Prague, Czechia.

CONFERENCE TALKS

Association of Contemporary African Linguistics 56 (ACAL)

May 2025

Accepted as Poster Presentation

Morphosyntactic Constraints on the Morphophonology of Reduplication

Authors: **Romi Hill**

ACADEMIC PRESENTATIONS

University of Edinburgh

September 2022

Invited Talk

Edinburgh, Scotland

The prosody-syntax interface: theoretical assumptions and computational implementation

Presenters: Tina Bögel, **Romi Hill**, Justin Hofenbitzer, Tianyi Zhao

Australian Languages Workshop

February 2020

Presentation

Queensland, Australia

Murrinhpatha noun classifiers: Syntax and discourse functions

Presenters: **Romi Hill**, Rachel Nordlinger, John Mansfield

Australian National University

January 2019

Presentation

Canberra, Australia

Two projects: Computational methods for corpus and dictionary-making: Warlpiri corpus and Dhurga dictionary

Presenters: **Romi Hill**, Eleanor Jorgensen

TEACHING ASSISTANT EMPLOYMENT

University of Konstanz

October 2023 - February 2024

Core Components of Language B

University of Konstanz

Graduate level subject covering the foundations of Syntax, Semantics and Pragmatics.

I prepared weekly exercises and solutions, and marked assignments.

Instructor: Junior Professor Colin Davis.

University of Konstanz

October 2023 - February 2024

Statistics

University of Konstanz

Graduate level subject for linguistics students covering linear and logistic regression.

I prepared weekly exercises and solutions.

Instructor: Dr. Massimiliano Canzi.

University of Konstanz

September 2023

Finite State Morphology

University of Konstanz

Undergraduate level subject covering finite state morphology with XFST.

I assisted students with theoretical and technical questions in class.

Instructor: Dr. Tina Bögel.

University of Konstanz

April 2023 - July 2023

English Syntax

University of Konstanz

Undergraduate level subject covering an introduction to syntactic theory, with a focus on English

I prepared weekly exercises and solutions, and marked assignments.

Instructor: Junior Professor Colin Davis.

University of Konstanz
Core Components of Language B

October 2022 - February 2023
University of Konstanz

Graduate level subject covering the foundations of Syntax, Semantics and Pragmatics.
I prepared weekly exercises and solutions, and marked assignments.
Instructor: Junior Professor Colin Davis.

University of Melbourne
PASS Leader

February 2019 - November 2019
University of Melbourne

I facilitated "Peer Assisted Study Sessions": weekly study sessions led by a high-achieving student who has successfully completed the subject in the previous year.
I was a PASS leader for two third-year linguistics subjects:

Semantics

Instructor: Lecturer Jonathon Lum

Exploring Linguistic Diversity

Instructor: Associate Professor Barbara Kelly

I prepared questions and discussion topics, and mentored students around university life.

INDUSTRY EMPLOYMENT

Appen
Computational Linguist

January 2020 - July 2021
Sydney, Australia

I primarily developed training data for natural language processing tasks. I employed my linguistic, technical, and client-management skills to prepare high quality training data for clients. I worked on a number of NLP projects, including pronunciation lexicons for Text To Speech & Automatic Speech Recognition systems, Named Entity Recognition annotation, text generation, academic support and mentoring for Python scripting, and tool development & process improvements. These projects all require proficient use and development of regex & Python scripts to target client- and language- specific requirements.

Real Thing AI
Conversational Designer

July 2021 - October 2021
Melbourne, Australia

I was the subject matter expert in linguistics, employed to improve the dialogue design for the technology underlying RealSAM, a voice command system designed for Blind and Visually Impaired users. I reported directly to the CEO, and identified key issues and proposed creative solutions to a team of engineers in order to improve the system's natural language understanding.

STUDENT CLUBS

Melbourne University Linguistics Club
Education Officer, General Committee Member

2015-2018
University of Melbourne

Apollo Music Society Choir
Alto Singer, General Committee Member

2015-2017
University of Melbourne

SUMMER SCHOOL PARTICIPATION

Basque Summer Tutorial in Language Documentation

August 2022
Larrau, France

CoEDL Summer School Workshop
Also volunteered as a helping assistant.

November 2019
Melbourne, Australia

CoEDL Summer School Workshop

November 2018
Canberra, Australia

TECHNICAL SKILLS

Computational Languages	Python, C, SQL, HTML & CSS, JavaScript, R, Praat scripting, XLE, XFST
Command Line Interfaces	bash/zsh, sed, git
Text Editing	LaTeX, vim, regex, VSCode
Linguistics Software	ELAN, Praat, FLEx, WebMAUS
Human Languages	Native Australian English, B2 German, Beginner's Mandarin Chinese

AWARDS AND RECOGNITION

Begabtenbefreiungssatzung	2022 - 2024
<i>Exemption to Pay International Fees</i>	<i>University of Konstanz</i>

Identified as a particularly gifted international student, and received an exemption to pay international fees.

DAAD Scholarship (Master Studies for All Academic Disciplines)	2022 - 2023
<i>Scholarship</i>	<i>University of Konstanz</i>

Monthly scholarship payments for excellently-qualified graduates. The scholarship is awarded by an independent selection committee, and selection is based on academic achievements, relevant internships and work experience, quality of the study project, and potential of the applicant.

Value: Monthly payments of €934.

R.G. Wilson Scholarship	2019
<i>Scholarship</i>	<i>University of Melbourne</i>

Three scholarships awarded to the highest average honours students in the Faculty of Arts.

Value: One-off payment of AU\$3,200.

Dwight Final Assessment Prize	2019
<i>Prize</i>	<i>University of Melbourne</i>

Awarded to the highest achieving honours student in Linguistics.

Value: One-off payment of AU\$1,600.

Dean's Honours List	2019
<i>Recognition</i>	<i>University of Melbourne</i>

Awarded to the top 3% of students within each year of the program in the Bachelor of Arts.

Appen Prize for Excellence in Linguistics	2018
<i>Prize</i>	<i>University of Melbourne</i>

Awarded to the highest achieving Linguistics student across all their Linguistics subjects, who goes on to enrol in honours.

Value: One-off payment of AU\$1,000.

Melbourne Global Scholars Award	2017
<i>Scholarship</i>	<i>University of Melbourne</i>

Awarded to high achieving students who are undertaking overseas study for 12 weeks or more.

Value: One-off payment of AU\$2,500.

REFERENCES

Colin Davis

Associate Professor of English Language and Linguistics

Nord University

colin.davis@nord.no

Dr. Tina Bögel

Postdoctoral Researcher in Linguistics

University of Konstanz

tina.boegel@uni-konstanz.de

Miriam Butt

Professor in Linguistics

University of Konstanz

miriam.butt@uni-konstanz.de