

Chihiro Taguchi

Updated June 21, 2022

Email: taguchi.chihiro.td0@is.naist.jp / c.taguchi@sms.ed.ac.uk / ctaguchi@end.edu

Phone: +44 7597 603895

Website: <https://chipewyan.github.io>

Research interests

Natural Language Processing (low-resource languages)
Linguistics (morphosyntax)

Education

**College of Engineering,
The University of Notre Dame** IN, United States
PhD in Computer Science and Engineering. Expected Aug 2022 – Jul 2027
Supervisor: Dr. David Chiang.

**School of Philosophy, Psychology and Language Sciences,
The University of Edinburgh** Edinburgh, United Kingdom
MSc by Research (MScR) in Linguistics. Sep 2021 – Present
Supervisors: Prof. Peter Ackema and Dr. Itamar Kastner.

**Division of Information Science,
Nara Institute of Science and Technology** Nara, Japan
MA in Engineering. Oct 2020 – Present
Supervisor: Prof. Taro Watanabe (Natural Language Processing Laboratory).
GPA: 3.58 (at the time of writing).

**School of Oriental and African Studies,
The University of London** London, United Kingdom
Exchange Student in Linguistics. Sep 2018 – Jun 2019
Mentor: Dr. Christopher Lucas.
Thesis: On Three Types of Predicates with Infinitival VP in Tatar:
kiräk, tiyeš, and *telä-* with LFG Analysis. (**Distinction**)
GPA: 3.75.

**Department of Political Science, Faculty of Law,
Keio University** Tokyo, Japan
Bachelor in Law. Apr 2015 – Sep 2019
Supervisor: Prof. Yoshihisa Hagiwara.
Thesis: Observations of the Process of Language Extinction in the Modern Society:
A Case Study on the Ikema Dialect of Miyako Ryukyuan. (in Japanese)
GPA: 3.77.

Training

Summer Intensive Language Program (Jinghpaw). Aug – Sep 2019

75-hour intensive training course of Jinghpaw (Tibeto-Burman, Sino-Tibetan).
Research Institute for Languages and Cultures of Asia and Africa (ILCAA),
Tokyo University of Foreign Studies (TUFS), Japan.

Lithuanian Language Summer Course. Jul 2019

Two-week summer school for Lithuanian (Balto-Slavic, Indo-European).
Vilnius University, Lithuania.

Summer Intensive Language Program (Mee (Ekagi)). Aug – Sep 2018

50-hour fieldwork training course for the documentation of Mee (Papuan).
ILCAA, TUFS, Japan.

Summer School in Languages and Linguistics. Jul 2018

Two-week summer courses on languages and linguistics.
Leiden University, Netherlands.

3rd Miyako Language Documentation Training Workshop. Feb 2018

One-week fieldwork for the documenting a dialect of Miyako (Ryukyuan, Japonic).
Ikemajima Island, Okinawa, Japan.

Summer School in Languages and Linguistics. Jul 2017

Two-week summer courses on languages and linguistics.
Leiden University, Netherlands.

Armenian Language Summer School. Jul – Aug 2016

One-month summer school on Eastern Armenian (Indo-European).
ARMACAD, Armenia.

Summer Intensive Language Program (Palestinian Arabic). Aug – Sep 2015

125-hour intensive training course of Palestinian Arabic (Semitic, Afroasiatic).
ILCAA, TUFS, Japan.

Summer Intensive Language Program (Chaga-Rombo). Aug – Sep 2014

132-hour fieldwork-style training on documenting Chaga-Rombo (Bantu).
ILCAA, TUFS, Japan.

Summer Intensive Language Program (Eastern Armenian). Aug – Sep 2013

100-hour intensive training course of Eastern Armenian (Indo-European).
ILCAA, TUFS, Japan.

Experience

Research Intern.

Apr 2021 – Present

Pluszero Co.

Project on building a semantic model of Japanese for artificial intelligence.

Research Assistant

Apr 2021 – Sep 2021

National Institute of Informatics, Japan.

Mentor: Dr. Saku Sugawara.

Research on the interface of natural language understanding and linguistics.

Cookpad Online Spring Internship 2021

Mar 2021

Cookpad Inc.

One-week program on front-end engineering.

Publications and Presentations

Consistent Grammatical Annotation of Turkic Languages for more universal Universal Dependencies. (*peer-reviewed*)

Chihiro Taguchi.

Workshop on Computational Linguistics on East Asian languages, 29th International Conference on Head-Driven Phrase Structure Grammar, 2022 (*to appear*).

Mermaid Construction in Lexical Functional Grammar. (*peer-reviewed*)

Chihiro Taguchi.

LFG22: the 27th International Lexical Functional Grammar Conference, 2022 (*to appear*).

名詞述語を伴う繰り上げ構文としての人魚構文 [Mermaid Construction as raising with a nominal predicate]. (in Japanese, *peer-reviewed*)

Chihiro Taguchi.

Proceedings of the 164th Conference of the Linguistic Society of Japan, 2022 (*to appear*).

Universal Dependencies Treebank for Tatar: Incorporating Intra-Word Code-Switching Information. (*peer-reviewed*)

Chihiro Taguchi, Sei Iwata, Taro Watanabe.

Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-resourced Languages in Eurasia (EURALI), Language Resources and Evaluation Conference, 2022 (*to appear*).

Mermaid Construction in Tatar. (*peer-reviewed*)

Chihiro Taguchi.

Proceedings of the 162nd Conference of the Linguistic Society of Japan, 2021.

Transliteration for Low-Resource Code-Switching Texts: Building an Automatic Cyrillic-to-Latin Converter for Tatar. (*peer-reviewed*)

Chihiro Taguchi, Yusuke Sakai, Taro Watanabe.

Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching, North American Chapter of the Association for Computational Linguistics, 2021.

So-Called “Prepositions” in Somali are Not Prepositions: A Linguistic Approach for Somali POS Tagging.

Chihiro Taguchi, Taro Watanabe.

Proceedings of the Annual Meeting of the Association of Natural Language Processing, 2021.

Raising to Quirky Subject in Tatar. (*peer-reviewed*)

Chihiro Taguchi.

Proceedings of the 161st Conference of the Linguistic Society of Japan, 2020.

Vowel Harmonies in Kazan Tatar: Element Theoretical Approach. (*peer-reviewed*)

Chihiro Taguchi.

Proceedings of the 159th Conference of the Linguistic Society of Japan, 2019.

近現代における言語衰退過程の観察と考察 —琉球語宮古池間方言を事例に—
[Observations of the Process of Language Extinction in the Modern Society: A Case Study of the Ikema Dialect of Miyako Ryukyuan] (in Japanese).

Chihiro Taguchi.

政治学研究 [*Studies in Political Science*], 2018. Keio University Press.

Projects

Universal Dependencies: Tatar NMCTT treebank.

Repository: https://github.com/UniversalDependencies/UD_Tatar-NMCTT

Talks

Mermaid Construction as raising with a nominal predicate.

Chihiro Taguchi.

The 74th Language Lunch At Edinburgh. Online. 8 Mar 2022.

Text Processing of Open-Access Jinghpaw Data on Google Colaboratory.

Chihiro Taguchi.

Follow-up meeting on Intensive Language Course 2019 (Jinghpaw) / The 3rd meeting on Kachin studies.

ILCAA, TUFS. Online. 16 Oct 2021.

Grants and Scholarships

JASSO Scholarship for Graduate-level Study Abroad. Sep 2021 – Aug 2022

Full tuition fee coverage and monthly stipends.

Ministry of Education, Culture, Sports, Science and Technology, Japan.

CICP NAIST Multilingual Corpus. (500,000 JPY) Jun 2021 – Mar 2022
Nara Institute of Science and Technology.

Grand Prix Winner April 2019
7th International Olympiad of the Tatar Language and Literature.
Ministry of Education and Science, Republic of Tatarstan, Russia.

Exchange Program Scholarship (1.5 mil JPY). Sep 2018 – Jun 2019
Gyomu Super Japan Dream Foundation.

Honorable Mention, Keio University Academic Writing Contest. Nov 2014
Keio University.

Finalist, Teen Polyglot Challenge. May 2014
Link: <https://www.youtube.com/watch?v=VndRKnEXiAQ>

Skills

Programming

Familiar: Python3.

Languages

Native: Japanese

Fluent: English

Advanced: Russian, Tatar

Intermediate: French, German, Korean, Mandarin Chinese