

Chihiro Taguchi

Updated February 4, 2024

田口 智大

✉ Email: ctaguchi@nd.edu

🐙 GitHub: [ctaguchi](https://github.com/ctaguchi)

💻 Website: <https://ctaguchi.github.io>

Education

**College of Engineering,
The University of Notre Dame** IN, United States
PhD in Computer Science and Engineering. Aug 2022 – Jul 2027

**School of Philosophy, Psychology and Language Sciences,
The University of Edinburgh** Edinburgh, United Kingdom
MSc by Research in Linguistics with Distinction. Sep 2021 – Aug 2022

**Division of Information Science,
Nara Institute of Science and Technology** Nara, Japan
MA in Engineering. Oct 2020 – Sep 2022

**School of Oriental and African Studies,
The University of London** London, United Kingdom
Exchange Student in Linguistics. Sep 2018 – Jun 2019

**Department of Political Science, Faculty of Law,
Keio University** Tokyo, Japan
Bachelor of Laws (LL.B.). Apr 2015 – Sep 2019

Training

Linguistic Society of America (LSA) Institute. Jun – Jul 2023
University of Massachusetts, Amherst, MA, United States.

Summer Language Abroad Program (Quechua), University of Notre Dame. May – Jun 2023
Quito, Ecuador.

Summer School in Languages and Linguistics. Jul 2022
Leiden University, Netherlands.

Advanced Core Training in Linguistics (ACTL) Summer School. Jun 2022
University of York, United Kingdom.

Summer Intensive Language Program (Jinghpaw). Aug – Sep 2019
Tokyo University of Foreign Studies (TUFS), Japan.

Lithuanian Language Summer Course. Jul 2019
Vilnius University, Lithuania.

Summer Intensive Language Program (Mee/Ekagi). Aug – Sep 2018
Tokyo University of Foreign Studies (TUFS), Japan.

Summer School in Languages and Linguistics. Jul 2018
Leiden University, Netherlands.

3rd Miyako Language Documentation Training Workshop. Feb 2018
Ikemajima Island, Okinawa, Japan.

Summer School in Languages and Linguistics. Jul 2017
Leiden University, Netherlands.

Armenian Language Summer School. Jul – Aug 2016
ARMACAD, Armenia.

Summer Intensive Language Program (Palestinian Arabic). Aug – Sep 2015
Tokyo University of Foreign Studies (TUFS), Japan.

Summer Intensive Language Program (Chaga-Rombo). Aug – Sep 2014
Tokyo University of Foreign Studies (TUFS), Japan.

Summer Intensive Language Program (Eastern Armenian). Aug – Sep 2013
Tokyo University of Foreign Studies (TUFS), Japan.

Experience **Student Researcher.** Aug 2023 – Dec 2023
AnalytiXIN.

Research Intern. Apr 2021
Pluszero Co.

Research Assistant. Apr 2021 – Sep 2021
National Institute of Informatics, Japan.

Cookpad Online Spring Internship 2021. Mar 2021
Cookpad Inc.

Publications **Chihiro Taguchi**, Yusuke Sakai, Parisa Haghani, David Chiang. Universal Automatic Phonetic Transcription into the International Phonetic Alphabet. *INTERSPEECH 2023*, 2023. ArXiv: <https://arxiv.org/abs/2308.03917>

Chihiro Taguchi, So Miyagawa. 形態論情報付き日本語 Universal Dependencies [Japanese Universal Dependencies with morphological information]. 第 29 回言語処理学会年次大会 [*The 29th Annual Meeting of the Association of Natural Language Processing (NLP2023)*], 2023.

So Miyagawa, Hiroshi Kanayama, **Chihiro Taguchi**, Nana Tohyama. 沖縄語の Universal Dependencies ツリーバンクコーパスの構築 [Building a treebank corpus for Okinawan Universal Dependencies]. 第 29 回言語処理学会年次大会 [*The 29th Annual Meeting of the Association of Natural Language Processing (NLP2023)*], 2023.

Chihiro Taguchi, David Chiang. Introducing Morphology in Universal Dependencies Japanese. *The Sixth Workshop on Universal Dependencies (UDW)*, 2023. Link: <https://aclanthology.org/2023.udw-1.8>

Chihiro Taguchi. Mermaid Construction in Lexical Functional Grammar. *Proceedings of the LFG'22 Conference*, 2022. Link: <https://lfg-proceedings.org/lfg/index.php/main/article/view/19>

Chihiro Taguchi. Consistent Grammatical Annotation of Turkic Languages for more universal Universal Dependencies. *Workshop on Computational Linguistics on East Asian languages*, 29th International Conference on Head-Driven Phrase Structure Grammar, 2022.

Chihiro Taguchi. 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) *Proceedings of the 164th Conference of the Linguistic Society of Japan*, 2022.

Chihiro Taguchi, Sei Iwata, Taro Watanabe. Universal Dependencies Treebank for Tatar: Incorporating Intra-Word Code-Switching Information. *Workshop on Resources and Technologies for Indigenous, Endangered and Lesser-resourced Languages in Eurasia (EURALI)*, Language Resources and Evaluation Conference (LREC), 2022. Link: <https://aclanthology.org/2022.euralli-1.17>

Chihiro Taguchi. Mermaid Construction in Tatar. *Proceedings of the 162nd Conference of the Linguistic Society of Japan*, 2021.

Chihiro Taguchi, Yusuke Sakai, Taro Watanabe. Transliteration for Low-Resource Code-Switching Texts: Building an Automatic Cyrillic-to-Latin Converter for Tatar. *Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching*, North American Chapter of the Association for Computational Linguistics, 2021. Link: <https://aclanthology.org/2021.calcs-1.18>

Chihiro Taguchi, Taro Watanabe. So-Called “Prepositions” in Somali are Not Prepositions: A Linguistic Approach for Somali POS Tagging. *Proceedings of the Annual Meeting of the Association of Natural Language Processing*, 2021.

Chihiro Taguchi. Raising to Quirky Subject in Tatar. *Proceedings of the 161st Conference of the Linguistic Society of Japan*, 2020.

Chihiro Taguchi. Vowel Harmonies in Kazan Tatar: Element Theoretical Approach. *Proceedings of the 159th Conference of the Linguistic Society of Japan*, 2019.

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

Projects

A study on the grammar of discourse in Eurasian languages. Nov 2023 – Present
Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

UD Turkic. Sep 2023 – Present
Repository: <https://github.com/ud-turkic>

Language Documentation with an AI Helper. Sep 2022 – Present
NSF Grant project in collaboration with Antonios Anastasopoulos, David Chiang, and Géraldine Walther.

Building a multilingual speech-to-IPA system. Jan 2023

SLT-Code Hackathon, IEEE.

Universal Dependencies: Tatar NMCTT treebank.

Aug 2021 – Jun 2022

Repository: <https://github.com/UniversalDependencies/UD-Tatar-NMCTT>

Talks

Chihiro Taguchi. El conjunto de datos de reconocimiento automático de voz para kichwa. El Taller: Tecnologías Digitales y Lenguas Indígenas. Universidad de Chile, Santiago de Chile, Chile, January 10, 2024.

Chihiro Taguchi, Jefferson Saransig. Non-Discourse-Configurationality in Imbabura Kichwa. *Linguistic Society of America Annual Meeting*, 2024.

Chihiro Taguchi, Jefferson Saransig, Dayana Velásquez. Reconocimiento automático del habla de kichwa para su inclusión y empoderamiento en el ciberespacio. VI Seminario Internacional Revitalizando Ando. Universidad Politécnica Salesiana, Quito, Ecuador, November 21, 2023.

Chihiro Taguchi. UD-Tatar NMCTT Treebank: Issues in annotation across Turkic UD. UD Turkic Workshop 2023. Istanbul Technical University, September 8, 2023.

Chihiro Taguchi. Grammaticalization of modal nominal predicates in Tatar. *International Conference on Turkish Linguistics*, 2023.

Chihiro Taguchi. Incorporating AI-based speech transcription into language documentation: A case study of Imbabura Kichwa. *LSA Institute*, 2023.

Chihiro Taguchi. Bridging Natural Language Processing and Descriptive Linguistics with Universal Dependencies. Field Linguistics Workshop: Grammatical Studies Workshops 23 “Grammatical Studies and Digital Humanities (2)”, Tokyo University of Foreign Studies, July 26, 2023.

Chihiro Taguchi. Japanese gapless relativization: The syntax–prosody interface to semantics. *LFG23: the 28th International Lexical Functional Grammar Conference*, 2023.

Chihiro Taguchi. Cross-linguistic analysis of the Mermaid Constructions in LFG. 33rd South of England LFG Meetings, October 29, 2022.

Chihiro Taguchi. Universal Dependencies Tatar for Code-Switching Detection. Notre Dame NL+, University of Notre Dame, September 19, 2022.

Chihiro Taguchi. Mermaid Construction in Lexical Functional Grammar. *LFG22: the 27th International Lexical Functional Grammar Conference*, 2022.

Chihiro Taguchi. Mermaid Construction as raising with a nominal predicate. The 74th Language Lunch At Edinburgh, March 8, 2022.

Chihiro Taguchi. Text Processing of Open-Access Jinghpaw Data on Google Colaboratory. Follow-up meeting on Intensive Language Course 2019 (Jinghpaw) / The 3rd meeting on Kachin studies. ILCAA, Tokyo University of Foreign Studies, October 16, 2021.

Awards

Conference Travel Grant. Kellogg Institute for International Studies, University of Notre Dame.	Jun 2023
Professionalization Grant. Kellogg Institute for International Studies, University of Notre Dame.	May 2023
Committee's Honorable Mention (委員特別賞). The 29th Annual Meeting of the Association for Natural Language Processing.	Mar 2023
Conference Presentation Grant. Graduate Student Government, University of Notre Dame.	Mar 2023
Summer Language Abroad Grant. The Center for the Study of Languages and Cultures, University of Notre Dame.	Mar 2023
JASSO Scholarship for Graduate-level Study Abroad. Ministry of Education, Culture, Sports, Science and Technology, Japan.	Sep 2021 – Aug 2022
CICP NAIST Multilingual Corpus. Nara Institute of Science and Technology.	Jun 2021 – Mar 2022
Grand Prix Winner, 7th International Olympiad of the Tatar Language and Literature. Ministry of Education and Science, Republic of Tatarstan, Russia.	Apr 2019
Exchange Program Scholarship. Gyomu Super Japan Dream Foundation.	Sep 2018 – Jun 2019
Honorable Mention, Keio University Academic Writing Contest. Keio University.	Nov 2014

Skills

Programming: Python, R, C, \LaTeX , HTML/CSS.
Softwares and libraries: PyTorch, Tensorflow, Wandb, Git, AWS, Docker, ELAN, Praat.
Languages:
Native: Japanese
Fluent: English
Advanced: Russian, Tatar, Spanish
Intermediate: French, German, Korean, Mandarin Chinese, Kichwa

Affiliations

Kellogg Institute Doctoral Student Affiliate
Linguistic Society of America
言語処理学会 [Association for Natural Language Processing]
日本言語学会 [Linguistic Society of Japan]