DOCTORAL THESIS WRITING QUALIFICATION REVIEW - CONTACT INFORMATION DOCUMENT

2025-08-12

Doctoral Thesis Writing Qualification Review

Contact Information Document for Review Committee

1 Applicant Information

Name: Sora Nagano

Student ID: [Student ID Number]

Enrollment Year: 2023 (Reiwa 5)

Affiliation: Graduate School of Humanities and Sociology, Department of Language and Information Sciences, The University of Tokyo

2 Contact Information

Address: 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan Graduate School of Humanities and Sociology, Department of Language and Information Sciences, The University of Tokyo

Phone Number: +81-3-5841-3839 (Department Office)

Email Address: s-oswld-n@g.ecc.u-tokyo.ac.jp

Emergency Contact: [Emergency Phone Number]

Alternative Email: [Alternative Email Address]

3 Thesis Title

English Title: Phonological Features in the Age of Deep Learning: A Multi-dimensional Exploration of Optimal Representational Units for Language Modeling

Japanese Translation:

4 Thesis Abstract (200 words)

This research explores the fundamental question of "optimal representational units" in computational phonology. The study aims to integrate symbolic approaches (phonemes, distinctive features) with modern deep learning approaches (continuous representations from wav2vec 2.0 and related models) by proposing a neuro-symbolic hybrid model utilizing vector quantization technology. The research evaluates different representational paradigms across four dimensions: predictive accuracy, linguistic interpretability, cognitive plausibility, and computational efficiency. Through

systematic comparison of continuous vectors, VQ codes, phonemes, and features across diverse phonological tasks, the study develops novel hybrid architectures that parameterize Maximum Entropy Harmonic Grammar using self-supervised learning representations. The cognitive plausibility is validated through developmental simulations using CHILDES corpus and ABX discrimination tasks. This work contributes to both theoretical linguistics and practical AI applications by establishing a new phonological representation theory that bridges symbolic precision with neural learning capabilities, advancing the integration of linguistic knowledge with data-driven machine learning for next-generation speech technologies.

5 Academic Supervisors

Primary Supervisor: [Supervisor Name] Professor, Graduate School of Humanities and Sociology, Department of Language and Information Sciences, The University of Tokyo

Secondary Supervisor: [Co-supervisor Name] Associate Professor, Graduate School of Humanities and Sociology, Department of Language and Information Sciences, The University of Tokyo

6 Research Overview

This doctoral thesis represents computational linguistics research that multi-dimensionally explores the optimality of representational units in phonology against the backdrop of rapid deep learning technology development. The study aims to bridge traditional symbolic approaches with modern neural network approaches through three major research questions:

1. **Empirical Landscape Elucidation (RQ1)**: Systematic comparative evaluation of different representational units
2. **Neuro-symbolic Integration (RQ2)**: Design and validation of hybrid architectures
3. **Cognitive Plausibility Validation (RQ3)**: Consistency evaluation with human language acquisition patterns

The research originality lies in proposing novel modeling methodologies that integrate symbolic precision with neural learning capabilities, aiming for interdisciplinary contributions through the fusion of linguistic theory and AI technology.

7 Preferred Review Schedule

Preferred Review Period: March 2024

Expected Thesis Submission: January 2028

Expected Defense: March 2028

8 Research Ethics and Data Management

This research will be conducted in accordance with the academic ethics guidelines of the Graduate School of Humanities and Sociology, Department of Language and Information Sciences, The University of Tokyo. Applications to the Ethics Review Committee will be submitted as necessary. All datasets used in the research are publicly available and will be utilized under appropriate licensing agreements.

The research maintains international collaboration potential, including partnerships with overseas research institutions, and aims for active international dissemination of research outcomes through publications in top-tier conferences and journals.

9 Language of Dissertation

Primary Language: English

Target Length: Approximately 200 pages

Expected Chapters: 7 chapters covering introduction, literature review, methodology, three empirical studies (RQ1-RQ3), and general discussion

10 Additional Information

The research involves development of novel computational models and requires access to high-performance computing resources. Collaboration agreements with international research institutions may be established to enhance the research scope and impact. All experimental procedures will follow established protocols for computational linguistics research and data usage guidelines.

Application Date: [Month] [Day], 2024

Applicant Signature: Sora Nagano

Supervisor Signature: [Supervisor Name]

Department Approval: [Department Stamp/Signature]