# Doctoral Thesis Writing Plan

Computational Linguistic Study on Japanese Accent Estimation

2025-08-19

# 1 Doctoral Thesis Writing Plan

## 1.1 Writing Schedule (36-month Plan)

### 1.1.1 Year 1 (Months 1-12): Foundation Phase

**Objective**: Establish theoretical foundation and baseline implementations

#### 1.1.1.1 Months 1-3: Literature Review and Theoretical Framework

- **Computational Phonology Survey**: Comprehensive review of symbolic approaches (OT, HG, SPE)
- **Neural Representation Analysis**: Review of self-supervised learning in speech (wav2vec 2.0, HuBERT, etc.)
- **Vector Quantization Literature**: Analysis of VQ-VAE and related discrete representation methods
- **Deliverable**: Complete literature review chapter (40-50 pages)

#### 1.1.1.2 Months 4-6: Methodology Development and Baseline Implementation

- **Evaluation Framework Design**: Multi-dimensional assessment protocols
- **Baseline Model Implementation**: Pure symbolic and pure neural baselines
- **Dataset Preparation**: Cross-linguistic phonological task datasets
- **Deliverable**: Methodology chapter draft (30-40 pages)

#### 1.1.1.3 Months 7-9: Initial Empirical Studies (RQ1)

- **Systematic Comparison**: Evaluation across representational units
- **Cross-linguistic Testing**: Initial validation across language families
- **Performance Analysis**: Comparative assessment of different approaches
- **Deliverable**: First empirical study chapter (35-45 pages)

#### 1.1.1.4 Months 10-12: Hybrid Architecture Design Foundation (RQ2)

- **Architecture Conceptualization**: Design principles for neuro-symbolic integration
- **Preliminary Implementation**: Basic hybrid model prototypes
- **Initial Testing**: Proof-of-concept evaluations
- **Deliverable**: Technical report and second empirical study foundation

### 1.1.2 Year 2 (Months 13-24): Core Development Phase

**Objective**: Develop and validate core hybrid architectures

#### 1.1.2.1 Months 13-15: Advanced Hybrid Architecture Implementation

- **MaxEnt-HG Integration**: Implementation of neural constraint parameterization
- **VQ-Code Integration**: Development of discrete neural representation methods
- **Architecture Optimization**: Performance tuning and optimization
- **Deliverable**: Complete hybrid model implementations

#### 1.1.2.2 Months 16-18: Comprehensive Evaluation (RQ2 Continued)

- **Task Performance Evaluation**: Systematic testing across phonological tasks
- **Interpretability Analysis**: Assessment of symbolic knowledge preservation
- **Cross-linguistic Validation**: Extended testing across diverse languages
- **Deliverable**: Second empirical study chapter (40-50 pages)

#### 1.1.2.3 Months 19-21: Cognitive Plausibility Studies (RQ3)

- **CHILDES Corpus Analysis**: Developmental trajectory simulation
- **ABX Discrimination Tasks**: Human-AI comparison studies
- **Acquisition Pattern Validation**: Consistency with developmental linguistics
- **Deliverable**: Third empirical study chapter (35-45 pages)

#### 1.1.2.4 Months 22-24: Integration and Optimization

- **Performance Optimization**: Model refinement and enhancement
- **Scalability Analysis**: Computational efficiency improvements
- **Reproducibility Preparation**: Code documentation and sharing preparation
- **Deliverable**: Technical documentation and optimization report

### 1.1.3 Year 3 (Months 25-36): Consolidation and Completion Phase

**Objective**: Finalize research, complete dissertation, and prepare defense

#### 1.1.3.1 Months 25-27: Comprehensive Analysis and Additional Studies

- **Gap Analysis**: Identification and addressing of remaining research gaps
- **Additional Experiments**: Supplementary studies as needed
- **Statistical Analysis**: Comprehensive statistical validation
- **Deliverable**: Complete empirical results and analysis

#### 1.1.3.2 Months 28-30: Dissertation Writing - Core Chapters

- **Introduction Chapter**: Problem statement, motivation, and contributions (20-25 pages)
- **Related Work Chapter**: Comprehensive literature review (40-50 pages)
- **General Discussion**: Cross-study analysis and theoretical implications (25-30 pages)
- **Deliverable**: Complete first draft of core chapters

#### 1.1.3.3 Months 31-33: Dissertation Writing - Final Chapters

- **Conclusion Chapter**: Summary, contributions, and future work (15-20 pages)
- **Abstract and Executive Summary**: Concise research overview (3-5 pages)
- **Bibliography and Appendices**: Complete reference compilation and supplementary materials
- **Deliverable**: Complete dissertation first draft (200+ pages)

#### 1.1.3.4 Months 34-36: Review, Revision, and Defense Preparation

- **Supervisor Review**: Comprehensive feedback incorporation
- **External Review**: Committee member feedback integration
- **Defense Preparation**: Presentation development and practice sessions
- **Final Submission**: Dissertation submission and defense scheduling
- **Deliverable**: Final dissertation and successful defense

## 1.2 Chapter Structure and Target Length

### 1.2.1 Chapter 1: Introduction (20-25 pages)

- Problem statement and motivation
- Research questions and hypotheses
- Contributions and significance
- Thesis organization

#### 1.2.2 Chapter 2: Related Work (40-50 pages)

- Computational phonology foundations
- Neural representation learning
- Neuro-symbolic integration approaches
- Evaluation methodologies

#### 1.2.3 Chapter 3: Methodology (30-40 pages)

- Theoretical framework
- Hybrid architecture design
- Evaluation protocols
- Experimental setup

#### 1.2.4 Chapter 4: Empirical Landscape Analysis (35-45 pages) [RQ1]

- Systematic comparison methodology
- Cross-linguistic evaluation results
- Performance analysis and insights

#### 1.2.5 Chapter 5: Neuro-symbolic Integration (40-50 pages) [RQ2]

- Hybrid architecture implementation
- Integration strategies and results
- Interpretability analysis

#### 1.2.6 Chapter 6: Cognitive Plausibility Validation (35-45 pages) [RQ3]

- Developmental simulation methodology
- Human-AI comparison studies
- Acquisition pattern analysis

#### 1.2.7 Chapter 7: General Discussion (25-30 pages)

- Cross-study synthesis
- Theoretical implications
- Practical applications
- Limitations and future work

#### 1.2.8 Chapter 8: Conclusion (15-20 pages)

- Summary of contributions
- Significance and impact
- Future research directions

**Total Target Length**: 240-305 pages

### 1.3 Milestone Schedule and Deliverables

#### 1.3.1 Academic Year Milestones

##### 1.3.1.1 Year 1 Milestones

- **Month 3**: Literature review completion
- **Month 6**: Methodology framework establishment
- **Month 9**: First empirical study completion
- **Month 12**: Baseline architecture validation

##### 1.3.1.2 Year 2 Milestones

- **Month 15**: Core hybrid architecture implementation
- **Month 18**: Second empirical study completion
- **Month 21**: Third empirical study completion
- **Month 24**: Comprehensive evaluation completion

#### 1.3.1.3 Year 3 Milestones

- **Month 27**: All empirical work completion
- **Month 30**: Dissertation first draft completion
- **Month 33**: Final draft completion
- **Month 36**: Defense and submission

### 1.3.2 Publication Timeline

#### 1.3.2.1 Conference Publications

- **Month 12**: Initial results submission to computational linguistics conference
- **Month 18**: Hybrid architecture results submission to machine learning conference
- **Month 24**: Cognitive plausibility results submission to cognitive science conference

#### 1.3.2.2 Journal Publications

- **Month 30**: Comprehensive methodology paper submission
- **Month 33**: Theoretical implications paper submission

## 1.4 Resource Requirements and Collaboration

### 1.4.1 Computational Resources

- High-performance GPU access for neural model training
- Large-scale storage for multilingual datasets
- Cloud computing resources for distributed experiments

### 1.4.2 Data Resources

- Cross-linguistic phonological datasets
- CHILDES corpus access and processing
- Multilingual speech corpora for self-supervised learning

### 1.4.3 Collaboration Opportunities

- International research partnerships
- Industry collaboration for practical applications
- Interdisciplinary connections with cognitive science researchers

## 1.5 Risk Management and Contingency Planning

### 1.5.1 Technical Risks

- **Model Convergence Issues**: Alternative optimization strategies prepared
- **Computational Limitations**: Cloud resource scaling and optimization plans
- **Data Availability**: Multiple dataset alternatives identified

### 1.5.2 Timeline Risks

- **Experiment Delays**: Parallel development strategies implemented
- **Writing Bottlenecks**: Incremental writing approach with regular deadlines
- **Review Delays**: Early submission for supervisor feedback

### 1.5.3 Quality Assurance

- Regular progress reviews with supervisors
- Peer review through conference submissions
- Code review and reproducibility verification

This comprehensive writing plan ensures systematic progress toward a high-quality doctoral dissertation while maintaining flexibility for unexpected challenges and opportunities that may arise during the research process.