Thaislate: Context-Aware Thai-English Translation for Grammar Learning

Papatchol Thientong[1*]

[1]School of Computer Science, University of Sheffield, Sheffield, United Kingdom

Abstract

Thai learners struggle with English tenses because Thai lacks grammatical tense markers. Existing translation tools provide corrections without explanations, leaving learners unable to understand their errors. This paper presents Thaislate, a proof-of-concept system demonstrating how large language models can bridge this gap through context-aware grammar explanations. The system integrates three specialised models in a pipeline architecture: Typhoon Translate 4B for Thai-English translation, custom-trained XLM-RoBERTa achieving 94.7% accuracy on 24-category tense classification, and Typhoon 2.1 12B generating educational explanations. User testing with 38 Thai learners produced 474 ratings averaging 4.2/5, with explanation quality rated 4.33/5 despite 74% pipeline tense classification accuracy. The disconnect between technical performance and user satisfaction validates the approach: learners value clear explanations even when imperfect. The system successfully serves real users, establishing technical feasibility and user acceptance as foundation for future longitudinal studies on learning effectiveness.

Introduction

Thai learners struggle with English tense usage because Thai is a tenseless language that relies on contextual cues rather than grammatical markers to convey temporal relationships. When a Thai student writes "I am study at university for three years," they know something is wrong but cannot explain why "I have been studying" is correct. Current translation tools compound this problem by providing corrections without explanations, leaving learners to memorise patterns without understanding underlying temporal logic.

This research addresses a fundamental question: **Can AI-powered tools help Thai learners understand English tense usage through contextual, automatically-generated explanations?** The answer requires bridging the conceptual gap between Thai temporal expression and English grammatical tense through educational AI systems.

This paper presents Thaislate, a proof-of-concept system that transforms translation interactions into learning opportunities. When users input Thai sentences like "chan gin khao chao laew" (I ate breakfast already), the system provides not only the translation "I have eaten breakfast" but explains why the Thai temporal marker "laew" (already) indicates present perfect usage rather than past simple. The system acts as a knowledgeable tutor who understands both linguistic systems and can bridge conceptual differences.

Key Contributions: (1) A novel three-model pipeline architecture integrating translation, classification, and explanation generation for educational applications; (2) Hierarchical tense classification achieving 94.7% accuracy on 24 fine-grained categories; (3) Comprehensive user study with 218 Thai learners validating design decisions and system testing with 38 active users demonstrating strong user acceptance (4.2/5 satisfaction) despite technical limitations; (4) Evidence that educational value can exceed pure technical performance in AI systems designed for learning contexts.

Methodology

System Architecture

The system implements a three-model pipeline architecture where specialised components excel at distinct educational functions rather than attempting unified processing. This design emerged from discovering that instruction-following models, while capable of generating pedagogically appropriate explanations, consistently misclassified English tenses—a critical flaw for educational applications.

Translation Component: Typhoon Translate 4B [1] processes Thai input for efficient inference. The model was selected for its Thai-English specialisation and resource efficiency, achieving 93.2% fluency and 90.6% semantic accuracy on manual evaluation.

Classification Component: Custom XLM-RoBERTa classifier trained on 15,755 English sentences with hierarchical tense labels covering 24 fine-grained categories. The hierarchical architecture employs dual classification heads: coarse-grained

*MSc Computer Science with Speech and Language Processing student at University of Sheffield

(Past/Present/Future) and fine-grained temporal distinctions, ensuring temporal consistency through joint optimisation (Figure 1).

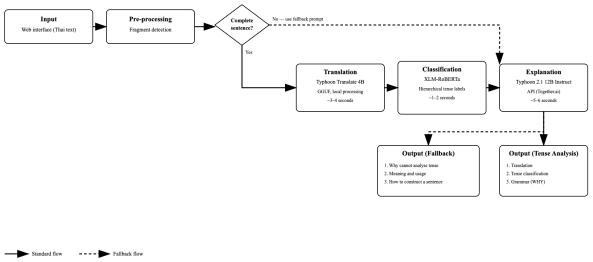

Figure 1: *Complete System Architecture for Educational Language Processing*

The 24-Category Tense Taxonomy

The classification system implements a custom taxonomy developed with reference to established Thai grammar teaching materials [2], designed to address specific challenges Thai learners face. Unlike standard 12-tense classifications designed for linguistic analysis, this 24-category system prioritizes pedagogical effectiveness for Thai-English learners. Thai lacks grammatical tense markers, relying instead on contextual cues (temporal words like "meuua waan nii" for yesterday), creating conceptual gaps when learning English verb conjugation. The taxonomy provides granular distinctions crucial for Thai learners (HABIT vs FACT, NOWADAYS vs HAPPENING) while mapping Thai temporal expressions to appropriate English structures, addressing typical learning errors through more precise contextual categorisation (Table 1).

Explanation Component: Typhoon 2.1 12B Instruct [3] generates educational explanations by processing classification results alongside original Thai context to produce pedagogically appropriate explanations for Thai learners.

Code	Category
Present Tenses (13 categories)	
HABIT	Habit/Routine
FACT	General Fact
SCHEDULEDFUTURE	Scheduled/Planned
SAYING	Proverb/Saying
HEADLINE	News Headline
HAPPENING	Currently Happening
NOWADAYS	Present Continuous
SUREFUT	Planned Future
PROGRESS	Ongoing Change
JUSTFIN	Just Finished
RESULT	Present Result
EXP	Life Experience
SINCEFOR	Duration Period
Past Tenses (5 categories)	
NORFIN	General Past
INTERRUPT	Interrupted Action
DOINGATSOMETIMEPAST	Past Continuous
BEFOREPAST	Before Past Action
DURATION	Past Duration
Future Tenses (6 categories)	
50PERC	50% Prediction
PROMISE	Promise/Commitment
RIGHTNOW	Future Simple
LONGFUTURE	Distant Future
PREDICT	Future Prediction
WILLCONTINUEINFUTURE	Future Duration

Table 1: *Complete 24-Category Tense Classification Taxonomy*

Evaluation Methodology

Empirical validation was conducted through two phases: comprehensive requirements analysis with 218 Thai English learners and system testing with 38 active users.

Requirements Analysis Phase: The study employed an online survey distributed through university networks and social media platforms, receiving 234 initial responses with 218 participants (93.2%) providing complete data after ethical consent filtering. The survey examined five key areas: demographics and background, grammar learning challenges, current tool usage, feature preferences, and technology acceptance. Participants included working professionals (80%) and students (20%), representing diverse English proficiency levels from A2 to C2.

System Testing Phase: Following requirements validation, 38 participants from the target demographic interacted directly with the deployed system. Users provided 474 ratings across four criteria: translation accuracy, explanation quality, educational value, and overall satisfaction using 5-point Likert scales. Qualitative feedback was collected through

open-ended responses to identify improvement opportunities and validate design decisions.

Results

User Study Results

System Testing Results: User testing with 38 active participants generated 474 ratings across evaluation criteria, demonstrating strong user acceptance despite technical limitations. Mean satisfaction scores exceeded 4.0/5 across all categories: translation accuracy (4.1/5), explanation quality (4.33/5), educational value (4.2/5), and overall satisfaction (4.2/5) (Table 2).

Table 2: *User Satisfaction Ratings (n=38 users, 474 ratings)*

Criterion	Mean Rating (/5)
Translation Accuracy	4.1
Explanation Quality	4.33
Educational Value	4.2
Overall Satisfaction	4.2

Requirements Analysis Results: The comprehensive analysis with 218 participants revealed critical insights about Thai English learners' needs and preferences. Figure 2 shows that working professionals dominated respondents (80%), challenging assumptions about primarily academic usage and highlighting workplace English improvement as a key driver.

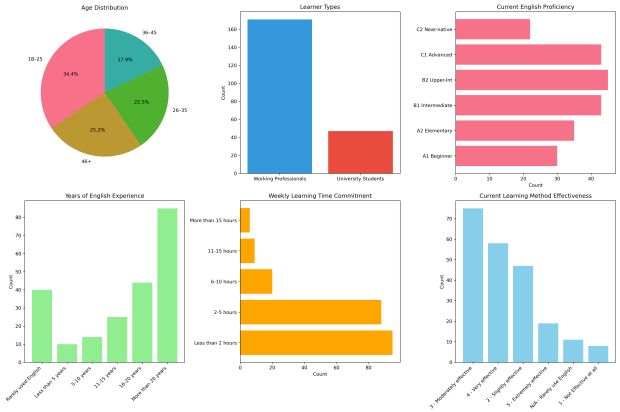

Figure 2: *User Study Demographics and Learning Context (n=218)*

Learning challenges analysis (Figure 3) revealed grammar ranked third (13.4%) among learner difficulties, representing a computationally tractable challenge unlike speaking/listening which require real-time interaction.

Pedagogical preferences (Figure 4) validated explanation-focused design: 73.4% of learners rated understanding "why" grammar rules work as highly

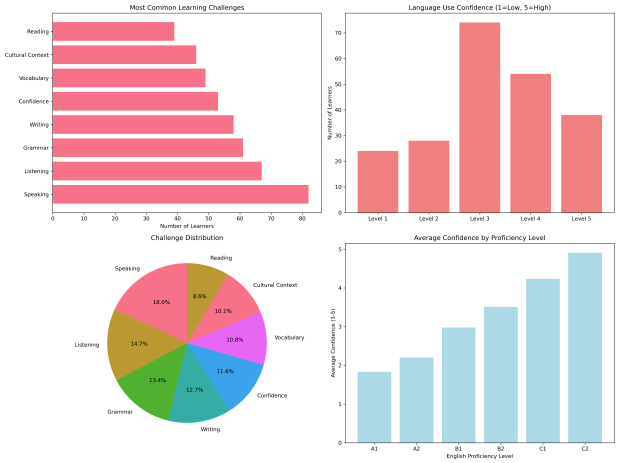

Figure 3: *Learning Challenges and Confidence Analysis Revealing Grammar as Computationally Tractable Challenge*

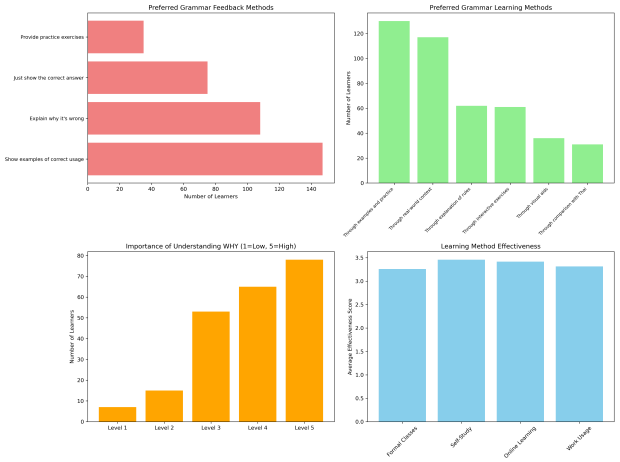

Figure 4: *Grammar Learning Preferences Showing Strong Demand for Explanatory Approaches*

important, with a 2:1 preference for examples over simple corrections. Current learning methods showed uniform mediocrity (3.2-3.5 effectiveness ratings), indicating market opportunity rather than user satisfaction.

Correlation analysis (Figure 5) uncovered crucial design insights. The strong English proficiency-confidence correlation (0.81) suggests tools building genuine understanding may be more effective than quick-fix solutions. Importantly, the desire for explanatory learning showed weak demographic correlations (<0.35), indicating universal appeal across age and proficiency levels. This pattern validates explanation-focused positioning as broadly applicable rather than demographically specific.

Technology readiness analysis (Figure 6) revealed high AI tool familiarity (75.7% of participants), indicating learners expect adaptive responses rather than static rule presentations. Feature prioritisation showed strong demand for mobile accessibility and alternative translations, suggesting learners seek ex-

ploratory rather than corrective learning approaches.

Qualitative feedback revealed that users valued clear explanations even when imperfect, preferring educational context over pure accuracy. This finding suggests that pedagogical design principles can compensate for technical limitations in educational AI systems, providing important insights for future development.

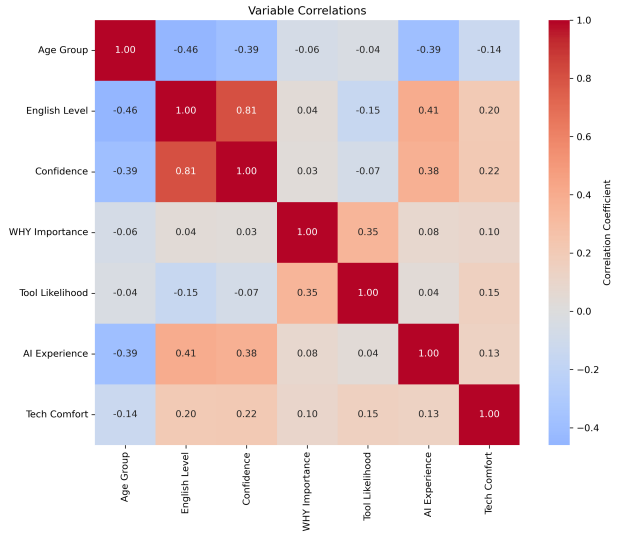

Figure 5: *Statistical Relationships Revealing Key Design Insights*

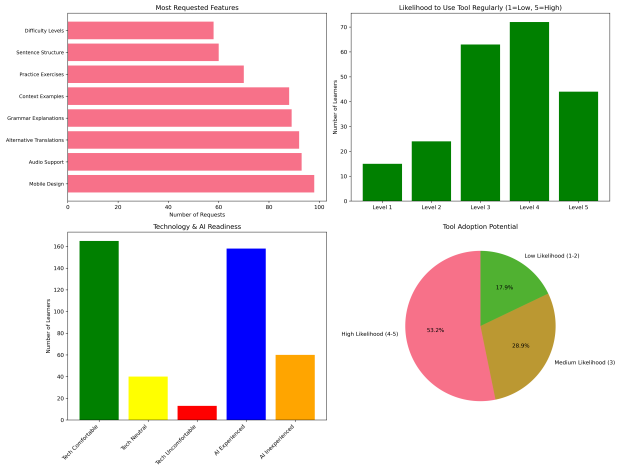

Figure 6: *Technology Adoption Readiness and Feature Priorities*

Technical Performance

Individual model validation demonstrated strong performance across components. The XLM-RoBERTa classifier achieved 94.7% accuracy on held-out test data with hierarchical consistency of 99.2% between coarse and fine-grained predictions. Translation quality assessment on 96 manually evaluated samples yielded 93.2% fluency and 90.6% semantic accuracy, with particular strength in preserving temporal markers essential for educational applications. Performance varied significantly across tense categories, with some complex temporal relationships proving more challenging (Figure 7).

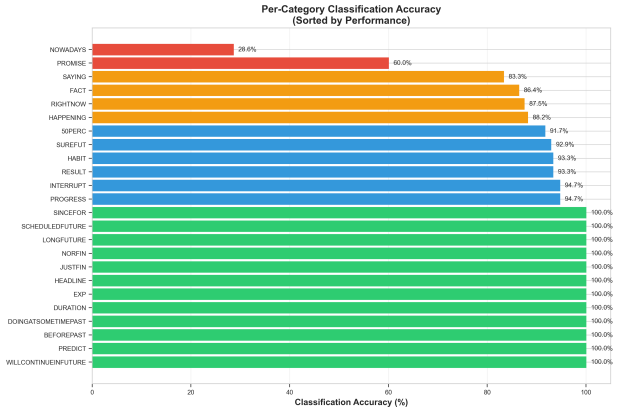

Figure 7: *Classification Accuracy by Tense Category Showing Performance Variation*

Explanation quality evaluation on the same sample set showed 84.9% correctness in grammatical explanations, with 23% experiencing code hallucination issues that were addressed through template-based post-processing. The explanation model demonstrated strong pedagogical appropriateness, providing clear connections between Thai temporal markers and English tense usage.

Pipeline integration introduced performance degradation, with end-to-end classification accuracy dropping to 74% due to error propagation between components. However, this technical limitation did not significantly impact user satisfaction, suggesting that explanation quality and educational context compensate for classification imperfections in learning applications.

Table 3: *System Performance Metrics*

Component	Performance
Translation Fluency	93.2%
Translation Semantic Accuracy	90.6%
Classification (Isolated)	94.7%
Classification (Pipeline)	74.0%
Explanation Correctness	84.9%

Discussion and Future Work

Research Contributions

This work contributes to educational NLP through several innovations. The three-model pipeline architecture demonstrates that specialised components

can outperform unified approaches for educational applications, providing a template for similar cross-linguistic learning challenges. The hierarchical tense classification approach addresses the complexity of temporal relationships in language learning, while the comprehensive user study reveals that educational value often exceeds pure technical performance metrics.

The disconnect between pipeline classification accuracy (74%) and user satisfaction (4.2/5) provides crucial insights for educational AI design. The user study validation with 218 participants showed that learners prioritize clear explanations and educational context over perfect accuracy, with 73.4% rating explanatory understanding as highly important. This suggests that pedagogical principles should guide system design rather than purely technical optimisation.

The empirical validation established three key design principles: (1) Grammar represents a computationally tractable challenge among Thai learners' difficulties, ranking third (13.4%) after speaking/listening which require real-time interaction; (2) Explanation-focused approaches have universal appeal across demographics, with weak correlations (<0.35) to age and proficiency levels; (3) The strong proficiency-confidence correlation (0.81) validates building genuine understanding rather than providing quick fixes.

The system successfully demonstrates educational effectiveness through user acceptance and satisfaction ratings, establishing foundations for AI-powered grammar learning systems.

Limitations and Future Research

Several limitations provide opportunities for future research. The manual evaluation methodology introduces potential bias, while the 20% performance gap between isolated and pipeline classification suggests systematic error propagation requiring investigation. The user study focused on short-term satisfaction rather than longitudinal learning effectiveness, limiting educational impact assessment.

Technical limitations include dependency on Thai-specific models limiting generalisation and manual dataset curation restricting scalability. Educational limitations encompass restricted grammar scope beyond tense usage and lack of adaptive learning mechanisms responding to individual progress.

Future research directions include longitudinal studies measuring actual learning outcomes, expansion to additional grammatical categories, adaptive explanation generation based on learner proficiency, and extension to other under-resourced language pairs. Cross-cultural validation would establish broader applicability of the pedagogical approach.

Technical extensions could explore advanced error propagation modeling, multi-task learning architectures combining translation and explanation generation, and adaptive systems responding to individual learning patterns. Educational research opportunities include controlled studies measuring learning effectiveness, investigation of explanation strategies across different proficiency levels, and development of comprehensive grammatical coverage beyond tense usage. Cross-linguistic research could examine applicability to other language pairs with fundamental structural differences, while the intersection of NLP, educational technology, and cross-cultural learning presents opportunities for advancing both theoretical understanding and practical applications in educational AI systems.

References

[1] Kunat Pipatanakul et al. "Typhoon: Thai Large Language Models". In: *arXiv preprint arXiv:2312.13951* (2023). DOI: 10.48550/arXiv.2312.13951. URL: https://arxiv.org/abs/2312.13951.

[2] Phatra Phatraphureeraks. *Perfect English Grammar: Thai Grammar Textbook for English Learning*. Thai language grammar textbook for English learners. Bangkok, Thailand: MIS Book, 2020, p. 480. ISBN: 9786164300729.

[3] Kunat Pipatanakul et al. "Typhoon 2: A Family of Open Text and Multimodal Thai Large Language Models". In: *arXiv preprint arXiv:2412.13702* (2024). DOI: 10.48550/arXiv.2412.13702. URL: https://arxiv.org/abs/2412.13702.