Mẫu 3A-4

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

THÔNG TIN VỀ LUẬN ÁN TIẾN SĨ

1. Họ và tên nghiên cứu sinh: Hà Mỹ Linh 2. Giới tính: Nữ

3. Ngày sinh: 12/05/1990 4. Nơi sinh: Ninh Bình

5. Quyết định công nhận nghiên cứu sinh: 2251/QĐ-ĐHKHTN ngày 19/07/2019 của Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội.

6. Các thay đổi trong quá trình đào tạo (nếu có): Quyết định gia hạn số 151/QĐ-ĐHQGHN, 1073/QĐ-ĐHQGHN của của Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội.

7. Tên đề tài luận án: Nghiên cứu các phương pháp biểu diễn và phát triển ngữ liệu, công cụ cho phân tích cú pháp và ngữ nghĩa tiếng Việt.

8. Chuyên ngành: Cơ sở Toán học cho Tin học 9. Mã số: 9460117.02

10. Cán bộ hướng dẫn khoa học:

- HDC: GS. TS Nguyễn Lê Minh, Viện Khoa học và Công nghệ tiên tiến Nhật Bản (JAIST) .
- HDP*:* TS Nguyễn Thị Minh Huyền, Trường Đại học Khoa học Tự nhiên – ĐHQGHN

11. Tóm tắt các kết quả mới của luận án:

11.1. Mục đích và đối tượng nghiên cứu của luận án

Ngữ liệu có chú giải thông tin ngôn ngữ đóng vai trò nền tảng trong xử lý ngôn ngữ tự nhiên, giúp huấn luyện và đánh giá các mô hình một cách chính xác, hiệu quả. Đồng thời, chúng là nguồn dữ liệu quý giá phục vụ nghiên cứu ngôn ngữ học, góp phần khám phá và lý giải các hiện tượng ngôn ngữ một cách có hệ thống. Tuy nhiên, đối với tiếng Việt, các tài nguyên ngôn ngữ có chú giải và các công cụ phân tích cú pháp và ngữ nghĩa vẫn còn hạn chế cả về quy mô lẫn mức độ hoàn thiện. Luận án này nhằm mục đích đóng góp vào sự phát triển của lĩnh vực Xử lý Ngôn ngữ Tự nhiên tiếng Việt thông qua việc (1) xây dựng một hệ thống tài nguyên ngôn ngữ tiếng Việt phong phú, đa dạng, được chú giải ở mức sâu, phù hợp với các lược đồ chú giải chuẩn quốc tế, tập trung vào

kho từ vựng động từ và các ngữ liệu chú giải cú pháp – ngữ nghĩa; và (2) nghiên cứu, phát triển và đánh giá các công cụ phân tích cú pháp và ngữ nghĩa tiên tiến, được tối ưu hóa cho đặc thù của tiếng Việt. Quá trình này mang tính tương hỗ: tài nguyên cung cấp dữ liệu cho việc phát triển công cụ, và công cụ lại trợ giúp việc xây dựng, mở rộng tài nguyên, tạo nền tảng vững chắc cho việc phát triển các ứng dụng phân tích văn bản tiếng Việt.

Các mục tiêu cụ thể của luận án bao gồm:

- Xây dựng ngữ liệu:
  + Kho văn bản: Thiết kế lược đồ chú giải ngữ pháp/ngữ nghĩa, xây dựng kho ngữ liệu được chú giải cú pháp (cú pháp thành phần, cú pháp phụ thuộc) và ngữ nghĩa. Các tài nguyên này được thiết kế dựa trên các mô hình chuẩn hóa quốc tế, bảo đảm tính nhất quán, khả năng mở rộng và tính tương thích với các hệ thống xử lý đa ngôn ngữ;
  + Từ vựng: Thiết kế, xây dựng mạng động từ tiếng Việt (viVerbNet).
- Phát triển công cụ:
  + Nghiên cứu, phát triển, tinh chỉnh và đánh giá các mô hình phân tích cú pháp và ngữ nghĩa cho tiếng Việt, nhằm vừa hỗ trợ quá trình gán nhãn ngữ liệu, vừa tận dụng chính các ngữ liệu này để cải thiện hiệu suất của các mô hình phân tích ngữ pháp/ngữ nghĩa tiếng Việt.

11.2. Các phương pháp nghiên cứu đã sử dụng

Luận án kết hợp phương pháp nghiên cứu lý thuyết, thực nghiệm và định lượng.

Về lí thuyết, luận án khảo sát, phân tích tài liệu ngôn ngữ học và các sơ đồ chú giải sẵn có để xây dựng các sơ đồ chú giải và hệ thống hóa kho từ vựng, ngữ liệu cú pháp, ngữ nghĩa tiếng Việt, đảm bảo tính tương thích với các kho ngữ liệu cho các ngôn ngữ khác và phù hợp với đặc thù tiếng Việt. Phương pháp thực nghiệm được triển khai qua việc xây dựng, thử nghiệm và đánh giá các mô hình phân tích cú pháp, ngữ nghĩa văn bản và mô hình phân cụm cho mạng động từ. Cuối cùng, phương pháp định lượng dùng để phân tích thống kê, đo lường, đánh giá hiệu quả mô hình và chất lượng kho ngữ liệu. Việc kết hợp các phương pháp này giúp nâng cao hiệu quả, độ tin cậy của các tài nguyên và công cụ được xây dựng.

11.3. Các kết quả chính

Luận án đã có những đóng góp cơ bản về hai hướng chính:

- Xây dựng các tài nguyên ngôn ngữ, gồm các lược đồ chú giải, hướng dẫn chú giải và kho ngữ liệu có chú giải theo lược đồ đã thiết kế:
  + Cú pháp phụ thuộc: Trên cơ sở tập nhãn cú pháp phụ thuộc tiếng Việt đã xây dựng trong giai đoạn trước, luận án tiến hành cập nhật, thiết kế lại và chỉnh sửa tập nhãn và hướng dẫn chú giải theo phiên bản 2.0 của Dự án cú pháp phụ thuộc phổ quát (*Universal Dependency - UD*). Kho ngữ liệu gồm hơn 9,000 câu (trong đó 3,000 câu đã được tích hợp vào kho UD hồi tháng 11

năm 2022) đã được xây dựng và sử dụng trong cuộc thi về phân tích cú pháp phụ thuộc tiếng Việt tại hội thảo về Xử lí ngôn ngữ tự nhiên và tiếng nói tiếng Việt (VLSP 2020).

+ Cú pháp thành phần: Kế thừa kho ngữ liệu cú pháp thành phần Viettreebank, luận án thực hiện việc rà soát, cập nhật và chuẩn hoá các nhãn cú pháp thành phần cũng như tài liệu hướng dẫn chú giải để có một bộ nhãn phù hợp với các nghiên cứu đối sánh đa ngữ. Trên cơ sở đó, kho ngữ liệu gồm hơn 9,000 câu đã được cập nhật theo tập nhãn mới và được sử dụng trong cuộc thi về phân tích cú pháp thành phần tiếng Việt tại hội thảo về Xử lý ngôn ngữ tự nhiên và tiếng nói tiếng Việt (VLSP 2022 và VLSP 2023).
+ Ngữ nghĩa nông (Vai nghĩa): Tập nhãn vai nghĩa cho tiếng Việt xây dựng trước đó đã được cập nhật và chỉnh sửa tương thích với khung chú giải vai nghĩa trong dự án Universal Proposition Bank 2.0. Tập nhãn này đã được sử dụng để xây dựng kho ngữ liệu tiếng Việt có chú giải vai nghĩa gồm 2,570 câu.
+ Ngữ nghĩa sâu: Thiết kế mô hình biểu diễn ngữ nghĩa sâu cho tiếng Việt và hướng dẫn chú giải dựa vào mô hình ngữ nghĩa trừu tượng của tiếng Anh (AMR) và tập hợp vai nghĩa LIRICS được thiết kế hướng chuẩn ISO. Kho ngữ liệu tiếng Việt gồm có 1,570 câu đã được xây dựng theo mô hình này.
+ Mạng động từ tiếng Việt: Xây dựng lược đồ chú giải mạng động từ cho tiếng Việt trên cơ sở tham chiếu VerbNet tiếng Anh, với 5 thành phần chính: vai nghĩa, ràng buộc lựa chọn, khung cú pháp, ràng buộc cú pháp và vị từ ngữ nghĩa. Sau đó, mạng động từ tiếng Việt (viVerbNet) gồm 100 cụm động từ được phát triển và gán nhãn theo lược đồ đã đề xuất.

- Về phương pháp và mô hình cho phân tích tiếng Việt:
  + Đánh giá, so sánh các mô hình véc-tơ từ huấn luyện sẵn cho tiếng Việt và một số phương pháp hiện đại để cải tiến hiệu quả của bài toán phân tích cú pháp.
  + Xây dựng công cụ chuyển đổi giữa cú pháp thành phần và cú pháp phụ thuộc, hỗ trợ quá trình gán nhãn dữ liệu.
  + Phát triển và đánh giá các thuật toán phân cụm động từ tiếng Việt.
  + Thử nghiệm các mô hình ngôn ngữ lớn cho bài toán gán nhãn vai nghĩa và phân tích ngữ nghĩa cho tiếng Việt, đánh giá kết quả đạt được trên bộ dữ liệu đã xây dựng.

12. Các hướng nghiên cứu tiếp theo:

Với các kết quả đạt được, luận án đã có những đóng góp quan trọng trong các bài toán phân tích cú pháp và ngữ nghĩa, đặc biệt là trong việc phát triển tài nguyên ngôn ngữ tiếng Việt. Thật vậy, trong khi các mô hình tính toán liên tục được cải tiến với tốc độ nhanh chóng, thì các kho ngữ liệu có chú giải vẫn giữ vai trò nền tảng, bền vững và không thể thay thế trong việc đánh giá cũng như hỗ trợ phát triển các công cụ phân tích ngôn ngữ.

Trong thời gian tới, hướng phát triển tiếp theo của luận án là:

- Tiếp tục phát triển kho ngữ liệu ngữ nghĩa, chia sẻ và công bố rộng rãi các tài nguyên này trong cộng đồng xử lý ngôn ngữ tiếng Việt.
- Tinh chỉnh và khai thác các mô hình ngôn ngữ lớn cho các bài toán phân tích cú pháp và ngữ nghĩa tiếng Việt, nhằm tăng hiệu quả và giảm chi phí xây dựng tài nguyên.
- Phát triển tiếp mạng động từ viVerbNet.
- Cải thiện chất lượng kho ngữ liệu:
  + Đánh giá và điều chỉnh sơ đồ chú giải để tăng tính nhất quán và khả năng sử dụng lại. Xây dựng bộ tiêu chí đánh giá chất lượng chú giải và đề xuất các công cụ hỗ trợ hiệu chỉnh bán tự động.
  + Tiếp tục khảo sát và phân tích các hiện tượng ngôn ngữ đặc thù của tiếng Việt trong ngữ liệu, từ đó đề xuất các biểu diễn ngữ nghĩa phù hợp hơn cho các cấu trúc cú pháp–ngữ nghĩa độc đáo của tiếng Việt.

13. Các công trình công bố liên quan đến luận án:

1. **Ha My Linh**, Nguyen Thi Minh Huyen, "*A Case Study on Meaning Representation for Vietnamese*", Proceedings of the First International Workshop on Designing Meaning Representations, pages 148-153, 2019, Italy, (*ISBN 978-1-950737-45-1*)
2. **Ha My Linh**, Nguyen Thi Minh Huyen, Vu Xuan Luong, Nguyen Thi Luong, Phan Thi Hue, Le Van Cuong, "*VLSP 2020 shared task: Universal dependency parsing for Vietnamese*", Proceedings of the 7th International Workshop on Vietnamese Language and Speech Processing, pages 77-83, 2020, Vietnam.
3. **Ha My Linh**, Le Van Cuong, Nguyen Thi Minh Huyen, "*Construction of a VerbNet style Lexicon for Vietnamese*", Proceedings of the 34th Pacific Asia Conference on Language, pages 84-91, 2020, Vietnam, (*ISSN 2619-7782*).
4. **Ha My Linh**, Do Duy Dao, Nguyen Thi Minh Huyen, Tran Thu Trang, "*Using rules for building Vietnamese AMR-based corpus*", Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông lần thứ XXIV, pages: 547-552, 2021, Vietnam.
5. Ishan Jindal, ALexandre Rademaker, Micha l ULewicz, **Ha My Linh**, Huyen Nguyen, Khoi-Nguyen Tran, Huaiyu Zhu, and Yunyao Li, "*Universal Proposition Bank 2.0* ". In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 1700–1711, 2022, MarseilLe, France, European Language Resources Association, (*ISBN: 979-10-95546-72-6*).
6. **Ha My Linh**, Do Duy Dao, Nguyen Thi Minh Huyen, Ngo The Quyen, Doan Xuan Dung, "*VLSP 2021 - NER Challenge: Named Entity Recognition for Vietnamese*", VNU Journal of Science: Computer Science and Communication Engineering, pages 87-97, Vol. 38.1, 2022, (*VNU Journal of Science ISSN: 0866-8612*).
7. The Quyen Ngo, Thi Anh Phuong Nguyen, **Ha My Linh**, Thi Minh Huyen Nguyen, Phuong Le-Hong, "*Improving Multi-label Classification of Similar Languages by Semantics-Aware Word Embeddings*", In Eleventh Workshop on NLP for Similar Languages, Varieties and DiaLects (VarDial 2024), pages 235-240, Mexico, Association for Computational Linguistics, (*ISBN 979-8- 89176-104-9*).

8. **Ha My Linh**, Thi Minh Huyen Nguyen, The Quyen Ngo, Tuan Thanh Le, Tran Thai Dang, Viet Hoang Ngo, Xuan Dung Doan, Thi Luong Nguyen, Van Cuong Le, Thi Hue Phan, Xuan Luong Vu, "*VLSP 2022 Challenge: Vietnamese Constituency Parsing*", **accepted** to Journal of Computer Science and Cybernetics, 2024, (*ISSN 2815-5939*).

9. Ha My Linh, Pham Thi Duc, Le Ngoc Toan, Thi Minh Huyen Nguyen, "*An Investigation of ISO-TimeML Applied to Vietnamese*", in Proceedings of the 38th Pacific Asia Conference on Language, Information, and Computation, PACLIC 38 (2024), Tokyo, Japan, pages 1387 - 1394.

Hà Nội, ngày ..... tháng.....năm.....

TM. Tập thể hướng dẫn	**Nghiên cứu sinh**
(Ký và ghi rõ họ tên)	*(Ký và ghi rõ họ tên)*

INFORMATION ON DOCTORAL THESIS

1. Full name: Ha My Linh
2. Sex: Female
3. Date of birth: 12/05/1990
4. Place of birth: Ninh Binh
5. Admission decision number 2251/QĐ-ĐHKHTN dated 19/07/2019 by VNU University of Science
6. Changes in academic process: 151/QĐ-ĐHQGHN, 1073/QĐ-ĐHQGHN
7. Official thesis title: Methods for designing meaning representations and developing linguistic resources and tools for Vietnamese syntactic and semantic analysis.
8. Major: Mathematical Foundations for Computer Science
9. Code: 9460117.02
10. Supervisors:

- Principal Supervisor: Prof. Dr. Nguyen Le Minh, Japan Advanced Institute of Science and Technology (JAIST).
- Co-Supervisor: Dr. Nguyen Thi Minh Huyen, University of Science – Vietnam National University, Hanoi.

11. Summary of the new findings of the thesis

11.1. Objectives and scopes of the thesis

Annotated corpora play a foundational role in Natural Language Processing (NLP), supporting the accurate and efficient training and evaluation of models. At the same time, these corpora serve as valuable resources for linguistic research, contributing to the systematic exploration and analysis of linguistic phenomena. However, for Vietnamese, annotated language resources and syntactic and semantic analysis tools are still limited in scale and completeness.

This dissertation aims to contribute to the advancement of Vietnamese NLP through: (1) the creation of a substantial and varied collection of Vietnamese language resources, annotated in-depth, adhering to international annotation standards, with a focus on a verb lexicon and corpora annotated for syntax and semantics; and (2) the investigation, development, and evaluation of advanced syntactic and semantic parsing tools optimized for the specific features of the Vietnamese language. This process is mutually reinforcing: the resources provide data for developing the tools, and these tools, in return, aid the building and broadening of the resources, establishing a solid foundation for the development of Vietnamese text analysis applications.

The objectives of the dissertation include:

- Language resource construction:
  + Text corpora: Design annotation schemes for syntax and semantics, and build distinct types of corpora: treebanks (annotated syntactically with both constituency and dependency structures) and a Sembank (annotated semantically). These resources are designed based on international standardization models, ensuring consistency, scalability, and compatibility with multilingual processing systems.
  + Lexical resources: Design and create a lexical resource for Vietnamese verbs compatible with the English VerbNet (viVerbNet).
- Tool development**:** Investigate, develop, fine-tune, and evaluate models for Vietnamese syntactic and semantic analysis, with the goals of both assisting the data annotation process and using these annotated corpora to improve the performance of Vietnamese grammatical and semantic analysis models.

11.2. Research methods

The thesis combines theoretical, empirical, and quantitative research methods.

Theoretically, the dissertation examines existing linguistic literature and annotation frameworks to propose new annotation schemes and cultivate an ecosystem encompassing Vietnamese lexical, syntactic, and semantic resources. Such a method ensures cross-linguistic compatibility while also being tailored to the characteristics of the Vietnamese language.

The empirical approach involves building, testing, and evaluating models for syntactic and semantic analysis, alongside clustering models for defining verb classes in a lexical resource for verbs.

Finally, quantitative methods are employed for statistical analysis, measurement, and evaluation of model performance and the quality of the language resource collections.

Combining these methods helps enhance the effectiveness and reliability of the developed resources and tools.

11.3. Main results

The dissertation has made fundamental contributions in two main directions:

- Development of language resources, including annotation schemes, annotation guidelines, and annotated corpora based on the designed schemes:
  + Dependency syntax: Based on the Vietnamese dependency labels developed in the previous phase, the dissertation revises, redesigns, and updates the dependency labels and annotation guidelines following the Universal Dependency (UD) version 2.0. A corpus comprising over 9,000 sentences (including 3,000 sentences integrated into the UD in November 2022) was developed and employed in the Vietnamese dependency parsing shared task

at the seventh international workshop on Vietnamese Language and Speech Processing (VLSP 2020).

+ Constituency syntax: Inheriting the Vietnamese constituency treebank (*VietTreebank*), the dissertation revises, updates, and standardizes the Vietnamese constituency labels and annotation guidelines to establish a label set suitable for cross-linguistic comparative studies. Based on this work, a corpus containing over 9,000 sentences was re-annotated using the updated label set. This corpus was utilized in the Vietnamese constituency parsing shared tasks at VLSP 2022 and VLSP 2023.
+ Shallow semantics (semantic role labelling): The existing Vietnamese semantic role label set was refined and updated to ensure compatibility with the Universal Proposition Bank 2.0. This label set served as the foundation for the construction of a Vietnamese semantic role labeling corpus of 2,570 sentences.
+ Deep semantics: A Vietnamese model for deep semantic representation was developed, along with its annotation guidelines, taking inspiration from both the Abstract Meaning Representation (AMR) model and LIRICS (The ISO-oriented semantic role set). Following this designed model, annotation was subsequently carried out on a Vietnamese corpus comprising 1,570 sentences.
+ Vietnamese VerbNet: A framework for Vietnamese verb annotation was formulated, drawing inspiration from English VerbNet and comprising five principal components: semantic roles, selectional restrictions, syntactic frames, syntactic constraints, and predicate semantics. Using this established scheme, the Vietnamese VerbNet (viVerbNet) was then built, featuring 100 verb classes that were duly annotated.

- Methods and models for analyzing the Vietnamese language:
  + Evaluation and comparative study of pre-trained Vietnamese word embeddings alongside various contemporary techniques for enhancing parsing performance.
  + Creation of a utility for converting between constituency and dependency syntactic representations to assist in data annotation tasks.
  + Development and evaluation of algorithms for clustering Vietnamese verbs.
  + Exploration of large language models applied to Vietnamese semantic role labeling and semantic parsing, with their effectiveness evaluated using the datasets developed in this research.

12. Future research directions

With the results achieved, this dissertation has made significant contributions to both syntactic and semantic analysis, most notably through the construction of Vietnamese language resources. In fact, even as computational models advance rapidly, annotated

corpora remain fundamental, enduring, and indispensable resources for the evaluation and enhancement of syntactic and semantic parsing tools.

Potential future research trajectories arising from this dissertation are as follows:

- Ongoing enhancement of Sembank, coupled with sharing and broadly distributing this resource among the Vietnamese language processing community.
- Fine-tuning and applying large language models to Vietnamese syntactic and semantic analysis tasks to boost performance while simultaneously decreasing resource creation costs.
- Sustained advancement and expansion of viVerbNet.
- Enhancing the quality of language corpora through measures such as:
  + Evaluating and revising annotation schemes to improve consistency and reusability; establishing criteria for the annotation quality evaluation; and proposing tools to facilitate semi-automatic revision processes.
  + Conducting deeper investigations and analyses of Vietnamese-specific linguistic phenomena unique to Vietnamese within corpora, in turn proposing more fitting semantic representations for the distinctive syntactic-semantic structures of Vietnamese.

13. Thesis-related publications:

1. **Ha My Linh**, Nguyen Thi Minh Huyen, "*A Case Study on Meaning Representation for Vietnamese*", Proceedings of the First International Workshop on Designing Meaning Representations, pages 148-153, 2019, Italy, (*ISBN 978-1-950737-45-1*)
2. **Ha My Linh**, Nguyen Thi Minh Huyen, Vu Xuan Luong, Nguyen Thi Luong, Phan Thi Hue, Le Van Cuong, "*VLSP 2020 shared task: Universal dependency parsing for Vietnamese*", Proceedings of the 7th International Workshop on Vietnamese Language and Speech Processing, pages 77-83, 2020, Vietnam.
3. **Ha My Linh**, Le Van Cuong, Nguyen Thi Minh Huyen, "*Construction of a VerbNet style Lexicon for Vietnamese*", Proceedings of the 34th Pacific Asia Conference on Language, pages 84-91, 2020, Vietnam, (*ISSN 2619-7782*).
4. **Ha My Linh**, Do Duy Dao, Nguyen Thi Minh Huyen, Tran Thu Trang, "*Using rules for building Vietnamese AMR-based corpus*", Một số vấn đề chọn lọc của công nghệ thông tin và truyền thông lần thứ XXIV, pages: 547-552, 2021, Vietnam.
5. Ishan Jindal, ALexandre Rademaker, Micha l ULewicz, **Ha My Linh**, Huyen Nguyen, Khoi-Nguyen Tran, Huaiyu Zhu, and Yunyao Li, "*Universal Proposition Bank 2.0* ". In Proceedings of the Thirteenth Language Resources and Evaluation Conference, pages 1700–1711, 2022, MarseilLe, France, European Language Resources Association, (*ISBN: 979-10-95546-72-6*).

6. **Ha My Linh**, Do Duy Dao, Nguyen Thi Minh Huyen, Ngo The Quyen, Doan Xuan Dung, "*VLSP 2021 - NER Challenge: Named Entity Recognition for Vietnamese*", VNU Journal of Science: Computer Science and Communication Engineering, pages 87-97, Vol. 38.1, 2022, (*VNU Journal of Science ISSN: 0866-8612*).
7. The Quyen Ngo, Thi Anh Phuong Nguyen, **Ha My Linh**, Thi Minh Huyen Nguyen, Phuong Le-Hong, "*Improving Multi-label Classification of Similar Languages by Semantics-Aware Word Embeddings*", In Eleventh Workshop on NLP for Similar Languages, Varieties and DiaLects (VarDial 2024), pages 235-240, Mexico, Association for Computational Linguistics, (*ISBN 979-8- 89176-104-9*).
8. **Ha My Linh**, Thi Minh Huyen Nguyen, The Quyen Ngo, Tuan Thanh Le, Tran Thai Dang, Viet Hoang Ngo, Xuan Dung Doan, Thi Luong Nguyen, Van Cuong Le, Thi Hue Phan, Xuan Luong Vu, "*VLSP 2022 Challenge: Vietnamese Constituency Parsing*", **accepted** to Journal of Computer Science and Cybernetics, 2024, (*ISSN 2815-5939*).
9. Ha My Linh, Pham Thi Duc, Le Ngoc Toan, Thi Minh Huyen Nguyen, "*An Investigation of ISO-TimeML Applied to Vietnamese*", in Proceedings of the 38th Pacific Asia Conference on Language, Information, and Computation, PACLIC 38 (2024), Tokyo, Japan, pages 1387 - 1394.

Date:

On behalf of academic supervisors **PhD. Student**