

THUYẾT MINH ĐỀ TÀI KH&CN CẤP CƠ SỞ NĂM 2017

1. TÊN ĐỀ TÀI: Xây dựng Module tổng hợp tiếng nói tiếng Mường cho hai phương ngữ Mường Bi – Hòa Bình và Mường Tân Sơn - Phú Thọ	2. MÃ SỐ T17-40		
3. LĨNH VỰC NGHIÊN CỨU Tự nhiên ☐ Kỹ thuật ☒ Môi trường ☐ Kinh tế ☐ Nông Lâm ☐ ATLĐ ☐ Giáo dục ☐ Y Dược ☐ SH trí tuệ ☐	4. LOẠI HÌNH NGHIÊN CỨU Cơ bản ☒ Ứng dụng ☒ Triển khai ☐		
5. THỜI GIAN THỰC HIỆN 12 tháng			
6. CHỦ NHIỆM ĐỀ TÀI Họ và tên: PHẠM VĂN ĐỒNG Chức danh khoa học: Thạc sỹ Địa chỉ cơ quan: Bộ môn Công nghệ Phần mềm Điện thoại cơ quan: 024 3752 1428 Di động: 0904656357 E-mail: phamvandong@humg.edu.vn Năm sinh: 1982			
7. DANH SÁCH THÀNH VIÊN THỰC HIỆN ĐỀ TÀI			
TT	Họ và tên	Đơn vị công tác và lĩnh vực chuyên môn	Chữ ký
1	Phạm Văn Đồng (Chủ nhiệm đề tài)	Bộ môn Công nghệ Phần mềm Công nghệ Thông tin	
2	Trần Thị Thu Thúy (Thành viên chính)	Bộ môn Mạng Máy Tính Công nghệ Thông tin	
3	Phạm Thị Hải Vân (Thành viên)	Bộ môn Công nghệ Phần mềm Công nghệ Thông tin	
4	Phạm Thị Nguyệt (Thành viên)	Bộ môn Tin học Kinh tế Công nghệ Thông tin	
5	Phạm Hữu Lợi (Thành viên)	Phòng thí nghiệm Địa tin học Công nghệ Thông tin	
8. ĐƠN VỊ PHỐI HỢP CHÍNH			
Tên đơn vị trong và ngoài nước	Nội dung phối hợp nghiên cứu	Họ và tên người đại diện đơn vị	

9. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU THUỘC LĨNH VỰC CỦA ĐỀ TÀI Ở TRONG VÀ NGOÀI NƯỚC

9.1. Ngoài nước

Trong những năm gần đây, sự phát triển của công nghệ thông tin, truyền thông, đặc biệt là các tiên bộ trong công nghệ tương tác người-máy thông qua tiếng nói đã mang lại nhiều lợi ích cho con người. Được quan tâm nghiên cứu và phát triển từ những năm 60 của thế kỷ trước, hiện nay, công nghệ xử lý ngôn ngữ tự nhiên, xử lý tiếng nói đã đạt được nhiều thành tựu đáng kể, với nhiều ứng dụng rộng rãi. Những ứng dụng sử dụng công nghệ này giúp con người có thể giao tiếp, tương tác với các hệ thống máy tính, robot thông qua ngôn ngữ tự nhiên của con người. Việc này giúp tăng năng suất lao động, hiệu quả học tập và mang lại nhiều tiện ích giá trị trong cuộc sống hàng ngày. Những ứng dụng chúng ta có thể dễ dàng bắt gặp hàng ngày như tổng đài điện thoại tự động sử dụng công nghệ nhận dạng tiếng nói, công nghệ tổng hợp tiếng nói, các phần mềm trợ lý ảo có thể trò chuyện với người dùng trên thiết bị di động (như *Siri*¹ trên iPhone, *Cortana*² trên Windows), các hệ thống thông minh điều khiển bằng tiếng nói trên xe ô tô (ví dụ hệ thống *Ford sync*³ trên các mẫu xe của Ford), hay điều khiển các thiết bị trong nhà thông qua các thiết bị thông minh (như *Amazon Echo*⁴, *Samsung smartThings*⁵), v.v... Sự phát triển của công nghệ cũng hỗ trợ giao tiếp giữa những người dùng có ngôn ngữ khác nhau thông qua các hệ thống dịch tự động hay còn gọi là dịch máy (machine translation). Trong các hệ thống như vậy thì thành phần tổng hợp tiếng nói là một hệ thống quan trọng tạo ra tiếng nói chất lượng cao cho người nghe.

Việc nghiên cứu về tổng hợp tiếng nói trên thế giới có nhiều kết quả khả quan và đã xây dựng thành công một số mô hình tổng hợp tiếng nói cho các ngôn ngữ Anh, Pháp, Đức, Trung Quốc...

Nghiên cứu cơ bản về tổng hợp tiếng nói điển hình có các nghiên cứu của giáo sư Thierry Dutoit [1], Benesty, Jacob, M. Mohan Sondhi, and Yiteng Huang với sổ tay xử lý tiếng nói [2], Paul Taylor [3].

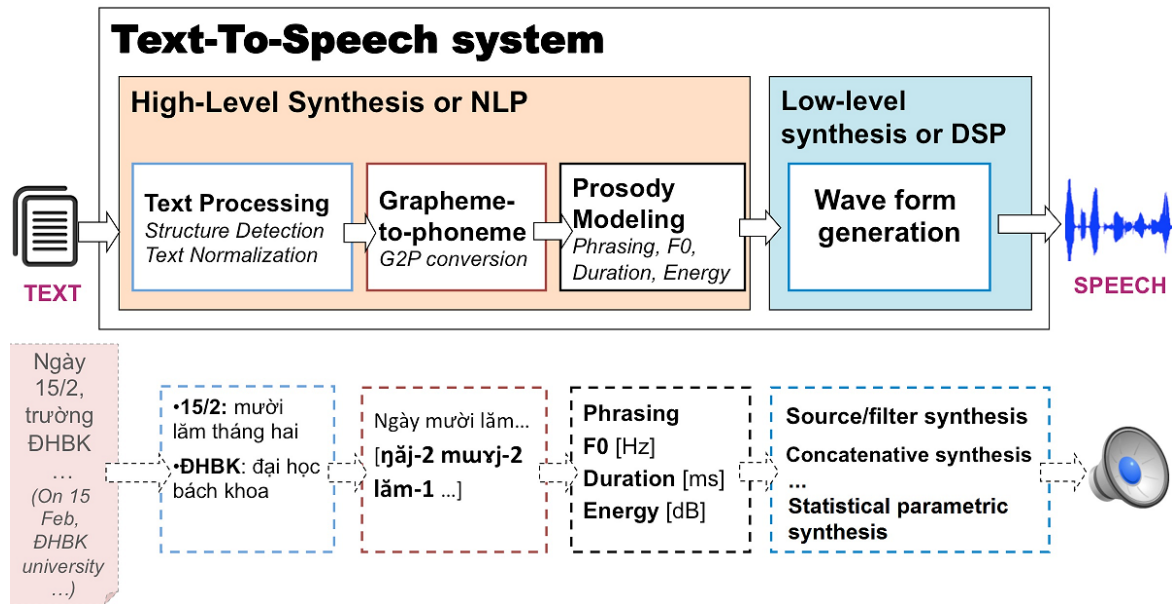
Việc nghiên cứu về ngôn ngữ nghèo tài nguyên trên thế giới còn rất nhiều hạn chế. Diễn hình có Hội thảo khoa học quốc tế về các ngôn ngữ nghèo tài nguyên SLTU'08, do Trung tâm Nghiên cứu Quốc tế MICA, Trường Đại học Bách Khoa Hà Nội và Phòng thí nghiệm tin học Grenoble, Viện Đại học Bách khoa Quốc gia Grenoble (Pháp) cùng phối hợp tổ chức, dưới sự tài trợ của Trung tâm Nghiên cứu khoa học Quốc gia Pháp (CNRS), Viện Đại học Bách khoa Quốc gia Grenoble (INPG), AFCEP và ISCA, diễn ra vào ngày 05/05/08 tại phòng hội thảo C2, trường ĐHBKHN. SLTU tiếp tục được tổ chức vào các năm chẵn, SLTU'10 ở Penang (Malaysia), SLTU'12 tại Cape Town (Nam Phi), SLTU'14 tại St Petersburg (Nga), SLTU'16 tại Yogyakarta, Indonesia ngày 09-12 tháng 5 năm 2016.

Giáo sư Harold Somers đã đưa ra nhiều công bố nghiên cứu về ngôn ngữ nghèo tài nguyên như tổng hợp văn bản thành giọng nói với ngôn ngữ nghèo tài nguyên [4], tự tổng hợp tiếng nói cho các bệnh nhân không nói tiếng Anh [5], thí nghiệm với Somali về phát triển tổng hợp tiếng nói cho ngôn ngữ nghèo tài nguyên [6].

Một hệ thống tổng hợp tiếng nói có thể coi là một hệ thống có thể tạo ra tiếng nói từ văn bản đầu vào (Text To Speech). Hệ thống thường gồm hai phần (xem *Hình 1*):

Mô đun tổng hợp mức cao (xử lý ngôn ngữ tự nhiên): có nhiệm vụ phân tích và chuyển đổi văn bản đầu vào thành chuỗi các tham số thích hợp. Mô đun này gồm các thành phần phân tích văn bản, phân tích ngữ âm, phân tích ngữ điệu. Mô đun tổng hợp mức cao thông thường phụ thuộc ngôn ngữ đầu vào.

Mô đun tổng hợp mức thấp (xử lý tín hiệu số): sẽ nhận đầu vào là các tham số thích hợp phân tích được từ mô đun mức cao, từ đó đưa vào thành phần xử lý tín hiệu số để tạo ra dạng sóng tín hiệu tiếng nói tương ứng.



Hình 1. Kiến trúc chung của một hệ thống tổng hợp tiếng nói từ văn bản [7] [8].

Được tập trung nghiên cứu phát triển từ những năm 60 của thế kỷ trước, lĩnh vực tổng hợp tiếng nói hay chuyển văn bản thành tiếng nói đã trải qua nhiều công nghệ và hướng tiếp cận khác nhau.

Có thể thấy hệ thống tổng hợp tiếng nói đã được nghiên cứu và phát triển từ lâu trên thế giới. Từ những năm đầu của thế kỷ 21, công nghệ tổng hợp tiếng nói đã có những bước phát triển vượt bậc.

Nhiều công cụ, nền tảng phần mềm được xây dựng và công bố rộng rãi, trong đó có nhiều công cụ mã nguồn mở. Đây là những công cụ rất hữu hiệu trong nghiên cứu và xây phát triển sản phẩm. Trong đó có thể kể đến các nền tảng phần mềm cho tổng hợp tiếng nói phổ biến gồm có: *HTS*⁶, *Festival*⁷, *Mary TTS*⁸. Các nền tảng này đều sử dụng phương pháp tổng hợp thống kê tham số và lựa chọn đơn vị âm. Một số nền tảng phần mềm mới ra đời

¹ <http://www.apple.com/ios/siri/>

² <https://www.microsoft.com/en/mobile/experiences/cortana/>

³ <http://www.ford.com/technology/sync/>

⁴ <http://amazon.com/echo>

⁵ <http://www.samsung.com/uk/smartthings/>

⁶ <http://hts.sp.nitech.ac.jp/>

⁷ <http://www.cstr.ed.ac.uk/projects/festival/>

⁸ <http://mary.dfki.de/>

gần đây như *Merlin*⁹ sử dụng phương pháp học sâu để xây dựng hệ thống tổng hợp tiếng nói.

9.2 Trong nước

a. Tổng quan về các nghiên cứu về xử lý tiếng nói, tổng hợp tiếng nói của một số đơn vị trong nước:

- Nhóm xử lý ngôn ngữ tự nhiên, Viện công nghệ thông tin, viện hàn lâm Việt Nam. Các bài báo và sản phẩm gần đây [10], [11] tập trung vào nghiên cứu và phát triển hệ thống tổng hợp tiếng nói tiếng Việt dựa trên phương pháp tổng hợp tiếng với các tham số thống kê (HMM based text to speech synthesis). Sản phẩm có sử dụng công nghệ tổng hợp tiếng nói tiếng Việt có thể kể đến là VnVoice 1.0 - Modul nhúng tổng hợp tiếng Việt

- Phòng thí nghiệm AiLab, ĐH KHTN Tp HCM, phó giáo sư Vũ Hải Quân đã đưa nghiên cứu của mình về xử lý ngôn ngữ tự nhiên trong đó có tổng hợp tiếng nói tiếng Việt bằng phương pháp ghép nối sử dụng cụm từ [12]. AiLab tập trung nghiên cứu về SPL cho tiếng Việt. Sản phẩm tiêu biểu là Speech Synthesis (Phiên bản 2.0). Đầu vào văn bản tiếng Việt sẽ được chuyển đổi thành tiếng nói sử dụng tổng hợp tiếng nói dựa trên tổng hợp đơn vị lựa chọn. Âm tiết được tự động phân đoạn bằng cách sử dụng nhận dạng tiếng nói và sau đó sử dụng như là đơn vị cơ bản trong hệ thống đường cơ sở. Hơn nữa, các đơn vị đã được mở rộng đến các cụm từ để tạo ra tiếng nói tự nhiên hơn.

- Phòng Giao tiếp tiếng nói, viện NCQT MICA. PGS TS. Trần Đỗ Đạt đã công bố hơn 44 nghiên cứu trong đó có nhiều nghiên cứu về TTS [13], [14], [15], [16], [17], [18]. Mục tiêu chính của đơn vị là thực hiện nghiên cứu cơ bản song song với phát triển ứng dụng. Đơn vị tập trung vào nghiên cứu về ngôn ngữ học, ngữ âm học và âm vị học của tiếng Việt và các ngôn ngữ khác ở Việt Nam và trong khu vực, để hiểu rõ hơn về cơ chế của sản xuất / nhận thức / hiểu biết, và để chuyển các kết quả cơ bản cho sự phát triển của công nghệ cao tại Việt Nam. Trong nhiều năm qua, đơn vị đã giải quyết đồng thời nghiên cứu cơ bản và nghiên cứu ứng dụng thông qua ngôn ngữ xử lý tự nhiên (NLP) và các công nghệ tiếng nói (Vocal Technologies). Tuy nhiên, do sự đa dạng và rộng lớn của lĩnh vực nghiên cứu nên phòng Giao tiếp tiếng nói Viện NCQT MICA không thực hiện nghiên cứu trong tất cả các chủ đề có thể trong NLP và xử lý tiếng nói (một lĩnh vực nghiên cứu rất lớn), mà tập trung vào các nghiên cứu cơ bản và công nghệ cho các bài toán phục vụ cho giao tiếp tiếng nói, đa phương ngôn ngữ (tiếng Việt, Khmer, Lào, Anh, Pháp), hữu ích cho tương tác người-máy hoặc máy tính hỗ trợ đối thoại giữa người với người.

b. Tình hình nghiên cứu về ngôn ngữ và xử lý ngôn ngữ dân tộc thiểu số ở Việt Nam

- Bộ môn Ngôn ngữ và Văn hoá các dân tộc thiểu số Việt Nam, Khoa ngôn ngữ học, Trường ĐH KHXH&NV, Đại học QGHN. GS Trần Trí Dõi đã đưa ra nghiên cứu và giảng dạy tiếng Việt và nhóm Việt - Mường (tiếng Việt thực hành, Ngữ âm tiếng Việt, Ngữ âm lịch sử và Lịch sử tiếng Việt, So sánh - lịch sử nhóm ngôn ngữ Việt - Mường); Nghiên cứu và giảng dạy ngôn ngữ - văn hoá các dân tộc thiểu số Việt Nam (Những vấn đề chung về ngôn ngữ - văn hoá các dân tộc thiểu số Việt Nam, nhóm Việt - Mường, ngôn ngữ và chữ viết cổ nhóm Tày - Thái); Nghiên cứu và giảng dạy chính sách ngôn ngữ - văn hoá, vấn đề

⁹ <http://www.cstr.ed.ac.uk/projects/merlin/>

giáo dục ngôn ngữ ở địa bàn các dân tộc thiểu số Việt Nam; Nghiên cứu và giảng dạy ngôn ngữ văn hoá và Địa danh học ở Việt Nam.

- Phòng giao tiếp tiếng nói, Viện MICA là một trong những đơn vị nghiên cứu trong nước tiến hành song song các nghiên cứu về ngôn ngữ và xử lý tiếng nói. Có thể tìm thấy nhiều nghiên cứu của Viện trong những năm gần đây về Ngữ âm và ngữ điệu tiếng Việt [19], tổng hợp tiếng nói [18], các ngôn ngữ thiểu số.

- Ngoài ra có thể tìm thấy một số nghiên cứu khác trong và ngoài nước về các ngôn ngữ Mon – Khmer [20], [21], [22]. Tạp chí Mon Khmer journal, tạp chí được thành lập từ năm 1964, đã đưa ra rất nhiều nghiên cứu về tiếng dân tộc. Hơn 50 năm qua, tạp chí đã cung cấp một diễn đàn học thuật cung cấp các bài báo chất lượng cao nghiên cứu về các ngôn ngữ và văn hóa Đông Nam Á.

c. Ngôn ngữ Mường và các nghiên cứu về tiếng Mường

Tại Việt Nam, dân tộc Mường là một trong 05 dân tộc có số dân đông nhất so với các dân tộc thiểu số khác (1.268.963 người theo kết quả điều tra dân số năm 2009 của Tổng cục thống kê). Cộng đồng dân tộc Mường sinh sống chủ yếu ở ba tỉnh Hòa Bình (người Mường chiếm 62% dân số toàn tỉnh), Phú Thọ (người Mường chiếm 13,1% dân số toàn tỉnh) và Thanh Hóa (chiếm 9,5%). Dân tộc Mường có nền văn hóa lâu đời, văn học dân gian rất phát triển. Những kho tàng truyện kể, các bài hát dân ca, và đặc biệt là các bài Mo Mường là những di sản văn hóa phi vật thể vô cùng quý giá cần được giữ gìn và phát huy.

Về mặt ngôn ngữ, tiếng Mường thuộc nhóm ngôn ngữ Việt Mường, nhánh Môn Khơ me, họ Nam Á. Tiếng Mường và tiếng Việt cùng một nguồn gốc, có quan hệ hết sức gần gũi và chặt chẽ với nhau. Vì vậy, tiếng Mường có nhiều đặc điểm giống với tiếng Việt. Những nghiên cứu về tiếng Mường hiện nay chủ yếu tập trung vào các vấn đề như: miêu tả đặc điểm ngữ âm các phương ngữ/thổ ngữ Mường (hệ thống âm đầu, âm chính, âm cuối, thanh điệu) ở các địa phương khác nhau [23] [24]; nghiên cứu lịch sử hình thành nguồn gốc tiếng Mường [25]–[27] ; nghiên cứu xây dựng chữ viết và phiên âm tiếng Mường [24], [28]–[31].

Bên cạnh những nghiên cứu cơ bản đã có về cấu trúc tiếng Mường thì những vấn đề nghiên cứu cần đặt ra đối với tiếng Mường hiện nay là cần tiếp tục có những nghiên cứu ứng dụng thành tựu của công nghệ thông tin trong xử lý tiếng nói tiếng Mường nhằm tạo ra những sản phẩm công nghệ tiện ích trong việc quảng bá và phổ biến tiếng Mường cũng như tạo cơ hội và môi trường cho người Mường học tiếng Việt nhằm nâng cao trình độ dân trí cho người Mường.

Nhìn chung, có khá nhiều nghiên cứu về ngôn ngữ dân tộc nói chung và tiếng Mường nói riêng, tuy nhiên việc ứng dụng các công nghệ về xử lý và giao tiếp tiếng nói cho các ngôn ngữ dân tộc tại Việt Nam còn khá ít ỏi và rời rạc, chưa có sự kết hợp giữa các nghiên cứu về ngôn ngữ và các kỹ thuật xử lý tiếng nói hiện đại.

9.3. Danh mục các công trình đã công bố liên quan đến đề tài **của chủ nhiệm và những thành viên** tham gia nghiên cứu

[1] Pham Van Dong, Mac Dang Khoa, Vu Thi Hai Ha, Tran Do Dat, *Rapid Development Of Text To Speech System For Unsupported Languages Using Faking Input Approach:*

Experiment With Muong, Proceedings of the 10th National Conference on Fundamental and Applied Information Technology Research (FAIR'10), ISBN: 978-604-913-614-6, 2017.

Tài liệu tham khảo

- [1] D. Thierry, *An introduction to text-to-speech synthesis*. Kluwer Academic Publishers, 1997.
- [2] J. Benesty, *Springer handbook of speech processing*. Springer Science & Business Media, 2008.
- [3] P. Taylor, “Text-To-Speech Synthesis”, *Camb. Univ. Press*, 2009.
- [4] H. Somers, “Faking it: Synthetic text-to-speech synthesis for under-resourced languages—Experimental design”, trong *Proceedings of the Australasian Language Technology Workshop*, 2005, tr 71–77.
- [5] H. Somers và H. Lovel, “Computer-based support for patients with limited English”, trong *Proceedings of the 7th International EAMT workshop on MT and other Language Technology Tools, Improving MT through other Language Technology Tools: Resources and Tools for Building MT*, 2003, tr 41–49.
- [6] H. Somers, G. Evans, và Z. Mohamed, “Developing speech synthesis for under-resourced languages by ‘faking it’: An experiment with Somali”, *Pap. Submitt. LREC*, 2006.
- [7] T. T. T. Nguyen, “HMM-based Vietnamese Text-To-Speech: Prosodic Phrasing Modeling, Corpus Design System Design, and Evaluation”, Paris 11, 2015.
- [8] T. T. T. Nguyen, “HMM-based Vietnamese Text-To-Speech: Prosodic Phrasing Modeling, Corpus Design System Design, and Evaluation”, Paris 11, 2015.
- [9] S. Lemmetty, “Review of Speech Synthesis Technology”, Helsinki University of Technology, 1999.
- [10] T. T. Vu, M. C. Luong, và S. Nakamura, “An HMM-based Vietnamese speech synthesis system”, trong *Speech Database and Assessments, 2009 Oriental COCODA International Conference on*, 2009, tr 116–121.
- [11] S. T. Phan, T. T. Vu, C. T. Duong, và M. C. Luong, “A study in vietnamese statistical parametric speech synthesis based on HMM”, *Int. J.*, vol 2, số p.h 1, tr 1–6, 2013.
- [12] V. H. Quân và C. X. Nam, “Tổng hợp tiếng nói tiếng Việt, theo phương pháp ghép nối cụm từ”, *Các Công Trình Nghiên Cứu Phát Triển Và Ứng Dụng CNTT-TT*, 2009.
- [13] T. Van Do, D.-D. Tran, và T.-T. T. Nguyen, “Non-uniform unit selection in Vietnamese speech synthesis”, trong *Proceedings of the Second Symposium on Information and Communication Technology*, 2011, tr 165–171.
- [14] N. T. T. Trang, C. D’ALESSANDRO, A. RILLIARD, và T. Do Dat, “HMM-based TTS for Hanoi Vietnamese: issues in design and evaluation”, 2013.
- [15] T. Do Dat, E. Castelli, J.-F. Serignat, T. Van Loan, và L. X. Hung, “Influence of F0 on Vietnamese syllable perception”, trong *9th European Conference on Speech Communication and Technology (Interspeech 2005)*, 2005, tr 1697–1700.
- [16] E. Castelli và others, “Generation of F0 contours for Vietnamese speech synthesis”, trong *Communications and Electronics (ICCE), 2010 Third International Conference on*, 2010, tr 158–162.
- [17] N. T. T. Trang, T. Do Dat, R. Albert, A. Christophe, và P. T. N. Yen, “Intonation issues in HMM-based speech synthesis for Vietnamese”, trong *The 4th International*

Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU'14), tr 98–104.

- [18] D.-K. Mac và D.-D. Tran, “Modeling Vietnamese Speech Prosody: A Step-by-Step Approach Towards an Expressive Speech Synthesis System”, trong *Trends and Applications in Knowledge Discovery and Data Mining*, Springer, 2015, tr 273–287.
- [19] T.-L. Nguyen, A. Michaud, D.-D. Tran, và D.-K. Mac, “The interplay of intonation and complex lexical tones: how speaker attitudes affect the realization of glottalization on Vietnamese sentence-final particles”, trong *Interspeech 2013*, 2013, tr the–Proceedings.
- [20] V. Kumar và c.s., “Asian and non-Asian origins of Mon-Khmer-and Mundari-speaking Austro-Asiatic populations of India”, *Am. J. Hum. Biol.*, vol 18, số p.h 4, tr 461–469, 2006.
- [21] P. Donegan, “Rhythm and vocalic drift in Munda and Mon-Khmer”, *Linguist. Tibeto-Burman Area*, vol 16, số p.h 1, tr 1–43, 1993.
- [22] H. L. Shorto, *Bibliographies of Mon-Khmer and Tai Linguistics*, vol 2. London: Oxford University Press, 1963.
- [23] Nguyễn Văn Tài, *Ngữ âm tiếng Mường qua các phương ngôn*. Hà Nội: Nxb Từ điển Bách khoa, 2005.
- [24] Nguyễn Văn Tài, *Ngữ âm tiếng Mường qua các phương ngôn*. Hà Nội: Nxb Từ điển Bách khoa, 2005.
- [25] Trần Trí Dõi, “Các ngôn ngữ thành phần của nhóm Việt - Mường”, *Ngôn Ngữ*, vol 3, 1996.
- [26] Trần Trí Dõi, *Một vài vấn đề nghiên cứu so sánh - lịch sử nhóm ngôn ngữ Việt Mường*. Hà Nội: Nhà xuất bản Đại học quốc gia Hà Nội, 2011.
- [27] D. John, “Mường is not a subgroup: Phonological evidence for a paraphyletic taxon in the Viet-Muong sub-family”, *Editor. Note*, tr 1, 2012.
- [28] Nguyễn Kim Thản, “Vài nét về hệ thống âm vị tiếng Mường và phương án phiên âm tiếng Mường”, *Ngôn Ngữ*, vol 1, 1971.
- [29] Milton Barker, “Bộ chữ Mường - Trao đổi khoa học về chữ Mường”. Tọa đàm Viện Ngôn ngữ học, 1994.
- [30] Nguyễn Như Ý, “Dự thảo phương án chữ Mường”. Tọa đàm Viện Ngôn ngữ học, 1994.
- [31] Hoàng Thị Châu, *Xây dựng bộ chữ phiên âm cho các dân tộc thiểu số ở Việt Nam*. Hà Nội: Nxb Văn hóa dân tộc, 2001.
- [32] Nguyễn Phan Cảnh, “Khảo sát về thanh điệu tiếng Mường (phương ngữ Mường Bi) trong các từ tách rời”, *Thông Báo Khoa Học*, vol 1, tr 36, 1962.
- [33] Nguyễn Minh Đức, “Một vài nét về các thổ ngữ của tiếng Mường Hòa Bình”, trong *Tìm hiểu ngôn ngữ các dân tộc thiểu số ở Việt Nam*, Hà Nội: Nhà xuất bản Khoa học xã hội, 1972.
- [34] Trần Trí Dõi, “Về quá trình hình thành thanh của một vài thổ ngữ/ ngôn ngữ Việt - Mường”, *Ngôn Ngữ*, vol 1, 1991.
- [35] Nguyễn Văn Khang, “Xây dựng bộ chữ Mường phục vụ cho việc bảo tồn, phát huy văn hóa Mường tại tỉnh Hòa Bình”. Đề tài cấp tỉnh, Sở khoa học và công nghệ tỉnh Hòa Bình - Viện Ngôn ngữ học, 2016.

10. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Ngày nay công nghệ xử lý tiếng nói đóng một vai trò quan trọng trong rất nhiều lĩnh vực của tương tác người máy. Trong thời gian gần đây, nhiều hệ thống tương tác bằng tiếng nói đã được giới thiệu, cho phép người sử dụng giao tiếp với thiết bị máy móc trên nhiều nền tảng khác nhau, như với điện thoại thông minh (Apple Siri, Google Now, Samsung S-Voice), xe hơi thông minh (BMW¹⁰, Ford¹¹, ...), nhà thông minh v.v.. Trong các hệ thống này thì chức năng tổng hợp tiếng nói, hay chuyển văn bản thành tiếng nói là một trong những chức năng cơ bản và quan trọng đảm bảo sự tương tác thuận lợi và thân thiện với người sử dụng.

Công nghệ tổng hợp tiếng nói hiện tại dựa trên kỹ thuật thống kê và học máy trên tập dữ liệu văn bản lớn. Vì vậy, các hệ thống tổng hợp tiếng nói trên hiện nay chỉ đang được áp dụng với các ngôn ngữ phổ biến và có chữ viết. Với các ngôn ngữ thiểu số như các ngôn ngữ dân tộc miền núi Việt Nam, với lượng dữ liệu hạn chế, thậm chí là không có dữ liệu văn bản (với các ngôn ngữ không có chữ viết) thì việc ứng dụng các hệ thống tổng hợp tiếng nói hiện tại là không thể. Việc xây dựng các hệ thống tổng hợp tiếng nói cho các ngôn ngữ thiểu số này yêu cầu các nghiên cứu về ngôn ngữ học để xây dựng dữ liệu và các nghiên cứu về công nghệ để thích nghi với các ngôn ngữ thiểu số. Việc nghiên cứu và phát triển công nghệ tổng hợp tiếng nói cho các ngôn ngữ ít nguồn tài nguyên là một trong những hướng nghiên cứu mới hiện nay trong lĩnh vực xử lý tiếng nói trên thế giới.

Tại Việt Nam, dân tộc Mường là một trong 05 dân tộc có số dân đông nhất so với các dân tộc thiểu số khác (trên một triệu người) và có nền văn hóa lâu đời, giàu bản sắc, có ngôn ngữ cùng cội nguồn với tiếng Việt, nhưng mãi đến ngày 8/9/2016 tiếng Mường ở Hòa Bình mới có bộ chữ viết chính thức. Cộng đồng dân tộc Mường sinh sống chủ yếu ở hai tỉnh Hòa Bình và Thanh Hóa, trong đó tại tỉnh Hòa Bình, dân tộc Mường chiếm đa số với 62% dân số toàn tỉnh. Ngoài ra, ngôn ngữ Mường là một trong những ngôn ngữ dân tộc với số lượng người khá đông đảo không chỉ ở Việt Nam mà còn trên nhiều quốc gia khác ở Châu Á.

Tháng 11/2015, Sở Khoa Học và Công Nghệ tỉnh Hòa Bình đã phê duyệt đề tài “Xây dựng bộ chữ Mường phục vụ cho việc bảo tồn, phát huy văn hóa Mường tại tỉnh Hòa Bình”, Đề tài do Viện Ngôn ngữ học, Viện Khoa học xã hội Việt Nam chủ trì, đã hoàn thành vào cuối năm 2016. Kết quả đề tài đang áp dụng ở tỉnh Hòa Bình sau đó sẽ đề xuất áp dụng trên địa bàn cả nước, tiến tới một bộ chữ thống nhất cho tiếng Mường.

Với việc bộ chữ viết tiếng Mường vừa được xây dựng tại tỉnh Hòa Bình sẽ mang đến các thuận lợi trong việc đánh giá kết quả nghiên cứu của đề tài. Cụ thể, bộ chữ viết tiếng Mường đang xây dựng có thể đưa vào thử nghiệm trong hệ thống tổng hợp tiếng nói tiếng Mường, từ đó so sánh kết quả đọc sử dụng bộ chữ này.

Do phụ thuộc chủ yếu vào dữ liệu, nên đối với các ngôn ngữ kém phổ biến, ít hay nghèo dữ liệu (low/under resourced languages) nói chung hay tiếng Mường nói riêng, việc xây dựng hệ thống tổng hợp tiếng nói phải đối mặt với nhiều vấn đề, trong đó khó khăn chính là việc thu thập và xây dựng cơ sở dữ liệu song song kích thước đủ lớn để huấn luyện hệ thống. Do đó, các phương pháp, công cụ phần mềm hiện tại không thể sử dụng trực tiếp với

các ngôn ngữ này. Có thể nói đây là một vấn đề mới và khó khăn trong lĩnh vực tổng hợp tiếng nói, đòi hỏi phải có nghiên cứu về phương pháp mới để giải quyết vấn đề tổng hợp tiếng nói với tiếng Mường.

Trong nhiều năm qua, các nhà nghiên cứu đã tiến hành hàng loạt các nghiên cứu về mặt ngôn ngữ cho các dân tộc thiểu số ở Việt Nam, với nhiều công trình được xuất bản dưới dạng sách, tạp chí và báo cáo khoa học. Tuy nhiên, hầu hết các nghiên cứu mới chỉ được tiến hành trên phương diện nghiên cứu cơ bản về mặt ngôn ngữ (cấu trúc ngôn ngữ, xác định quan hệ cội nguồn, loại hình,...) mà hầu như chưa có các nghiên cứu ứng dụng công nghệ đối với các thứ tiếng này. Do đó, hướng nghiên cứu xây dựng hệ tổng hợp tiếng nói tiếng Mường có thể coi là một hướng mới và có nhiều cơ sở thuận lợi để thực hiện.

Do yêu cầu của đề tài đặt ra nhiều vấn đề mới trong nghiên cứu cả về công nghệ và ngôn ngữ, nên trong thời gian một năm, đề tài chỉ tập trung vào mục tiêu tổng hợp tiếng nói tiếng Mường cho hai phương ngữ Mường Bi - Hòa Bình và Mường Tân Sơn - Phú Thọ.

11. MỤC TIÊU ĐỀ TÀI

Xây dựng Module tổng hợp tiếng nói tiếng Mường cho hai phương ngữ Mường Bi – Hòa Bình và Mường Tân Sơn – Phú Thọ.

12. ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU

12.1. Đối tượng nghiên cứu

- Ngôn ngữ tiếng Mường
- Hệ thống tổng hợp tiếng nói

12.2. Phạm vi nghiên cứu

- Ngôn ngữ tiếng Mường: 02 phương ngữ ở Mường Bi Hòa Bình và Tân Sơn Phú Thọ, hệ thống ngôn ngữ có 5 thanh điệu
- Hệ thống tổng hợp tiếng nói: Tổng hợp ghép nối, tổng hợp theo tham số thống kê, hệ thống MarryTTS

13. CÁCH TIẾP CẬN, PHƯƠNG PHÁP NGHIÊN CỨU

13.1. Cách tiếp cận

Trên cơ sở nghiên cứu về ngôn ngữ tiếng Mường, xét độ tương đồng về phiên âm IPA quốc tế, đề xuất phương pháp ánh xạ ngữ âm giữa tiếng Việt và tiếng Mường để xây dựng Module tổng hợp tiếng nói tiếng Mường. Chứng minh tính đúng đắn của phương pháp thông qua đánh giá cảm thụ và đánh giá chấp nhận được (MOST Test và Intelligibility Test).

13.2. Phương pháp nghiên cứu

Để đạt được mục tiêu nghiên cứu đề tài sử dụng các phương pháp nghiên cứu sau:

- Nghiên cứu lý thuyết: Thu thập tài liệu, phân tích, đánh giá, đề xuất và chứng minh.
- Nghiên cứu thực tiễn: Khảo sát ngôn ngữ tiếng Mường, xây dựng Module tổng hợp tiếng nói cho tiếng Mường và đánh giá chất lượng Module

13.3. Phương tiện kỹ thuật phục vụ nghiên cứu

14. NỘI DUNG NGHIÊN CỨU VÀ TIẾN ĐỘ THỰC HIỆN

¹⁰ http://www.bmw.com/com/en/insights/technology/technology_guide/articles/voice_control_system.html

¹¹ <http://www.ford.co.uk/experience-ford/Technology/Bluetooth-Voice-Control>

14.1. Nội dung nghiên cứu

- **Nội dung 1** Nghiên cứu phương pháp tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi – Hòa Bình và Mường Tân Sơn – Phú Thọ
- **Nội dung 2** Phát triển bộ công cụ thu thập và xử lý dữ liệu tiếng nói tiếng Mường
- **Nội dung 3** Xây dựng cơ sở dữ liệu tiếng Mường ở Mường Bi và Tân Sơn
- **Nội dung 4** Xây dựng Module tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi và Tân Sơn

14.2. Tiến độ thực hiện

TT	Nội dung NC	Sản phẩm đạt được	Thời gian Thực hiện	Người thực hiện (chính -tham gia)
1	Nội dung 1: Nghiên cứu phương pháp tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi – Hòa Bình và Mường Tân Sơn – Phú Thọ	Bài báo khoa học về phương pháp tổng hợp tiếng nói cho tiếng Mường	11/2017 – 12/2017	Phạm Văn Đồng, Trần Thị Thu Thúy, Phạm Thị Hải Vân, Phạm Thị Nguyệt, Phạm Hữu Lợi
2	Nội dung 2: Phát triển bộ công cụ thu thập và xử lý dữ liệu tiếng nói tiếng Mường	Bộ công cụ thu thập và xử lý dữ liệu tiếng nói	1/2018 – 3/2018	Phạm Văn Đồng, Trần Thị Thu Thúy, Phạm Thị Hải Vân
3	Nội dung 3: Xây dựng cơ sở dữ liệu tiếng Mường ở Mường Bi và Tân Sơn	Cơ sở dữ liệu tiếng Mường	4/2018 – 6/2018	Phạm Văn Đồng, Trần Thị Thu Thúy, Phạm Thị Hải Vân, Phạm Thị Nguyệt, Phạm Hữu Lợi
4	Nội dung 4: Xây dựng Module tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi và Tân Sơn	Module tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi và Tân Sơn	7/2018 – 9/2018	Phạm Văn Đồng, Trần Thị Thu Thúy, Phạm Thị Hải Vân
5	Tổng kết, đánh giá		10/2018	Phạm Văn Đồng, Trần Thị Thu Thúy, Phạm Thị Hải Vân, Phạm Thị Nguyệt, Phạm Hữu Lợi

15. SẢN PHẨM

15.1 Sản phẩm dạng khoa học

Sách chuyên khảo ☐

Sách tham khảo ☐

Giáo trình ☐

Bài báo đăng tạp chí nước ngoài ☐

Bài báo đăng tạp chí trong nước ☒

Bài đăng kỷ yếu hội nghị, hội thảo quốc tế ☐

15.2 Sản phẩm dạng đào tạo

Nghiên cứu sinh ☒

Cao học ☐

Đại học ☒

15.3 Sản phẩm dạng ứng dụng

☐ Mẫu, mô hình

☐ Phần mềm máy tính

☐ Tiêu chuẩn, bộ tiêu chí

☐ Tài liệu dự báo, đánh giá

☐ Phương pháp, thuật toán

☒ CT máy tính

☐ Vật liệu

☐ Dây chuyền công nghệ

☐ Qui phạm, đề xuất

☐ Đề án, dự án

☐ Báo cáo phân tích

☐ Bản quy hoạch

☐ Thiết bị máy móc

☐ Quy trình công nghệ

☐ Sơ đồ, bản đồ

☐ Luận chứng kinh tế

☐ Bản kiến nghị

☒ Sản phẩm khác (*)

15.4 Các sản phẩm khác

15.5 Tên sản phẩm, số lượng và yêu cầu khoa học đối với sản phẩm

TT	Tên sản phẩm	Số lượng	Yêu cầu khoa học, thông số kỹ thuật, chỉ tiêu cụ thể
1	Bài báo khoa học	01	Bài báo đăng tạp chí trong nước
2	Module xử lý dữ liệu, Module tổng hợp tiếng nói tiếng Mường	02	
3	Báo cáo tổng kết đề tài	01	Theo quy định
4	Đào tạo NCS	01	Hỗ trợ tài liệu trong quá trình đào tạo NCS của chủ nhiệm đề tài

16. HIỆU QUẢ

+ Phục vụ cho công tác giáo dục và đào tạo

- Kết quả nghiên cứu của đề tài được dùng làm tài liệu giảng dạy/tham khảo cho các môn học *Trí tuệ nhân tạo*, *Công nghệ phần mềm* và *Mã nguồn mở*.
- Góp phần đào tạo 01 nghiên cứu sinh chuyên ngành Khoa học máy tính, 01 sinh viên chuyên ngành Công nghệ phần mềm.

+ Phục vụ phát triển kinh tế - xã hội

Góp phần bảo tồn, phát huy ngôn ngữ tiếng Mường.

+ Phục vụ an ninh - quốc phòng

17. PHƯƠNG THỨC CHUYỂN GIAO KẾT QUẢ NGHIÊN CỨU VÀ ĐỊA CHỈ ỨNG DỤNG

- Phương thức chuyển giao kết quả NC

- Làm làm tài liệu tham khảo cho các môn học *Trí tuệ nhân tạo, Công nghệ phần mềm và Mã nguồn mở*.
- Làm công cụ hỗ trợ cho việc thu thập và xử lý dữ liệu tiếng nói tiếng Mường, công cụ tổng hợp tiếng nói tiếng Mường.

- Tên và địa chỉ đơn vị sử dụng kết quả NC

- Khoa Công nghệ Thông tin, Trường Đại học Mỏ - Địa chất.

18. KINH PHÍ THỰC HIỆN ĐỀ TÀI VÀ NGUỒN KINH PHÍ

18.1 Tổng kinh phí: 25.000.000 Trong đó:

Hỗ trợ từ Nguồn thu của Nhà trường (đ): 25.000.000 (Bằng chữ) Hai mươi lăm triệu đồng chẵn

Các nguồn kinh phí khác: 0 (đ)..... (Bằng chữ) Không

18.2 Dự trù kinh phí theo các mục chi: (Đơn vị tính: đồng)

TT	Các khoản chi,	Kinh phí thực hiện		
		Tổng số	Trong đó:	
			Nguồn thu của Nhà trường	Nguồn khác
I	Chi công lao động tham gia trực tiếp thực hiện đề tài	19.526.000	19.526.000	
II	Chi mua vật tư, nguyên, nhiên, vật liệu, tài liệu, tư liệu, ...			
III	Chi sửa chữa, mua sắm thiết bị phục vụ nghiên cứu			
IV	Chi khác			
1	Công tác phí			
2	Hội nghị, hội thảo			
3	Văn phòng phẩm, in ấn, tài liệu	2.224.000	2.224.000	
4	Nghiệm thu cấp Trường	2.000.000	2.000.000	
5	Chi phí quản lý chung	1.250.000	1.250.000	
Tổng cộng		25.000.000	25.000.000	

Ngày... tháng... năm 2017

HIỆU TRƯỞNG
(Kí tên, đóng dấu)

Ngày...tháng...năm 2017

Phó Trưởng Bộ môn
(ký, họ và tên)

Ngày 23 tháng 11 năm 2017

Chủ nhiệm đề tài
(ký, họ và tên)

Hoàng Anh Đức

Phạm Văn Đồng

DỰ TOÁN CHI TIẾT

Tên đề tài: Xây dựng Module tổng hợp tiếng nói tiếng Mường cho hai phương ngữ Mường Bi – Hòa Bình và Mường Tân Sơn – Phú Thọ

A. Cơ sở dự toán:

1. Quyết định số 1804/QĐ-MĐC ngày 31/12/2015 của Hiệu trưởng Trường Đại học Mỏ - Địa chất quy định định mức xây dựng, phân bổ dự toán và quyết toán kinh phí áp dụng đối với nhiệm vụ KHCN cấp Cơ sở của Trường Đại học Mỏ - Địa chất;

I. Công lao động thực hiện đề tài

T T	Nội dung	Sản phẩm	Cán bộ thực hiện	Hệ số tiền công Hstcn	Số ngày thực hiện Snc	Thành tiền (đ) $\Sigma = 1.300.000 * Hstcn * Snc$		
						Tổng	Trong đó	
							Nhà trường	Khác
1	<u>Nội dung 1:</u> Nghiên cứu đề xuất phương pháp tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi – Hòa Bình và Mường Tân Sơn – Phú Thọ	Sản phẩm 1: Bài báo khoa học về phương pháp tổng hợp tiếng nói cho tiếng Mường	Phạm Văn Đồng	0.42	6	3.276.000	3.276.000	
			Trần Thị Thu Thúy	0.30	3	1.170.000	1.170.000	
			Phạm Hữu Lợi	0.16	3	624.000	624.000	
2	<u>Nội dung 2:</u> Phát triển bộ công cụ thu thập và xử lý dữ liệu tiếng nói tiếng Mường	Sản phẩm 2: Công cụ thu thập và xử lý dữ liệu tiếng nói	Phạm Văn Đồng	0.42	3	1.638.000	1.638.000	
			Trần Thị Thu Thúy	0.30	2	780.000	780.000	
			Phạm Hữu Lợi	0.16	1	208.000	208.000	
3	<u>Nội dung 3:</u> Xây dựng cơ sở dữ liệu tiếng Mường ở Mường Bi và Tân Sơn	Sản phẩm 3: Cơ sở dữ liệu tiếng Mường	Phạm Văn Đồng	0.42	4	2.184.000	2.184.000	
			Trần Thị Thu Thúy	0.30	2	780.000	780.000	
			Phạm Hữu Lợi	0.16	2	416.000	416.000	

4	<u>Nội dung 4:</u> Xây dựng Module tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi và Tân Sơn	Sản phẩm 4: Module tổng hợp tiếng nói cho tiếng Mường cho hai phương ngữ Mường Bi và Tân Sơn	Phạm Văn Đồng	0.42	8	4.368.000	4.368.000	
			Trần Thị Thu Thúy	0.30	4	1.560.000	1.560.000	
			Phạm Thị Nguyệt					
				0.16	4	832.000	832.000	
5	<u>Nội dung 5:</u> Tổng kết, đánh giá	Báo cáo tổng kết đề tài	Phạm Văn Đồng	0.42	2	1.092.000	1.092.000	
			Trần Thị Thu Thúy	0.30	1	390.000	390.000	
			Phạm Thị Hải Vân	0.16	1	208.000	208.000	
Cộng						19.526.000	19.526.000	

II. Chi mua vật tư, nguyên nhiên vật liệu

T T	Nội dung chi	Đơn vị tính	Đơn giá (đ)	Số lượng	Thành tiền (đ)		
					Tổng số	Trong đó	
						Nhà trường	khác
1							
2							

III. Chi sửa chữa, mua sắm thiết bị phục vụ nghiên cứu

T T	Nội dung chi	Đơn vị tính	Đơn giá (đ)	Số lượng	Thành tiền (đ)		
					Tổng số	Trong đó	
						Nhà trường	khác
1							
2							

IV. Chi khác

T T	Nội dung chi	Đơn vị tính	Đơn giá (đ)	Số lượng	Thành tiền (đ)		
					Tổng số	Trong đó	
						Nhà trường	Khác
1	Công tác phí						
	<i>Tiền tàu, xe vận chuyển</i>						
	<i>Tiền phòng nghỉ</i>						
	<i>Phụ cấp lưu trú</i>						
2	Hội nghị, hội thảo khoa học						
	<i>Người chủ trì</i>						
	<i>Thư ký hội thảo</i>						
	<i>Báo cáo khoa học</i>						
	<i>Thành viên tham gia hội thảo</i>						
3	Văn phòng phẩm, in ấn, dịch tài liệu				2.224.000	2.224.000	
4	Hội đồng nghiệm thu cấp cơ sở				2.000.000	2.000.000	
5	Chi phí quản lý chung nhiệm vụ				1.250.000	1.250.000	
	Cộng				5.474.000	5.474.000	

Tổng 25.000.000 đ

(số tiền bằng chữ: Hai mươi lăm triệu đồng chẵn)

**HIỆU TRƯỞNG
TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT**

CHỦ NHIỆM ĐỀ TÀI

Phạm Văn Đồng