



MINISTÈRE DE L'ÉDUCATION  
NATIONALE, DE L'ENSEIGNEMENT  
SUPÉRIEUR ET DE LA RECHERCHE

>>>

## AVIS DE SOUTENANCE DE THESE DE DOCTORAT

**Monsieur Hong Quang NGUYEN** soutiendra une thèse  
**05 novembre 2008 à 14h 30**

**L.I.A.**

**Amphi B. Pascal**

**SPÉCIALITÉ : INFORMATIQUE ED 166**

**Titre de la thèse : Reconnaissance automatique de la parole continue grand vocabulaire en vietnamien**

**Membres du jury :**

HIRST Daniel, DR Informatique, Université de Provence,  
BONASTRE Jean-François, PR Informatique, Université d'Avignon et des Pays de Vaucluse  
BESACIER Laurent, HDR Informatique, Université Joseph Fourier, Grenoble,  
LAMEL Lori, DR2 Informatique, CNRS Orsay,  
CASTELLI Eric, HDR Informatique, IP Hanoi - INP Grenoble,  
NOCERA Pascal, MCF Informatique, Université d'Avignon et des Pays de Vaucluse

**Résumé de la thèse :**

La reconnaissance de la parole vietnamienne est juste au début de son développement. Nous constatons que les différences entre la langue vietnamienne et les langues occidentales sont telles que les techniques de reconnaissance de la parole communément employées pour ces dernières (anglais, français par exemple) ne suffisent pas pour développer directement un système de reconnaissance performant. La prise en compte des caractéristiques de la langue vietnamienne au niveau de la représentation des données (lexique, modèle de langage) et des modèles (modèle de tons) permettent par contre d'obtenir des résultats prometteurs.

La première différence est la segmentation des entités sémantiques des phrases. En vietnamien, les mots/concepts peuvent être composés d'une ou plusieurs syllabes qui sont systématiquement découpés en syllabes séparées par un espace (langue syllabique). La segmentation en mots/concepts de la phrase est une tâche importante pour les langues isolantes telles que le mandarin, le cantonais, le thaï mais aussi pour le vietnamien. Pour améliorer les résultats des traitements automatiques de ces langues, nous avons construit un module de segmentation en mots multi syllabiques des phrases syllabiques. Deux approches ont été utilisées pour cela : la première utilise un dictionnaire de mots vietnamiens multi syllabiques alors que la seconde construit automatiquement un lexique multi syllabique à l'aide d'un algorithme utilisant l'information mutuelle des mots comme critère de regroupement, et la programmation dynamique pour simplifier les traitements.

La deuxième différence entre ces langues est l'importance du ton dans la langue vietnamienne. La reconnaissance des tons est donc un aspect fondamental du traitement des langues tonales. Dans cette thèse, nous avons étudié en détail différentes méthodes pour représenter de manière optimale la fréquence fondamentale et l'énergie, d'une part, et pour trouver un moyen d'atténuer l'influence du phénomène de coarticulation entre les tons. Nous avons utilisé deux approches pour effectuer cette reconnaissance : une approche trame à trame à l'aide des modèles de Markov caché et une méthode globale à l'aide d'un perceptron multicouche.

En cumulant les traitements des caractéristiques linguistiques (lexique multi syllabique) et acoustiques (reconnaissance des tons), les résultats ont été améliorés de pratiquement 50 % (par rapport au système initial). Ces résultats prouvent que l'ajout d'informations supplémentaires, caractéristiques de la langue vietnamienne, améliore considérablement les performances des systèmes de reconnaissance de la parole.