

**AVIS DE SOUTENANCE DE THÈSE DE  
DOCTORAT**

**Monsieur PEILLON Stéphane** soutiendra une thèse  
**Le Jeudi 28 Novembre 2002 à 11 heures**

**Amphi ADA LOVELACE**

**IUP GMI - Université d'Avignon et des Pays de Vaucluse**

**SPÉCIALITÉ : Informatique**

Titre de la thèse : Indexation vocale à vocabulaire illimité à base de décodage phonétique. Application à la détection de clés vocales dans un flux de parole.

Membres du jury :

**Mme Régine ANDRÉ-OBRECHT**, Professeur, IRIT, Université Paul Sabatier, Toulouse.

**M. Pietro LAFACE**, Professeur, DIP Informatica, Politecnico di Torino, Turin.

**M. Denis JOUVET**, Ingénieur, France Télécom R&D, LANNION.

**M. Alexandre FERRIEUX**, Ingénieur, France Télécom R&D, LANNION.

**M. Renato DE MORI**, Professeur, Laboratoire d'Informatique (EA 931 FRE 2487), Université d'Avignon et des Pays de Vaucluse.

**M. Marc EL BEZE**, Professeur, Laboratoire d'Informatique (EA 931 FRE 2487), Université d'Avignon et des Pays de Vaucluse.

Résumé de la thèse :

Le stockage croissant de données multimédia est aujourd'hui confronté au manque d'outils efficaces d'extraction et de tri de documents. Dans le cadre particulier de données vocales, nous proposons une technique d'indexation permettant de retrouver des documents sonores à partir de leur seul contenu. La pose d'index appropriés sur le médium permet de réduire considérablement l'information nécessaire a posteriori pour la phase de recherche d'une clé.

Nous comparons deux modes d'indexation utilisant tous deux des index phonétiques : l'un repose sur la meilleure séquence possible de phonèmes, l'autre sur des échelles d'hypothèses phonétiques produites sur une segmentation automatique a priori du corpus. Ce deuxième mode que nous désignons sous l'appellation "treillis synchrone de phonèmes", offre de meilleures performances avec un faible surcoût de calcul, et présente un apprentissage réduit des paramètres du moteur de recherche. La technique présentée permet en outre la détection de clés vocales, soit dans un corpus vocal, soit dans un corpus textuel. La clé est alors elle-même indexée.