



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

>>>

## AVIS DE SOUTENANCE DE THESE DE DOCTORAT

**Monsieur Gilles POUCHOULIN soutiendra une thèse  
le 19 décembre 2008 à 14h 30**

**LIA**

SPÉCIALITÉ : INFORMATIQUE ED 166

Titre de la thèse : Approche Statistique pour l'Analyse Objective et la Caractérisation de la Voix Dysphonique

Membres du jury :

ADDA-DECKER Martine, CR-HDR Informatique, LIMSI-CNRS  
GRENEZ Francis, PR Traitement du signal, Université Libre de Bruxelles, Belgique  
CREVIER-BUCHMAN Lise, CR O.R.L. Phoniatrie, CNRS/Sorbonne-Nouvelle, Paris,  
GIOVANNI Antoine, PR O.R.L. CHU La Timone, Marseille,  
GHIO Alain, IR Traitement du signal, Université de Provence, Marseille,  
MELONI Henri, PR Informatique, Université d'Avignon et des Pays de Vaucluse,  
BONASTRE Jean-François, PR Informatique, Université d'Avignon et des Pays de Vaucluse,  
FREDOUILLE Corinne, MCF Informatique, Université d'Avignon et des Pays de Vaucluse.

Résumé de la thèse : Dans notre société où la communication verbale est essentielle, l'évaluation de la qualité de la voix pathologique et des causes de sa dégradation occupent une place de plus en plus importante pour le corps médical. Une personne atteinte d'un trouble vocal momentané ou durable comme une dysphonie peut subir de graves conséquences dans ses relations sociales, tant sur le plan professionnel que personnel : manque d'efficacité dans la communication, arrêt ou perte du travail, exclusion sociale, voire perte identitaire. Cependant, face aux limites du jugement à l'oreille (analyse perceptive) du dysfonctionnement vocal, les thérapeutes de la voix ressentent le besoin de plus en plus pressant d'une méthode d'évaluation objective de la qualité de la voix pathologique, complémentaire à l'analyse perceptive.

Cette thèse s'inscrit dans ce cadre et plus précisément dans l'adaptation des techniques de Reconnaissance Automatique du Locuteur (RAL) à la classification automatique des voix dysphoniques suivant le grade global de l'échelle perceptive GRBAS. Toutes les études sont conduites sur un corpus de 80 voix dysphoniques (dont 20 voix de contrôle) fourni par le département ENT du Centre Hospitalier et Universitaire de La Timone (Marseille). L'objectif des travaux présentés ici est d'acquérir une meilleure compréhension des phénomènes acoustiques liés à la dysphonie.

L'originalité d'une telle approche réside dans l'utilisation d'un système de classification automatique comme outil de caractérisation des phénomènes pathologiques dans le signal de parole afin d'apporter aux experts humains de nouvelles connaissances sur les altérations de la voix. En ce sens, les spécialistes de la voix comme les phonéticiens pourront valider et/ou enrichir ces nouvelles connaissances en les approfondissant. Le cas échéant, les experts pourront en retour suggérer des indications/directives permettant au système automatique d'explorer de nouvelles pistes d'investigation. Cette démarche se distingue des méthodologies proposées dans la littérature qui visent davantage à améliorer les performances du système pour la tâche visée.

Les travaux réalisés dans cette thèse se subdivisent en deux volets : un premier volet qui décrit le système automatique adapté au contexte pathologique et un deuxième volet qui s'intéresse à la recherche de l'information pertinente. Dans cette optique, trois axes de recherche sont proposés.

Le premier axe est consacré à l'étude de différentes représentations paramétriques du signal de parole utilisées classiquement en RAL et appliquées ici dans un contexte pathologique. Les analyses spectrale, cepstrale et prédictive sont comparées ainsi que la complémentarité des coefficients statiques et dynamiques (dérivées 1<sup>ière</sup>, 2<sup>ième</sup> et 3<sup>ième</sup>) avérée utile en RAL. Cette étude a montré l'intérêt de l'analyse spectrale dans notre contexte expérimental, ainsi que celui des informations dynamiques.

Le deuxième axe de recherche étudie la manière dont les caractéristiques acoustiques de la dysphonie sont dispersées sur l'ensemble de l'espace fréquentiel. Cette étude a montré la pertinence de la bande de fréquences [0-3000]Hz. Partant de ce résultat, une évaluation perceptive des signaux de parole du corpus des voix dysphoniques filtrés en [0-3000]Hz a été proposée ainsi qu'un parallèle avec la bande téléphonique.

Dans le dernier axe de cette thèse, les manifestations de la dysphonie sont étudiées en observant le comportement du système de classification par phonème ou classe de phonèmes. La principale observation concerne la pertinence de la classe des consonnes sur les deux bandes fréquentielles [0-8000]Hz et [0-3000]Hz. Le comportement « peu attendu » des consonnes et plus particulièrement des consonnes sourdes (vis à vis du type de pathologie étudié) permet ici au système automatique de remplir pleinement son rôle d'outil caractérisant les phénomènes pathologiques. En effet, l'analyse du comportement du système a permis de mettre en évidence des phénomènes (comme par exemple le VOT) qui nécessitent à présent une expertise phonétique et clinique approfondie.