

**PROGRAMME THEMATIQUE EN SCIENCES HUMAINES ET SOCIALES**

**CORPUS ET OUTILS DE LA RECHERCHE EN SCIENCES  
HUMAINES ET SOCIALES**

**APPEL A PROJETS 2006**

Destiné aux différentes disciplines des sciences humaines et sociales

**FORMULAIRE**

Date limite d'envoi des dossiers : **12 mai 2006**

**Les dossiers (si possible au format PDF) doivent être envoyés par courrier électronique à**  
corpus-anr@ens-lsh.fr

**et**

**par courrier postal en 4 exemplaires (un original et 3 copies) cachet de la poste faisant foi**

ENS LSH  
Programme ANR Corpus et outils de la recherche en sciences humaines et sociales  
15 Parvis René Descartes BP 7000  
69342 Lyon cedex 07

RENSEIGNEMENTS ADMINISTRATIFS  
corpus-anr@ens-lsh.fr

RENSEIGNEMENTS SCIENTIFIQUES  
pierre-olivier.pin@agencerecherche.fr

FORMULAIRE DE SOUMISSION ET RENSEIGNEMENTS  
<http://www.agence-nationale-recherche.fr/>

## I - FICHE D'IDENTITE DU PROJET

**Titre du projet** (*maximum 120 caractères*)

***Données Cinéradiographiques Valorisées et recherches sur la Coarticulation, l'Inversion et l'évaluation de Modèles Physiques***

**Acronyme ou titre court** (*12 caractères maxi*)

*DOCVACIM*

**Mots-clés**

***cinéradiographie, multimédia, production-perception, parole, coarticulation, inversion, modélisation, valorisation***

**Résumé du projet** (*maximum 5000 caractères*)

Ce projet de recherche, en production-perception de la parole, a pour objectif général de contribuer à la compréhension et à la modélisation du contrôle des gestes de la parole, par l'analyse de la *coarticulation*, l'étude de l'*inversion* et l'*évaluation de modèles physiques*.

La production de la parole nécessite une coordination précise des articulateurs dans un espace donné ; le chevauchement gestuel ou la *coarticulation* se fait selon des contraintes spécifiques aux sons, aux langues et aux locuteurs. Cet aspect des recherches sur l'organisation spatio-temporelle des gestes linguistiques est donc central aux travaux en production-perception de la parole (cf. axe de recherche 1).

Appréhender les dimensions perceptives – auditive et visuelle – de la parole requiert une bonne maîtrise de la relation non-linéaire (mal-posée) entre la sortie acoustique et l'origine articulaire des sons. L'*inversion* acoustique articulaire consiste ainsi à retrouver l'évolution de la forme du conduit vocal à partir du signal de parole (cf. axe de recherche 2).

La compréhension des mécanismes de contrôle de la production de la parole passe, selon nous, par le développement et l'*évaluation de modèles* du contrôle moteur et de modèles physiques de l'appareil de production de la parole. C'est la comparaison de données mesurées sur des locuteurs, dans des conditions bien contrôlées avec les prédictions proposées par ces modèles, qui permet d'évaluer, de manière quantitative et objective, les différentes hypothèses attestées dans la littérature sur le contrôle de la production de la parole (cf. axe de recherche 3).

Pour étudier le contrôle moteur des gestes humains en général, ainsi que leur efficacité perceptive, le substrat de base est constitué d'une quantité très importante de données physiologiques, cinématiques, acoustiques et éventuellement dynamiques, collectées sur des locuteurs humains. Cette étape de collecte et d'analyse de données est indispensable ; elle a constamment progressé au cours des vingt dernières années.

Malgré ces avancées, le *manque de données cinéradiographiques* se fait cruellement ressentir dans la communauté scientifique, préoccupée par l'étude de la production et de la perception de la parole. En effet, elles sont *les seules* à offrir actuellement, en même temps, une résolution spatio-

temporelle correcte, sur l'ensemble du conduit vocal, dans le plan sagittal.

Or, l'acquisition de nouvelles données cinéroradiographiques est aujourd'hui restreinte, voire interdite dans certains pays.

Notre projet de recherche sur le contrôle moteur des gestes humains en général, ainsi que sur leur efficacité perceptive vise, en conséquence, à valoriser et à diffuser une partie des nombreux films cinéroradiographiques (ou radiofilms) sur la production et la perception de la parole qui ont été réalisés à l'Institut de Phonétique de Strasbourg (IPS), depuis la fin des années '50. Le programme concerne une vingtaine de films de très bonne qualité, portant sur des problématiques linguistiques de langues parlées en Europe, en Afrique, en Asie et en Amérique Latine.

Il est organisé autour de 4 objectifs principaux, trois spécifiques et un général qui correspond aux axes de recherche:

/1/ la *sauvegarde* de ces données cinéroradiographiques patrimoniales ;

/2/ leur *traitement* (grâce au développement d'outils) ;

/3/ leur *valorisation* par l'intégration dans une base, l'organisation d'ateliers et de colloques ;

/4/ et leur *analyse*, ainsi que leur *exploitation* pour des objectifs de recherche fondamentale sur la production de la parole (la *coarticulation*, l'étude de l'*inversion* et l'évaluation de *modèles physiques de la production de la parole*).

À l'issue du projet, nous mettrons à la disposition de la communauté scientifique :

/1/ un ensemble de données *multilingues* et multimédia, *unique au monde*, sur la production et la perception de la parole, comprenant images cinéroradiographiques du conduit vocal, signal acoustique et tracés sagittaux des contours du conduit vocal, le tout synchronisé, qui sera accessible au sein d'une plateforme de traitement ;

/2/ des outils et des logiciels d'exploitation adaptés, permettant d'extraire les informations linguistiques de ces données ;

/3/ et, bien entendu, les résultats de nos productions scientifiques, sous forme d'ouvrages, d'articles, de publications électroniques, etc.

Les partenaires du présent projet constituent un remarquable ensemble de compétences transdisciplinaires, mobilisant des savoir-faire de groupes ayant de fortes connaissances théoriques en production et en perception de la parole, en phonologie, en traitement du signal et de l'image, en inversion acoustique articulatoire et en modélisation.

**Abstract** (Do not exceed 5000 car.)

The main thrust of the present research proposal in speech production and perception is to contribute to understanding and modelling of control of speech gestures, by analysing *coarticulation*, by studying *inversion* and by *evaluating physical models*.

Speech production requires a precise coordination of articulators in a given space; gestural overlap, or *coarticulation*, follows constraints that are sound-, language- and speaker-specific. This aspect of research on the spatiotemporal organisation of linguistic gestures is thus central to works on speech production and perception (cf. research axe 1).

Understanding the auditory and visual perceptual dimensions of speech, calls for an adequate knowledge of the *ill-posed* relationship between the acoustic output and the articulatory origins of sounds. Acoustic *inversion* consists in recovering vocal tract changes from the speech signal (cf. research axe 2).

Comprehending the mechanisms that underlie speech production control cannot bypass development and *evaluation* of motor control and *physical models* of the speech production apparatus. It is by comparing data obtained from speakers, in well-controlled conditions, with predictions proposed by these models, that one could evaluate, in a quantitative and objective manner, various hypotheses attested in the literature on speech production control (cf. research axe 3).

In order to study the motor control of human speech gestures and their perceptual efficiency, the basic substrate is composed of a substantial amount of physiological, kinematic, acoustic and eventually dynamic data, collected from human speakers. This stage of data acquisition and analyses is indispensable, and has continually made significant technological progress during the last twenty years.

Despite such advances, *lack of X-ray data* is still a real setback in the speech community, preoccupied with the study of speech production and perception. Indeed, X-ray data are the *only type of data* that offer a correct and complete spatio-temporal resolution of the entire vocal tract in the sagittal plane.

However, acquisition of new cineradiographic data is today restricted and even prohibited in certain countries.

Our project on motor control of human gestures in general, and also on their perceptual efficiency aims, consequently, to develop and to share part of the several X-ray movies on production and perception of speech that were made at the Phonetics Institute of Strasbourg (IPS), as from the end of the fifties. The project concerns 20 movies of very high quality, dealing with linguistic issues in languages spoken in Europe, in Africa, in Asia and in Latin America.

The project comprises 4 major objectives, three specific goals, and a general one that corresponds to the research axes:

- /1/ saving patrimonial X-ray data;
- /2/ their processing (by developing tools);
- /3/ their enhancement by integrating them into a data base, by holding workshops and meetings;
- /4/ and their analysis, together with their exploitation, with basic research aims in speech production (coarticulation, inversion and evaluation of physical models of speech production).

At the end of the project, the following will be made available to the scientific community:

- /1/ a set of multilingual and multimedia data, *unique* in the world, on speech production and perception, containing cineradiographic images of the vocal tract, acoustic signals, tracings and contours of the vocal tract, all synchronised and accessible within a processing platform;
- /2/ adapted tools and softwares that allow extracting linguistic information from these data;
- /3/ and, of course, results of our scientific productions in the form of books, articles, electronic publications, etc.

The partners of the present project make up an outstanding blend of trans-disciplinary competencies, harnessing know-how of groups with a strong theoretical background in speech production and perception, phonology, speech signal and image processing, acoustic-to-articulatory inversion and modelling.

--

**Coordinateur du projet (Partenaire 1)**

Civilité	Nom	Prénom	Discipline	Laboratoire (nom complet)	Type établissement
Monsieur	SOCK	Rudolph	Phonétique Générale et Expérimentale	Composante Parole et Cognition de l'Equipe d'Accueil 1339 Linguistique, Langues et Parole - LiLPa / Institut de Phonétique de Strasbourg (IPS)	public

**Nom des responsables scientifiques des autres partenaires**

	Civilité	Nom	Prénom	discipline	Laboratoire (nom complet)	Type : établissement public
Partenaire 2 <sup>1</sup>	M.	PERRIER	Pascal	Electronique / Signal, Image, Parole	Institut de la Communication Parlée de Grenoble (ICP)	public
Partenaire 3	M.	LAPRIE	Yves	Informatique – Ingénieur Civil / Signal, Image, Parole	Laboratoire Lorrain de Recherches en Informatique et ses Applications (LORIA)	public
Partenaire 4						

**Nombre de personnes impliquées dans ce projet (en équivalent temps plein : ETP)<sup>2</sup>:**

Chercheurs et enseignants-chercheurs permanents \_\_\_\_\_  
 Post-doctorant(s) déjà recruté(s) \_\_\_\_\_ Doctorant(s) \_\_\_\_\_ Ingénieurs et techniciens \_\_\_\_\_  
 Personnes à recruter \_\_\_\_\_

**Durée du projet :** 48 mois

<sup>1</sup> Les cases peuvent être multipliées en fonction du nombre de partenaires.

<sup>2</sup> Quelque soit la catégorie de personnel, il s'agit ici, pour chaque personne impliquée dans le projet, de multiplier son temps de recherche par le pourcentage de temps qu'il consacra à ce projet.

## Dimensionnement total du projet

<b>Coût complet du projet :</b>	€ Reporter le total indiqué au tableau (a) du récapitulatif global (section D du document)
<b>Aide financière demandée :</b>	€ Reporter le total indiqué au tableau (b) du récapitulatif global (section D du document)
<b>Effort en personnel demandé :</b>	<b>homme. mois</b> Reporter le total indiqué au tableau (c) du récapitulatif global (section D du document)

Je déclare exactes toutes les informations contenues dans ce document et m'engage à envoyer une copie de ce dossier à chacun des établissements ou organismes de rattachement de mon laboratoire.

Visa du directeur du laboratoire

Nom, Prénom  
Date et signature du **coordinateur du projet** précédé de la mention « Lu et approuvé »

Nom prénom, date et signature du directeur du laboratoire

**En cas de recouvrement thématique avec d'autres appels à projets (AAP) lancés par l'ANR, les coordinateurs de projet devront veiller à choisir l'AAP le mieux adapté à leur projet. Les personnes impliquées dans plusieurs AAP soumis à l'ANR devront le mentionner dans le tableau « demandes de contrats en cours d'évaluation » (Section D du document).**

# APPEL A PROJETS DE RECHERCHE

## II - PRESENTATION DETAILLEE DU PROJET

### A - Identification du coordinateur et des autres partenaires du projet

Acronyme ou titre court du projet :	<b>DOCVACIM</b>
-------------------------------------	-----------------

#### A-1 – Partenaire 1 = Coordinateur du Projet

Un coordinateur, responsable scientifique du projet, doit être désigné par les partenaires.

**\* champ obligatoire**

Civilité *	Nom *	Prénom *	
Monsieur	SOCK	Rudolph	
Grade *	Professeur	Employeur *	Université Marc Bloch – Strasbourg 2
Mail *	sock@umb.u-strasbg.fr		
Tél *	03 88 41 73 68	Fax 03 88 41 73 69	

<b>Laboratoire (nom complet) *</b>			
Institut de Phonétique de Strasbourg / Composante Parole et Cognition Equipe d'Accueil Linguistique, Langues et Parole (LiLPa)			
N° Unité (s'il existe)	E.A. 1339		
<b>Adresse complète du laboratoire *</b>			
Université Marc Bloch 22, rue Descartes			
Code postal *	67084	Ville *	Strasbourg
<b>Etablissements de tutelle</b> (indiquer le ou les établissements et organismes de rattachement, souligner l'établissement susceptible d'assurer la gestion du projet) :			
MISHA....			

#### Principales publications :

Liste des 10 principales publications ou brevets de l'équipe partenaire 1 (définie tableau ci-dessous) au cours des cinq dernières années, relevant du domaine de recherche couvert par la présente demande dans l'ordre suivant : Auteurs (en soulignant les auteurs faisant effectivement partie de la demande), Année, Titre, Revue, N°Vol, Pages. N'indiquez pas les publications soumises.

- [1] SOCK R. VAXELAIRE B. ROY J.-R. HIRSCH F. HECKER V. (2005) Temporal and spatial correlates of quantity contrasts in Wolof. X-ray, kinematic and acoustic data. Psychology Press. Taylor & Francis Group. New York and Hove, 253 – 274.
- [2] WIOLAND F. (2005) La vie sociale des sons du français, L'Harmattan, 216 p.
- [3] SOCK R. VAXELAIRE (2004) Émergence de structures phonétiques et catégorisation phonologique. À paraître in Mélanges de linguistique générale et française, offerts au Professeur Georges KLEIBER. M. RIEGEL C. SCHNEDECKER & A. THEISSEN. (Eds.), PEETERS, Liège, Paris.
- [4] SOCK R. VAXELAIRE B. (2004) Le diable cognitif dans les détails sensori-moteurs anticipatoires. In « L'Anticipation à l'horizon du présent », 141 – 157. SOCK R. & VAXELAIRE B. (Eds.), Psychologie et Sciences Humaines, Mardaga, Liège.
- [5] HIRSCH F. SOCK R. CONNAN P.-Y. BROCK G. (2003). Auditory effects of anticipatory rounding in relation with vowel height in French, Proceeding of the International Congress of Phonetics Sciences, Barcelone, 1445-1448
- [6] CONNAN P.-Y. ROY J.-P. BROCK G. HIRSCH F. (2003). Using Digital Cine-Radiography to Study Anticipatory Labial Activity in French, ICPHS Proceeding, Barcelone, 3153-3156.
- [7] SOCK R. (2003) Comprendre les gestes anticipatoires audibles et inaudibles. In Stratégies et parcours. De l'anticipation à la didactique du FLE. Revue Sciences Cognitives, Linguistique et Intelligence Artificielle (SCOLIA) – SOCK R. KLEIBER G. KASHEMA L. (Eds.), Publication de l'Université Marc Bloch, vol. 17, 11 – 28.
- [8] VAXELAIRE B. & SOCK R. ROY J.-P. ASCI A. HECKER V. (2003) Audible and inaudible anticipatory gestures in French. 15th International Congress of Phonetic Sciences, Barcelona, 3 – 9 August, 447 - 450.

[9] VAXELAIRE B. HECKER V. SOCK R. 2002 La perception auditive de gestes vocaliques anticipatoires. *XXIV<sup>èmes</sup> Journées d'Etudes sur la Parole de l'Association Francophone de la Communication Parlée* (JEP de l'AFCP), Nancy, juin 2002, 109 – 112.

[10] SOCK R. VAXELAIRE (2001) Peut-on travailler sans représentations en production-perception de la parole ? In *Par monts et par vaux. Itinéraires linguistiques et grammaticaux. Mélanges de linguistique générale et française, offerts au Professeur Martin RIEGEL*. In BURIDANT C. KLEIBER G. & PELLAT J.-C. (Eds.), PEETERS, Liège, Paris, 379-391.

Ce projet fait-il partie des projets labellisés (ou en cours de labellisation) par un pôle de compétitivité (ou par plusieurs, en cas de projet interpôle) ? **NON**  
**Si oui, nom du pôle ou des pôles :**

**Partenaire 1 = Coordinateur du Projet : Institut de Phonétique de Strasbourg / Composante Parole et Cognition**

	Nom	Prénom	Emploi actuel	Discipline	% de temps de recherche consacré au projet	Rôle/Responsabilité dans le projet 4 lignes max
<i>exemple</i>	<i>MARTIN</i>	<i>Charlotte</i>	<i>Professeur</i>		<i>30%</i>	
Coordinateur	SOCK	Rudolph	Professeur	Phonétique Générale et Expérimentale	30%	Coordination générale du projet ; formation et encadrement de Post-Doctorants pour l'analyse des radiofilms ; exploitation des données pour l'élaboration de contraintes coarticulatoires ; valorisation en collaboration avec la MISHA UMS 2552
Membres de l'équipe	VAXELAIRE	Béatrice	Maître de Conférences	Phonétique Générale et Expérimentale	20%	Formation et encadrement de Post-Doctorants pour l'analyse des radiofilms, la réalisation des croquis et la numérisation des croquis ; exploitation des données pour l'étude de la coarticulation.
	ZERLING	Jean-Pierre	Professeur	Phonétique Générale et Expérimentale	15%	Sélection des films et repérage des séquences pertinentes par rapport au corpus ; inventaire des tracés déjà existants ; exploitation des données pour l'étude de la coarticulation.
	WIOLAND	François	Professeur	Phonétique Générale et Expérimentale	10%	Sélection des films et repérage des séquences pertinentes par rapport au corpus ; inventaire des tracés déjà existants ; exploitation des données pour l'étude de la coarticulation.
	BROCK	Gilbert	Ingénieur d'Etudes	Audiovisuelle	30%	Sauvegarde et mise en forme des données ; transfert des films du standard 35 mm vers standard vidéo BetacamSP ; synchronisation des images et du son.

Pour chacun des membres de l'équipe du projet, fournir une biographie **d'une page maximum** qui comportera :

A/ Nom, prénom, âge, cursus, situation actuelle

B/ Autres expériences professionnelles

C/ Liste des 10 publications (ou brevets) comprenant les 5 publications les plus récentes et les 5 publications les plus significatives

D/ Prix, distinctions



SOCK Rudolph  
49 ans

Professeur des Universités 2°CI – 7° section CNU

- Responsable de la Composante Parole et Cognition  
E.A. 1339 – Linguistique, Langues et Parole (LILPA)  
Institut de Phonétique de Strasbourg, Université Marc Bloch – Strasbourg 2

### Formation

- Doctorat de 3ème Cycle de Linguistique et Phonétique : 1983  
Institut de Phonétique de Grenoble -Université Stendhal
- Doctorat d'Etat en Linguistique - Sciences du Langage & Communication Parlée : 1998  
Institut de la Communication Parlée de Grenoble

### Cursus professionnel

- Chargé de Cours - Université de Neuchâtel - Suisse : 1985 - 1986.
- Maître de Conférences 2°CI - Université Marc Bloch - Strasbourg : 1994 à 1996
- Maître de Conférences 1°CI - Université Marc Bloch - Strasbourg : 1996 à 2001
- Professeur des Universités : 2001 –
- Missions d'Enseignement (TS -Teaching Staff Mobility - Socrate) Institut de Phonétique de Munich - Séminaires, cours et Travaux Dirigés
- Chercheur-Associé aux Haskins Laboratories - U.S.A. : 1991 – 1996
- Boursier de la Fondation Fyssen — Sciences Cognitives : éthologie :1991 – 1992
- Participant au projet NIH "Sources of Sound in Speech" (attribué à A. Löfqvist, Haskins Labs. U.S.A.) : 1991 - 1992.
- Directeur de l'Equipe d'Accueil 3403 Institut de Phonétique de Strasbourg : 2001 - 2005
- Enseignement : Phonétique Générale et Expérimentale, Phonologie
- Recherche : Organisation Spatio-Temporelle en Production-Perception de la Parole

### 5 PUBLICATIONS LES PLUS RECENTES

- [1] SOCK R. VAXELAIRE B. ROY J.-R. HIRSCH F. HECKER V. (2005) Temporal and spatial correlates of quantity contrasts in Wolof. X-ray and acoustic data. Psychology Press. Taylor & Francis Group. New York and Hove, 253 – 274.
- [2] SOCK R. VAXELAIRE (2004) Émergence de structures phonétiques et catégorisation phonologique. À paraître in Mélanges de linguistique générale et française, offerts au Professeur Georges KLEIBER. M. RIEGEL C. SCHNEDECKER & A. THEISSEN. (Eds.), PEETERS, Liège, Paris.
- [3] SOCK R. VAXELAIRE B. (2004) Le diable cognitif dans les détails sensori-moteurs anticipatoires. In « L'Anticipation à l'horizon du présent », 141 – 157. SOCK R. & VAXELAIRE B. (Eds.), Psychologie et Sciences Humaines, Mardaga, Liège.
- [4] SOCK R. (2003) Comprendre les gestes anticipatoires audibles et inaudibles. In Stratégies et parcours. De l'anticipation à la didactique du FLE. Revue Sciences Cognitives, Linguistique et Intelligence Artificielle (SCOLIA) – SOCK R. KLEIBER G. KASHEMA L. (Eds.), Publication de l'Université Marc Bloch, vol. 17, 11 – 28.
- [5] SOCK R. VAXELAIRE (2001) Peut-on travailler sans représentations en production-perception de la parole ? In Par monts et par vaux. Itinéraires linguistiques et grammaticaux. Mélanges de linguistique générale et française, offerts au Professeur Martin RIEGEL. In BURIDANT C. KLEIBER G. & PELLAT J.-C. (Eds.), PEETERS, Liège, Paris, 379-391.

### 5 PUBLICATIONS LES PLUS SIGNIFICATIVES

- [1] ABRY C. ORLIAGUET J.P. SOCK R. (1990) Patterns of Speech Phasing. Their Robustness in the Production of a Timed Linguistic Task: Single vs. Double (Abutted) Consonants in French. European Bulletin of Cognitive Psychology 10, 269-288.
- [2] PERRIER P. BOË L.J. SOCK R. (1992) Vocal Tract Area Function Estimation from Midsagittal Dimensions with CT Scans and a Vocal Tract Cast: Modelling the Transition with Two Sets of Coefficients. J. Speech Hearing Res. 35, 53-67.
- [3] SOCK R. LÖFQVIST A. (1995) Some Timing Constraints in the Production of Bilabial Stops. Journal of Phonetics, 23, 129-138.
- [4] SOCK R. (1998) Organisation temporelle en production de la parole. Émergence de catégories sensori-motrices phonétiques. Septentrion, Villeneuve d'Ascq – 479 p.
- [5] SOCK R. (2001) La théorie de la viabilité en production-perception de la parole. In KELLER D. DURAFOR J.-P. BONNOT J.-F. & SOCK R. (Eds.), Psychologie et Sciences Humaines, Mardaga, Liège, 285 – 316.

VAXELAIRE Béatrice

Maître de Conférences - 7° section CNU

42 ans

• Composante Parole et Cognition

E.A. 1339 – Linguistique, Langues et Parole (LILPA)

Institut de Phonétique de Strasbourg, Université Marc Bloch – Strasbourg 2

### Responsabilité

• Responsable de l'Equipe Cinéradiographie de l'IPS

### Formation

• Doctorat de Linguistique, Phonétique & Informatique : 1993

Institut de Phonétique de Strasbourg -Université Marc Bloch – Strasbourg 2

### Cursus professionnel

• ATER Sciences du Langage : 1992 à 1994.

• Maître de Conférences - Université Marc Bloch - Strasbourg : 2002 –

• Postdoctoral Researcher à Queen Margaret College - Contrat Européen ACCOR - SPHERE : 1996

• Participante au projet ACCOR - SPHERE « EMA – EPG Data Combined Analyses » : 1996

• Enseignement : Phonétique Générale et Expérimentale, Phonologie

• Recherche : Cinéradiographie et gestes phonétiques

### 5 PUBLICATIONS LES PLUS RECENTES

[1] SOCK R. VAXELAIRE B. ROY J.-R. HIRSCH F. HECKER V. (2005) Temporal and spatial correlates of quantity contrasts in Wolof. X-ray and acoustic data. Psychology Press. Taylor & Francis Group. New York and Hove, 253 – 274.

[2] SOCK R. VAXELAIRE B. (2004) Le diable cognitif dans les détails sensori-moteurs anticipatoires. In « L'Anticipation à l'horizon du présent », 141 – 157. SOCK R. & VAXELAIRE B. (Eds.), Psychologie et Sciences Humaines, Mardaga, Liège.

[3] VAXELAIRE B. & SOCK R. ROY J.-P. ASCI A. HECKER V. (2003) Audible and inaudible anticipatory gestures in French. 15th International Congress of Phonetic Sciences, Barcelona, 3 – 9 August, 447 - 450.

[4] VAXELAIRE B. HECKER V. SOCK R. 2002 La perception auditive de gestes vocaliques anticipatoires. *XXIV<sup>èmes</sup> Journées d'Etudes sur la Parole de l'Association Francophone de la Communication Parlée* (JEP de l'AFCP), Nancy juin 2002, 109 – 112.

[5] SOCK R. VAXELAIRE (2001) Peut-on travailler sans représentations en production-perception de la parole ? In Par monts et par vaux. Itinéraires linguistiques et grammaticaux. Mélanges de linguistique générale et française, offerts au Professeur Martin RIEGEL. In BURIDANT C. KLEIBER G. & PELLAT J.-C. (Eds.), PEETERS, Liège, Paris, 379-391.

### 5 PUBLICATIONS LES PLUS SIGNIFICATIVES

[1] VAXELAIRE B. (1996) La notion de contexte en phonétique : un aperçu. Sciences Cognitives, Linguistique & Intelligence Artificielle – SCOLIA 6, 87 – 101.

[2] VAXELAIRE B. SOCK R. (1996) A cineradiographic and acoustic study of velar gestures in French. First European Speech Communication Association (ESCA) Tutorial and Research Workshop, Autrans-France, 65-68.

[3] VAXELAIRE B. (1997) Contrôle du velum et nasalité en français. Sciences Cognitives, Linguistique & Intelligence Artificielle – SCOLIA 10, 39 – 60.

[4] VAXELAIRE B. SOCK R. (1997) Laryngeal movements and speech rate. An X-ray investigation. Eurospeech '97. 5<sup>th</sup> European Conference on Speech Communication and Technology, Rhodes, Greece. Vol. 2, 1039-1042.

[5] VAXELAIRE B. SOCK R. (1998) Perturbing phonetic configurations and gestures using speech rate. Editions. L'Harmattan, Paris, 447- 453.

**Jean-Pierre ZERLING**

né le 31 juillet 1946 à Nancy (54)

**Professeur de Phonétique Générale et Expérimentale,  
Directeur de l'Institut de Phonétique de Strasbourg (iPS)**

• **Titres universitaires**

- **Doctorat d'État de Phonétique** (Mention Très Honorable), Strasbourg II, 4 juillet 1990,  
"Aspects articulatoires de la labialité vocalique en français. Contribution à la modélisation à partir de labio-photographies, labiofilms et films radiologiques. Étude statique, dynamique et contrastive".  
Directeur : Péla SIMON (Strasbourg). Jury : P.SIMON (Strasbourg), F.CARTON (Nancy), M.ROSSI (Aix-en-Provence), F.WIOLAND (Strasbourg), L.J.BOË (Grenoble).
- **Doctorat de 3<sup>e</sup> Cycle de Phonétique** (Mention Très Bien), Nancy II, 6 sept. 1979,  
"Articulation et coarticulation dans des groupes occlusive-voyelle en français. Etude ciné-radiographique et acoustique ; contribution à la modélisation articulatoire".  
Directeur : F. CARTON (Nancy). Jury : P.SIMON (Strasbourg), M.ROSSI (Aix-en-Provence)

• **Activités d'enseignement**

- 1976 à 1981 **Assistant** à l'Université de **Lille III** (Villeneuve d'Ascq).
- 1981 à 1993 **Maître de Conférences de Phonétique** à l'Université Marc Bloch - **Strasbourg II**.
- Depuis 1993 **Professeur de Phonétique** à l'Université Marc Bloch - **Strasbourg II**.  
Lettres Modernes, Sciences du Langage, FLE (Français Langue Étrangère).
- Depuis 1993 **Chargé de cours à l'École d'Orthophonie de Strasbourg (ULP, Faculté de Médecine)**.
- Depuis 1993 **Chargé de cours à l'I.I.E.F.** (l'Institut International d'Études Françaises de Strasbourg, UMB).

• **Activités de recherche et participation à des programmes de recherche**

- 1999-2007 **Contrat d'encadrement doctoral et de recherche** à l'UMB
- 1997-99 **Contrat CNRS** de 3 ans entre ICP (Institut de la Communication Parlée, Grenoble, et l'iPS (Institut de Phonétique de Strasbourg). **Programme ingénierie des langues**. Thème : Valorisation de la banque de données ciné-radiographiques de l'IPS : Numérisation des données et élaboration d'une plate-forme multimédia pour leur analyse.
- 2001-04 **Contrat Cognitif (Action) du MRT** de 3 ans, entre ICP (Institut de la Communication Parlée, Grenoble, et l'iPS). Thème : Contrôle du mouvement et de l'efficacité perceptive de gestes anticipatoires.
- 2001-05 Équipe d'Accueil : **EA-3403 - Institut de Phonétique de Strasbourg**.
- Depuis 2005 Équipe d'Accueil : **EA-1339 - Linguistique, Langues & Parole (LiLPA)** Composante Parole et Cognition

• **Publications et communications (5 parmi les plus significatives)**

- BOTHOREL A., SIMON P., WIOLAND F. & ZERLING J.P. (1986) Ciné-radiographie des voyelles et consonnes du français. Recueil de documents synchronisés pour 4 sujets : vues latérales du conduit vocal, vues frontales de l'orifice labial, données acoustiques, Publication de l'Institut de Phonétique de Strasbourg, 298 p.
- ZERLING J.P. (1998) Mini-dico verlan, Coll. Le "Trois-Demi", Édition Biotop, Paris
- ZERLING J.P. (1999) "Base articulatoire et contraintes en production de la parole. Approche expérimentale et comparative en français et en portugais du Brésil", (conférence invitée), V Congresso Nacional de Fonética e Fonologia, Universidade Federal Fluminense (UFF), Niterói - Rio de Janeiro, Brésil, 2-4 décembre 1996, 20 p.
- ZERLING J.P. (1999) " Structure syllabique et morphologique du verlan. Analyse phonétique à partir d'un lexique de verlan citadin. ", Travaux de l'Institut de Phonétique de Strasbourg (TIPS), 29, pp. 73-94.
- ZERLING J.P. (2000) "Structure syllabique et morphologique des mots à caractère onomatopéique et répétitif en français. Étude phonétique", Travaux de l'Institut de Phonétique de Strasbourg (TIPS), 30, pp. 115-162.

• **Publications et communications (5 récentes)**

- ZERLING J.P. (2000) "Contraintes et stratégies phonétiques. Illustration expérimentale comparée.", Conférence d'ouverture, Actes du VI Congresso Nacional de Fonética e Fonologia, Sociedade Brasileira de Fonética, Université Fédérale Fluminense, Niteroi-RJ, Brésil, 27-29 novembre 2000, 18 p.
- ZERLING J.P. (1998) "Verlan 2000 à la Bastille. Une approche phonétique", Travaux de l'Institut de Strasbourg (TIPS), 28, pp. 203-232.
- COUTINHO G., MESSIAS L. & ZERLING J.P. (2001) "Le Brésil et l'expression des sentiments. Onomatopées, interjections et exclamations", Travaux de l'Institut de Phonétique de Strasbourg (TIPS), 31, pp. 23-60.
- ZERLING J.P. & DE CASTRO L. (2002) "Analyse comparée de trois patrons prosodiques en français et en portugais européen", Travaux de l'Institut de Strasbourg (TIPS), 32, pp. 1-35.
- MAURY D. & ZERLING J.P. (2004) "Le comportement des voyelles du français par rapport à un modèle de référence", Colloque MIDL 2004 - Identification des langues et des variétés dialectales par les humains et par les machines, LIMSI-CNRS, Paris, 29-30 nov, p. 85

François WIOLAND  
Né le 1/4/39 à Strasbourg  
Professeur de phonétique générale et expérimentale  
Directeur de l'Institut International d'Etudes Françaises  
Chevalier des Palmes Académiques  
Conseiller du Président de la FIPF  
Prof. des Universités 1<sup>o</sup> classe  
Docteur d'Etat

1) Activités de recherche (10 dernières années)

1997 – 1999 : Contrat CNRS programme ingénierie des langues  
2001 – 2004 : Contrat Cognitique du MRT  
2001 – 2005 : EA 3403 Institut de Phonétique de Strasbourg  
2005 → : EA 1339 Linguistique, Langues et Parole (LiLPA) Composante Parole et Cognition

**5 PUBLICATIONS LES PLUS RECENTES**

- [1] Wioland F. (2004) La réduction des mots longs : quelques observations d'ordre phonétique, Colloque « Perturbations et réajustements », COGNIEST, Haguenau, 1-3 décembre 2004, publié dans Travaux de l'Institut de Phonétique de Strasbourg (TIPS), 2004.
- [2] Wioland F. (2002) Les structures syllabiques et la distribution des phonèmes en coréen parlé, publié dans les Travaux de l'Institut de Phonétique de Strasbourg (TIPS), 32, en col. avec Eun-Yung LEE, 149 – 169.
- [3] Wioland F. (2001) Que faire de la graphie « e » ? Le Français dans le Monde 318, 31 – 33.
- [4] Wioland F. (2000) Vers un modèle prosodique du français parlé? *Apprendre enseigner, acquérir : La prosodie au cœur du débat*, Collection Dyalang, Publications de l'Université de Rouen, CNRS, octobre 2000, 13 – 19.
- [5] Wioland F. (1999) La graphie E muette, Actes du XIIIe Congrès brésilien des professeurs de français, *L'enseignement pluriel du français*, Salvador, Bahia, Brésil, Elos, n°3, 1999, 60 – 62.

**5 PUBLICATIONS LES PLUS SIGNIFICATIVES**

Ouvrages

- [1] Bothorel A., Simon P., Wioland F. & Zerling J.P. (1986) Cinéradiographie des voyelles et consonnes du français. Recueil de documents synchronisés pour 4 sujets : vues latérales du conduit vocal, vues frontales de l'orifice labial, données acoustiques, Publication de l'Institut de Phonétique de Strasbourg, 298 p.
- [2] Wioland F. (1990) Prononcer les sons du français. Hachette.
- [3] Wioland F. Wenk B.J. (1982) Is French really syllable-timed? *Journal of Phonetics* 10.
- [4] Wioland F. (2001) La graphie E : quelles prononciations ? Publications de l'UFSC, Florianopolis, Brésil.
- [5] Wioland F. (2005) La vie sociale des sons du français, L'Harmattan, 216 p.

Gilbert Brock  
Né le 8 mars 1951  
Institut de Phonétique de Strasbourg /  
Linguistique Langues et Parole (LiLPa) E.A. 1339  
Composante Parole et Cognition  
22, rue Descartes - 67084 Strasbourg  
Mél. : brock@umb.u-strasbg.fr

**Fonctions** Ingénieur d'Etudes 1 C  
Responsable technique : Audiovisuelle - Informatique

### Activités de recherche et participation à des programmes de recherche

- 1997-99** Contrat CNRS de 3 ans entre ICP (Institut de la Communication Parlée, Grenoble, et l'iPS (Institut de Phonétique de Strasbourg). Programme ingénierie des langues. Thème : Valorisation de la banque de données cinéradiographiques de l'IPS : Numérisation des données et élaboration d'une plate-forme multimédia pour leur analyse.
- 2001-04** Contrat Cognitique (Action du MRT) de 3 ans, entre ICP (Institut de la Communication Parlée, Grenoble, et l'iPS). Thème : Contrôle du mouvement et de l'efficacité perceptive de gestes anticipatoires.
- 2001-05** Equipe d'Accueil : EA-3403 - Institut de Phonétique de Strasbourg.
- Depuis 2005** Equipe d'Accueil : EA-1339 - Linguistique, Langues & Parole (LiLPA) Composante Parole et Cognition

### Publications

1. BROCK G., 1977. Méthode de synchronisation graphique images-son pour l'exploitation des films radiologiques. Présentation de l'appareillage réalisé à l'Institut de Phonétique de Strasbourg. TIPS 9, 221-232.
2. P. SIMON, G. BROCK & M.H. HAN, 1978. Description et utilisation d'un équipement à rayons X pour l'étude de certains aspects articulatoires. Application au coréen in Articulatory Modeling and Phonetics 1978 (pp. 223-242).
3. G. BROCK, F. WIOLAND. & P; SIMON, 1978. La chaîne de segmentation et l'analyse en temps réel. Présentation et application. TIPS 10, 133-164.
4. A. BOTHOREL, G. BROCK & MAILLARD-SALIN, 1980. Mouvement des lèvres, déplacement du larynx et variation de la longueur du conduit vocal. Séminaire international sur la la bialité. Lannion, février 1980.
5. A. BOTHOREL, C. BROCK, G. MAILLARD-SALIN, 1980. Contribution à l'étude des rapports entre les mouvements de l'os hyoïde et le déplacement du larynx. TIPS 12, 225-269.
6. BROCK G., NAJAFIZADETH H., SIMON P., & WOLFF F., 1984. Numérotation et codage synchrones graphique et phonique pour l'exploitation des films radiologiques. TIPS 16, 181-190.
7. BROCK G., NAJAFIZADETH H., SIMON P., & WOLFF F., 1987. Numérotation synchronisée radiologique. Radiologie, 381-383.
8. BROCK G., 1989. Optimisation d'une station analogique de traitement du signal acoustique. Mélanges Péla SIMON, 147-155.
9. METZ-LUTZ M.N., WIOLAND F., BROCK G. 1991. A Real-time Approach to Spoken Language Processing in Aphasia. Brain and Language
10. METZ-LUTZ M.N., WIOLAND F., BROCK G. 1991. Auditory Lexical Access in Aphasia : Effects of Phonological, Morphemic and Semantic Priming. Journal of Clinical and Experimental Neuropsychology, 1991, 13, 3, 435.
11. CONNAN P.Y., WIOLAND F., METZ\_LUTZ M.N., BROCK G., 1992. Analyse acoustico-phonétique du message verbal. Son rôle dans la reconnaissance lexicale. Proceedings of the XII International Congress of Phonetic Science, vol.5, 78-81.
12. METZ LUTZ M.N., WIOLAND F., BROCK G. 1993 What is wrong with auditory comprehension in aphasia? Inter disciplinary Perspectives in Speech and Language Pathology, M.M. Leahy, J.L. Kallen SCSLS / TCD Dublin, 1993? 187 - 195
13. N. WIOLAND, J.P. ZERLING, G. BROCK, G. RUDOLF, D. KURTZ - 1994 The acoustic-phonetic analysis of French non sense syllables. An ERP study. European Congress of Psychophysiology, Barcelone, 7-10 avril 1994. Contribution to the 15th International Congress on Acoustics 26-30 June 1995, Trondheim, Norway :
14. BADIN P, BEAUTEMPS D. , LALLOUACHE T. M. , BAILLY G. (GRENOBLE) GABIOUD B.(LAUSANNE) , MAEDA S. (PARIS), ZERLING J.P. , & BROCK G. (STRASBOURG) 1995 Cineradiography of VCV sequences : Articulatory-Acoustic Data for a speech production model.

Acronyme ou titre court du projet	DOCVACIM
-----------------------------------	----------

**A-2 : Autres partenaires du projet (remplir une fiche par partenaire)**

Un responsable scientifique de l'équipe partenaire doit être désigné

**Partenaire 2 : ICP****\* champ obligatoire**

<b>Civilité *</b>	<b>Nom *</b>	<b>Prénom *</b>	
Monsieur	PERRIER	Perrier	
<b>Grade *</b>		<b>Employeur *</b>	Institut National Polytechnique de Grenoble
<b>Mail *</b>	perrier@icp.inpg.fr		
<b>Tél *</b>		<b>Fax</b>	04 76 57 47 10

<b>Laboratoire * (nom complet)</b>	
Institut de la Communication Parlée	
<b>N° Unité (s'il existe)</b>	UMR CNRS 5009
<b>Adresse complète du laboratoire *</b>	
Institut National Polytechnique de Grenoble 46 Avenue Félix Viallet	
<b>Code postal *</b>	38031
<b>Ville *</b>	Grenoble
<b>Etablissements de tutelle (indiquer le ou les établissements et organismes de rattachement, souligner l'établissement susceptible d'assurer la gestion du projet) :</b>	

**Principales publications :**

Liste des 10 principales publications ou brevets de l'équipe partenaire 2 (définie tableau ci-dessous) au cours des cinq dernières années, relevant du domaine de recherche couvert par la présente demande dans l'ordre suivant : Auteurs (en soulignant les auteurs faisant effectivement partie de la demande), Année, Titre, Revue, N°Vol, Pages. N'indiquez pas les publications soumises.

- Perrier, P. (2006). About speech motor control complexity. In *Speech Production: Models, Phonetic Processes, and Techniques* (J. Harrington & M. Tabain, editors), New York: Psychology Press. *In Press*
- Fuchs, S., Perrier, P., Geng, C. & Mooshammer, C. (2006). What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. In *Speech Production: Models, Phonetic Processes, and Techniques* (J. Harrington & M. Tabain, editors), New York: Psychology Press. *In Press*
- Tabain, M. & Perrier, P. (2005). Articulation and acoustics of /i/ in preboundary conditions in French. *Journal of Phonetics*, Vol. 33(1), 77-100.
- Perrier, P., Payan, Y. & Marret, R. (2004). Modéliser le physique pour comprendre le contrôle: le cas de l'anticipation en production de parole. In *L'anticipation à l'horizon du Présent* (R. Sock & B. Vaxelaire, editors), pp. 159-177. Sprimont, Belgique: Pierre Margala.
- Apostol, L., Perrier, P. & Bailly, G. (2004). A model of acoustic interspeaker variability based on the concept of formant-cavity affiliation. *Journal of Acoustical Society of America*, Vol 115 (1), 337-351.
- Berthommier, F. (2004). A phonetically neutral model of the low-level audiovisual interaction, *Speech Communication*, Vol.44, 31-41.
- Schwartz, J.-L., Berthommier, F. & Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition*, Vol. 93, B69-B78.
- Perrier, P., Payan, Y., Zandipour, M. & Perkell, J.S. (2003). Influences of tongue biomechanics on speech movements during the production of velar stop consonants: a modeling study. *Journal of Acoustical Society of America*, Vol. 114 (3), 1582-1599.
- Badin, P., Bailly, G., Revéret, L., Baci, M., Segebarth, C. & Savariaux, C. (2002). Three-dimensional linear articulatory modeling of tongue, lips and face based on MRI and video images. *Journal of Phonetics*, Vol.30(3): 533-553.
- Matthies L.M., Perrier P., Perkell J.S. & Zandipour M. (2001). Variation in speech movement kinematics and temporal patterns of coarticulation with changes in clarity and rate. *Journal of Speech Language and Hearing Research*, Vol. 44 (2), 340-353.

**Partenaire 2 : Institut de la Communication Parlée de Grenoble**

	Nom	Prénom	Emploi actuel	Discipline (à renseigner uniquement pour SHS)	% de temps de recherche consacré au projet	Rôle/Responsabilité dans le projet 4 lignes max
<i>exemple</i>	<i>MARTIN</i>	<i>Charlotte</i>	<i>Professeu</i>		3	
Responsable	PERRIER	Pascal	Professeur		1	Coordination du projet ; structuration de la base de données ; exploitation des données pour élaborer et évaluer des modèles de contrôle de la production de la parole (évaluation de modèles de génération de séquences de parole sur des modèles physiques).
Membres de l'	BERTHOMMIER	Frédéric	CR -CNRS		1	Détection automatique de contours radiographiques ; application à l'inversion du signal de parole et à la synthèse articulatoire à partir de la transcription phonétique.
	SAVARIAUX	Christophe	IR - CNRS		1	Mise en forme des données ; numérisation des films radiographiques ; développement de la plateforme logicielle pour le traitement des signaux de mouvements et des signaux acoustiques

Pour chacun des membres de l'équipe du projet, fournir une biographie **d'une page maximum** qui comportera :

A/ Nom, prénom, âge, cursus, situation actuelle

B/ Autres expériences professionnelles

C/ Liste des 10 publications (ou brevets) comprenant les 5 publications les plus récentes et les 5 publications les plus significatives

D/ Prix, distinctions



- Fuchs, S., Perrier, P., Geng, C. & Mooshammer, C. (In press). What role does the palate play in speech motor control? Insights from tongue kinematics for German alveolar obstruents. In *Speech Production: Models, Phonetic Processes, and Techniques* (J. Harrington & M. Tabain, editors), New York: Psychology Press.
- Tabain, M. & Perrier, P. (2005). Articulation and acoustics of /i/ in preboundary conditions in French. *Journal of Phonetics*, Vol. 33(1), 77-100.
- Gérard, J.-M., Ohayon, J., Luboz, V., Perrier, P. & Payan, Y. (2005). Non linear elastic properties of the lingual and facial tissues assessed by indentation technique. Application to the biomechanics of speech production. *Medical Engineering & Physics*, 27, 884–892.

#### **CINQ PUBLICATIONS LES PLUS SIGNIFICATIVES**

- Perrier, P., Payan, Y. & Marret, R. (2004). Modéliser le physique pour comprendre le contrôle: le cas de l'anticipation en production de parole. In *L'anticipation à l'horizon du Présent* (R. Sock & B. Vaxelaire, editors), pp. 159-177. Sprimont, Belgique: Pierre Mardaga.
- Perrier P., Payan Y., Zandipour M. & Perkell J. (2003) Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114(3), 1582-1599.
- Gérard J.M., Wilhelms-Tricarico R., Perrier P. & Payan Y. (2003) A 3D dynamical biomechanical tongue model to study speech motor control. *Recent Res. Devel. Biomechanics*, 1, 49-64.
- Matthies L.M., Perrier P., Perkell J.S. & Zandipour M. (2001). Variation in speech movement kinematics and temporal patterns of coarticulation with changes in clarity and rate. *Journal of Speech Language and Hearing Research*, 44 (2), 340-353.
- Perkell J.S., Guenther F.H., Lane H., Matthies L.M., Perrier P., Vick J., Wilhelms-Tricarico R., & Zandipour M. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics*, 28 (3), 233-272.

## Curriculum Vitae

<b>NOM</b> Frédéric Berthommier 46 ans	<b>Profession</b> Chargé de Recherche 1 <sup>ère</sup> Classe (CNRS) Laboratoire de Recherche : Institut de la Communication Parlée (I.C.P.), UMR CNRS 5009
--	---

### FORMATION

Etablissement	Diplôme	Année	Discipline
Faculté Lariboisière Saint Louis – Paris 7 (France)	Internat		Médecine
Paris 7 (France)	Maîtrise		Biologie Humaine
Paris 6 (France)	Licence		Informatique
Université Joseph Fourier - Grenoble (France)	Doctorat		Génie biomédical

### CURSUS PROFESSIONNEL

Chargé de Recherche 1<sup>ère</sup> Classe – CNRS

### CINQ PUBLICATIONS LES PLUS RECENTES

- Schwartz J.-L., Berthommier, F., Savariaux, C. (2004) Seeing to hear better: Evidence for audio-visual interactions in speech identification, *Cognition*, 93, 69-78.
- Berthommier, F. (2004), A phonetically neutral model of the low-level audiovisual interaction, *Speech Communication*, 44, 31-41.
- Heckmann M., Berthommier F., Kroschel K. (2002) Noise adaptive stream weighting in audio-visual speech recognition, *Applied Signal Processing*, 11, 1260-1273.
- Giraudet, P., Berthommier, F. & Chaput, M. (2002). Mitral cell temporal response patterns evoked by odor mixtures in the rat olfactory bulb. *Journal of Neurophysiology* 88(2), 829-838.
- Choi, S., Hong, H., Glotin, H. & Berthommier, F. (2002). Multichannel signal separation for cocktail party speech recognition: a dynamic recurrent network. *Neurocomputing*, 49, 299-314

### CINQ PUBLICATIONS LES PLUS SIGNIFICATIVES

- Berthommier, F. (2000). Les représentations auditives. In *La parole, des modèles cognitifs aux machines communicantes* (P. Escudier & J.-L. Schwartz, éditeurs), Paris: Hermes.
- Lorenzi, C., Berthommier, F., & Demany, L. (1999). Discrimination of amplitude-modulation phase spectrum, *Journal of the Acoustical Society of America*, 105, 2987-2990
- Lorenzi, C., Berthommier, F., Apoux, F. & Bacri, N. (1999). Effects of envelope expansion on speech recognition, *Hearing Research*, 136, 131-138.
- Berthommier, F. & Lorenzi, C. (1998) Implications of physiological mechanisms of amplitude modulation processing for modelling complex sounds analysis and separation, in *CASA*, Rosenthal, D. & Okuno, H. (Eds), LEA, pp. 59-70.
- Berthommier, F., Buonviso, N. & Chaput, M. (1995) A probabilistic model of temporal processing in the olfactory bulb, In *Le Neuromimétisme : Epistémologie, Neurobiologie, Informatique*, H. PaugamMoisy, J.P., Royet & D.A., Zighed (Eds), Hermès, pp. 107-123.

## SAVARIAUX Christophe

Né en 1968 à Paris IX<sup>e</sup>  
5 Avenue Romain Rolland  
38400 S<sup>t</sup> Martin d'Hères

Situation actuelle : Ingénieur de recherche CNRS 2<sup>ème</sup> classe en poste à l'ICP depuis septembre 1999.

Diplômes Universitaires :

- DEUG A, Mathématiques et Physique : 1987, Poitiers.
- Licence Électronique Électrotechnique Automatique, option "traitement du signal et de l'information" : 1989, Poitiers. **Mention AB.**
- Maîtrise Électronique Électrotechnique Automatique, option "traitement du signal et de l'information" : 1990, Poitiers. **Mention AB.**
- Diplôme d'Étude Approfondie en Signal Image Parole : 1991, INP Grenoble. **Mention B.**
- Thèse de Doctorat de l'Institut National Polytechnique de Grenoble, spécialité : Signal Image Parole, janvier 1995. **Mention très honorable avec félicitations<sup>3</sup>.**

Laboratoire de recherche :

Institut de la Communication Parlée (ICP)  
Université Stendhal

Activité majeure :

Responsable du bloc expérimental de l'ICP : acquisition et traitement des données audio, vidéo et électromagnétiques.

Actions principales :

- Acquisition vidéo multi-sources. Numérisation sur station DPS ou par logiciel via une carte matrox. Développement d'un logiciel de traitement d'images couleurs (chroma-key numérique) pour extraction du contour des lèvres face/profil..
- Acquisition et traitement de données électromagnétiques enregistrées à partir d'un articulographe (AG100 de Carstens) 10 canaux.
- Étude des perturbations pathologiques de l'articulation et du système de perception : étude acoustique de locuteurs ayant subi une exérèse de la cavité endobuccale avec reconstruction (Projet ACI Cognitive du CNRS).
- Gestion du parc informatique PC du site ICP Campus.

5 publications les plus récentes

- BEAUTEMPS D., CATHIARD M-A., ATTINA V., **SAVARIAUX C.** & ARNAL A. (2006) Temporal organization of cued speech production. In Bailly, Bateson & Perrier (Eds.).
- ABRY C., CATHIARD MA., LABOISSIÈRE R., LOEVENBRUCK H., **SAVARIAUX C.**, SCHWARTZ JL. & VILAIN A. (2006) Some insights in bimodal perception given for free by the natural time course of speech production. In Bailly, Bateson & Perrier (Eds.).
- BAILLY G., ELISEI F., BADIN P. & **SAVARIAUX C.** (2006) Degrees of freedom of facial movements in face-to-face conversational speech. To be published in the *LREC 2006 Workshop on Multimodal Corpora*, Genoa, Italy.
- ZMARICH C., GILI FIVELA B., PERRIER P. & **SAVARIAUX C.** (2006) Acoustic and kinematic correlates of phonological length contrast in Italian consonants. To be published in the *10th Conference on Laboratory Phonology*, Paris, France.
- SCHWARTZ J.L., BERTHOMMIER F. & **SAVARIAUX C.** (2004) Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition*, Volume 93, B69-B78.

5 publications les plus significatives

- LEBEAU J., **SAVARIAUX C.**, PERRIER P., BETTEGA G. & RAPHAEL B. (2000) Evaluation fonctionnelle des reconstructions endobuccales. Un outil intéressant : l'interprétation articulatoire du signal acoustique. *Rev. Stomato. Chir. Maxillofac.*, 101, Vol. 2, pp 60-64.
- SAVARIAUX C.**, PERRIER P., ORLIAGUET J.P & SCHWARTZ J.L. (1999) Compensation for the perturbation of French [u] using a lip tube: II. Perceptual analysis. *Journal of the Acoustical Society of America*, 106 (1), 381-393.
- SAVARIAUX C.**, PERRIER P. & ORLIAGUET J.P. (1995) Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube: A study of the control space in speech production. *Journal of the Acoustical Society of America*, 98 (5), 2428-2442.
- MÉNARD L., PERRIER P. & **SAVARIAUX C.** (2004) Exploring production-perception relationships for 4-year-old children: a study of compensation strategies to a lip-tube perturbation. [ASA Meeting 2004](#), 5pSC11. New York, USA.

<sup>3</sup> En application de la décision du Conseil Scientifique de l'INPG du 24 mars 1994, la mention "Très Honorable avec Félicitations" est attribuée à l'unanimité du jury à l'issue d'un vote à bulletin secret.

**SAVARIAUX C.**, PERRIER P., PAPE D. & LEBEAU J. (2001) Speech production after glossectomy and reconstructive lingual surgery: a longitudinal study. *Proceedings of the 2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA)*. Firenze, Italy.

Acronyme ou titre court du projet	<b>DOCVACIM</b>
-----------------------------------	-----------------

**A-2 : Autres partenaires du projet (remplir une fiche par partenaire)**

*Un responsable scientifique de l'équipe partenaire doit être désigné*

**Partenaire 3 : LORIA**

**\* champ obligatoire**

<b>Civilité *</b>	<b>Nom *</b>	<b>Prénom *</b>	
Monsieur	LAPRIE	Yves	
<b>Grade*</b>	CR1	<b>Employeur *</b>	CNRS
<b>Mail *</b>	Yves.Laprie@loria.fr		
<b>Tél *</b>	03 83 59 20 36	<b>Fax</b>	03 83 27 83 19

<b>Laboratoire *</b> (nom complet)	
Laboratoire Lorrain de Recherche en Informatique et ses Applications	
<b>N° Unité</b> (s'il existe)	UMR 7503
<b>Adresse complète du laboratoire *</b>	
LORIA - Campus Scientifique - BP 239	
<b>Code postal *</b>	54506
<b>Ville *</b>	Vandoeuvre-lès-Nancy Cedex
<b>Établissements de tutelle</b> (indiquer le ou les établissements et organismes de rattachement, souligner l'établissement susceptible d'assurer la gestion du projet) :	

**Principales publications :**

Liste des 10 principales publications ou brevets de l'équipe partenaire 2 (définie tableau ci-dessous) au cours des cinq dernières années, relevant du domaine de recherche couvert par la présente demande dans l'ordre suivant : Auteurs (en soulignant les auteurs faisant effectivement partie de la demande), Année, Titre, Revue, N°Vol, Pages. N'indiquez pas les publications soumises.

- YVES LAPRIE, SLIM OUNI. – « Introduction of constraints in an acoustic-to-articulatory inversion ». – In : 7th International Conference on Spoken Language Processing - ICSLP 2002, Denver, USA. – septembre 2002.

- YVES LAPRIE, ANNE BONNEAU. – « A copy synthesis method to pilot the Klatt synthesiser ». – In : *International Conference on Speech and Language Processing, Denver, USA.* – septembre 2002.
- SLIM OUNI, YVES LAPRIE. – « A study of the French Vowels Through The Main Constriction of the Vocal Tract Using an Acoustic-to-articulatory inversion method ». – In : *15th International Congress of Phonetic Sciences 2003 - ICPHS'2003, Barcelone, Espagne.* – aug 2003.
- YVES LAPRIE, SLIM OUNI, BLAISE POTARD, SHINJI MAEDA. – « Inversion experiments based on a descriptive articulatory model ». – In : *6th International Seminar on Speech Production.* – Sydney, Australia, décembre 2003.
- JINGYING CHEN, MARIE-ODILE BERGER, YVES LAPRIE. – « An Effective Lip Tracking Algorithm for Acoustic-to-Articulatory Inversion ». – In : *5th International Workshop on Image Analysis for Multimedia -WIAMIS'2004, Lisbon, Portugal.* – apr 2004.
- YVES LAPRIE. – « A concurrent curve strategy for formant tracking ». – In : *Interspeech 2004 – International Conference on Spoken Language Processing, Jeju, Corée du sud.* – oct 2004.
- POTARD B., LAPRIE Y. – « Using phonetic constraints in acoustic-to-articulatory inversion ». – In : *Interspeech, Lisboa.* – septembre 2005.
- V. ROBERT, B. WROBEL-DAUTCOURT, Y. LAPRIE, A. BONNEAU. – « Inter-speaker variability of labial coarticulation with the view of developing a formal coarticulation model for French ». – In : *Proceedings of International Conference on Auditory-Visual Speech Processing (AVSP'05)*, pp. 65–70. – Vancouver, 2005.
- S. Ouni et Y. Laprie, *Modeling the articulatory space using a hypercube codebook for acoustic-to-articulatory inversion*, Journal of the Acoustical Society of America (JASA). vol. 118, p.444--460. (2005)
- MAJ J.-B., BONNEAU A., FOHR D., LAPRIE Y. – « An elitist approach for extracting automatically wellrealized speech sounds with high confidence ». – In : *Interspeech, Lisboa.* – septembre 2005.
- B. WROBEL-DAUTCOURT, M. O. BERGER, B. POTARD, Y. LAPRIE, S. OUNI. – « A based system for acquisition of visible articulatory data ». – In : *Proceedings of International on Auditory-Visual Speech Processing (AVSP'05)*, pp. 145–150. – Vancouver, 2005.

### Partenaire 3 : LORIA

	Nom	Prénom	Emploi actuel	Discipline (à renseigner uniquement pour SHS)	% de temps de recherche consacré au projet	Rôle/Responsabilité dans le projet 4 lignes max
<i>exemple</i>	<i>MARTIN</i>	<i>Charlotte</i>	<i>Professeur</i>		<i>30%</i>	
Responsable	LAPRIE	Yves	CR1		20%	Coordination du projet ; exploitation des données cinéroradiographiques pour l'inversion acoustique articulaire ; élaboration d'un modèle de contrôle
Membres de l'équipe	OUNI	Slim	MCF		20%	Exploitation des données cinéroradiographiques pour l'inversion acoustique articulaire ; élaboration d'un modèle de contrôle

Pour chacun des membres de l'équipe du projet, fournir une biographie **d'une page maximum** qui comportera :

A/ Nom, prénom, âge, cursus, situation actuelle

B/ Autres expériences professionnelles

C/ Liste des 10 publications (ou brevets) comprenant les 5 publications les plus récentes et les 5 publications les plus significatives

D/ Prix, distinctions

## Biographie

**Yves Laprie** (44 ans) est actuellement chargé de recherche (CR1) au CNRS.

Il a obtenu le diplôme d'Ingénieur civil des mines en 1985 à Nancy, le Doctorat en informatique en 1990 et l'habilitation à diriger des recherches en 2003.

Yves Laprie est le responsable de l'équipe Parole (28 personnes en 2005) au LORIA et a été membre du bureau du GFCP, puis de l'Association Francophone de la Communication Parlée entre 1998 et 2005. Ses thèmes de recherche sont l'inversion acoustique articulatoire (c'est-à-dire la récupération de l'évolution temporelle de la forme du conduit vocal), l'analyse de la parole et la recherche d'indices acoustiques caractéristiques des sons de la parole.

L'approche de l'inversion acoustique articulatoire développée dans le cadre des thèses de Bruno Mathieu, Slim Ouni et Blaise Potard, repose sur la construction d'une table articulatoire qui assure une très bonne résolution articulatoire et acoustique. Cette approche permet d'introduire facilement des contraintes pour réduire l'indétermination du problème.

Yves Laprie a aussi développé un logiciel d'analyse de la parole (WinSnoori) à partir des algorithmes d'analyse qu'il a élaborés (détection de la fréquence fondamentale, suivi automatique de formants, synthèse par copie pour un synthétiseur à formants, transformation du débit ou de la fréquence fondamentale de signaux de parole à l'aide de PSOLA).

Yves Laprie est actuellement le coordinateur du projet Européen ASPI sur l'inversion audiovisuelle articulatoire.

Quelques publications choisies :

1. « Un cadre articulatoire pour la reconnaissance automatique de la parole ». – *In : La reconnaissance de la parole : du signal à son interprétation*, J.P. Haton, C. Cerisara, D. Fohr, Y. Laprie, et K. Smaïli (éd.). Dunod, Paris, 2006.
2. A. OUNI, Y. LAPRIE. – « Modeling the articulatory space using a hypercube codebook for acoustic-toarticulatory inversion ». – *Journal of the Acoustical Society of America* 118, 1 (2005), pp. 444–460
3. A. BONNEAU, L. DJEZZAR, Y. LAPRIE. – « Perception of the Place of Articulation of French Stop Bursts ». – *Journal of the Acoustical Society of America* 100, 1 (1996), pp. 555–564.
4. YVES LAPRIE. – « A concurrent curve strategy for formant tracking ». – *In : Interspeech 2004 – International Conference on Spoken Language Processing, Jeju, Corée du sud.* – oct 2004.
5. Y. LAPRIE, M.-O. BERGER. – « Cooperation of Regularization and Speech Heuristics to Control Automatic Formant Tracking ». – *Speech Communication* 19, 4 (octobre 1996), pp. 255–269.
6. Logiciel WinSnoori (déposé à l'APP en 1999, la dernière version date de novembre 2005) sur l'analyse de la parole (<http://www.winsnoori.fr>)
7. POTARD B., LAPRIE Y. – « Using phonetic constraints in acoustic-to-articulatory inversion ». – *In : Interspeech, Lisboa.* – septembre 2005.
8. MAJ J.-B., BONNEAU A., FOHR D., LAPRIE Y. – « An elitist approach for extracting automatically well realized speech sounds with high confidence ». – *In : Interspeech, Lisboa.* – septembre 2005.
9. V. ROBERT, B. WROBEL-DAUTCOURT, Y. LAPRIE, A. BONNEAU. – « Inter-speaker variability of labial coarticulation with the view of developing a formal coarticulation model for French ». – *In : Proceedings of International Conference on Auditory-Visual Speech Processing (AVSP'05)*, pp. 65–70. – Vancouver, 2005.
10. ROBERT V., WROBEL-DAUTCOURT B., LAPRIE Y., BONNEAU A. – « Strategies of labial coarticulation ». – *In : Interspeech, Lisboa.* – septembre 2005.

**Slim Ouni** (32 ans) est actuellement Maître de Conférences à l'IUT Nancy-Charlemagne, Université de Nancy 2. Il a obtenu un doctorat en informatique de l'Université Henri Poincaré à Nancy en 2001. Ses travaux de thèse portent sur l'étude de l'inversion acoustique-articulatoire. En 2002, il a rejoint l'équipe de

recherche de *Dominic Massaro* (renommé pour ces travaux en perception de la parole et têtes parlantes) à l'université de Californie à Santa Cruz, aux Etats-Unis. Pendant deux ans, il a mené au sein de cette équipe des recherches dans le domaine de la parole audiovisuelle (analyse, perception et synthèse des têtes parlantes). Les domaines de recherche aux quels il s'intéresse sont la production de la parole (modélisation du conduit vocal, l'inversion acoustique-articulaire, phonétique articulaire), la parole audiovisuelle (analyse, perception et synthèse, modélisation de la coarticulation, intelligibilité de la parole) et l'apprentissage des langues (acquisition et apprentissage d'une deuxième langue).

#### **Quelques Publications choisies :**

1. Ouni, S. & Laprie, Y. (2005). *Modeling the articulatory space using a hypercube codebook for acoustic-to-articulatory inversion*. J. Acoust. Soc. Am., July 2005 – V. 118(1), pp. 444-460.
2. Ouni, S., Cohen, M. M., & Massaro, D. W. (2005). *Training Baldi to be multilingual: A case study for an Arabic Badr*. Speech Communication, 45(2), 115-137.
3. Ouni, S. (2005). *Can We Retrieve Vocal Tract Dynamics that Produced Speech? Toward a Speaker Articulatory Strategy Model*. Interspeech 2005, 4-8 Septembre 2005, Lisbonne, Portugal.
4. Ouni, S., Cohen, M.M., Ishak, H. & Massaro, D.W. (2005). *Visual Contribution to Speech Perception: Measuring the Intelligibility of Talking heads*. AVSP'05, 24-27 Juillet 2005, British Columbia, Canada.
5. Ouni, S. & Laprie, Y. (2003). *A study of the main constriction of the vocal tract for French vowels using an acoustic-to-articulatory inversion method*. International Congress of Phonetic Sciences, Août 2003, Barcelona, Espagne.
6. Ouni, S., Laprie, Y. (2001). *Exploring the null space of the acoustic-to-articulatory inversion using a hypercube codebook*. Eurospeech'2001, Aalborg, Danemark.
7. Ouni, S., Laprie, Y. (2001). *Studying articulatory effects through hypercube sampling of the articulatory space*. International Congress on Acoustics, Rome, Italie.
8. Ouni, S., Laprie, Y. (1999). *Design of Hypercube Codebooks for the Acoustic-to-Articulatory Inversion Respecting the Non-Linearities of the Articulatory-to-Acoustic Mapping*. Eurospeech'99, Budapest, Hongrie.

# Programmes SHS 2006

## B - Description du projet

Acronyme ou titre court du projet : <i>DOCVACIM</i>
---

*Les objectifs, l'originalité du projet, la problématique, les méthodologies employées et les modalités d'accès aux terrains, le programme des travaux et ses différentes phases, la bibliographie et l'état de l'art, les modalités de valorisation des connaissances doivent être présentées. Les modalités de mise en œuvre de l'interdisciplinarité éventuelle et des diverses collaborations doivent être précisées et justifiées en accord avec l'orientation du projet. Les modalités de coordination et de travail en commun des différents partenaires doivent être décrites.*

*La capacité de ou des équipes « porteuse(s) » doit être attestée par la qualification et les productions scientifiques antérieures de leurs membres. Leurs rôles dans les différentes phases du projet doivent être précisés et la valeur ajoutée des collaborations entre les différentes équipes sera argumentée. Les moyens demandés doivent être justifié au regard des objectifs scientifiques du projet et du programme des travaux.  
(Arial 11, simple interligne)*

mardi 9 mai 2006

### PROJET SCIENTIFIQUE

#### **B-1 – Enjeux et objectifs, contexte, problématique, originalité :**

*Données Cinéradiographiques Valorisées et recherches sur la Coarticulation, l'Inversion et l'évaluation de Modèles physiques*

#### *Objectifs et contexte*

Ce projet de recherche, en production-perception de la parole, a pour objectif général de contribuer à la compréhension et à la modélisation du contrôle des gestes de la parole, par l'analyse de la *coarticulation*, l'étude de l'*inversion* et l'évaluation de *modèles physiques*.

La production de la parole nécessite une coordination précise des articulateurs dans un espace donné ; le chevauchement gestuel ou la *coarticulation* se fait selon des contraintes spécifiques aux sons, aux langues et aux locuteurs. Cet aspect des recherches sur l'organisation spatio-temporelle des gestes linguistiques est donc central aux travaux en production-perception de la parole (cf. axe de recherche 1).

Appréhender les dimensions perceptives – auditive et visuelle – de la parole requiert une bonne maîtrise de la relation entre la sortie acoustique et l'origine articulatoire des sons. L'*inversion* articulatoire consiste ainsi à retrouver l'évolution de la forme du conduit vocal à partir du signal de parole (cf. axe de recherche 2).

La compréhension des mécanismes de contrôle de la production de la parole passe, selon nous, par le développement et l'*évaluation de modèles* du contrôle moteur et de modèles physiques de l'appareil de production de la parole. C'est la comparaison de données mesurées sur des locuteurs,

dans des conditions bien contrôlées avec les prédictions proposées par ces modèles, qui permet d'évaluer, de manière quantitative et objective, les différentes hypothèses proposées dans la littérature sur le contrôle de la production de la parole (cf. axe de recherche 3).

Pour étudier le contrôle moteur des gestes humains en général, ainsi que leur efficacité perceptive, le substrat de base est constitué d'une quantité très importante de données physiologiques, cinématiques, acoustiques et éventuellement dynamiques, collectées sur des locuteurs humains. Cette étape de collecte et d'analyse de données est indispensable ; elle a constamment progressé au cours des vingt dernières années, cela grâce au développement et à l'utilisation de techniques de mesure efficaces et sophistiquées, telles que l'endoscopie, la transillumination glottique, l'électropalatographie, l'électromagnétométrie bi- ou tri-dimensionnelle, l'imagerie par Résonance Magnétique ou encore les micro-capteurs de pression mécanique.

L'intérêt d'une telle démarche d'investigation expérimentale est indéniable, puisqu'elle a permis de déterminer les corrélats physiques, articulatoires et acoustiques essentiellement, des sons de la parole et de leurs coarticulations. Ces travaux ont ainsi fait avancer de manière significative les débats sur les relations entre phonologie et phonétique (voir en particulier : Perkell & Klatt, 1986 ; Stevens: 1972, 1989), ainsi que la typologie des langues (Ladefoged & Maddieson, 1996), ou encore, plus récemment, la synthèse de visages parlants (Elisei *et al.*, 2001 ; Badin *et al.*, 2002, Engwall, 2003).

Malgré ces avancées dans les recherches en Sciences de la Parole, le *manque de données cinéradiographiques* se fait cruellement ressentir dans la communauté scientifique, préoccupée par l'étude de la production et de la perception de la parole. En effet, elles sont *les seules* à offrir actuellement, en même temps, une résolution spatio-temporelle correcte, sur l'ensemble du conduit vocal, dans le plan sagittal.

Or, l'acquisition de nouvelles données cinéradiographiques est aujourd'hui restreinte, voire interdite dans certains pays.

Notre projet de recherche sur le contrôle moteur des gestes humains en général, ainsi que sur leur efficacité perceptive vise, en conséquence, à valoriser et à diffuser une partie des nombreux films cinéradiographiques (ou radiofilms) sur la production et la perception de la parole qui a été réalisée à l'Institut de Phonétique de Strasbourg (IPS), depuis la fin des années '50. Le programme concerne une vingtaine de films de très bonne qualité, portant sur des problématiques linguistiques de langues parlées en Europe, en Afrique, en Asie et en Amérique Latine.

Il est organisé autour de 4 objectifs principaux, trois spécifiques et un général (les axes de recherche) :

- /1/ la *sauvegarde* de ces données cinéradiographiques patrimoniales ;
- /2/ leur *traitement* (grâce au développement d'outils) ;
- /3/ leur *valorisation* par l'intégration dans une base, l'organisation d'ateliers et de colloques ;
- /4/ et leur *analyse*, ainsi que leur *exploitation* pour des objectifs de recherche fondamentale sur la production de la parole (la *coarticulation*, l'étude de l'*inversion* et l'*évaluation de modèles physiques de la production de la parole*).

*Objectif /1/* : La sauvegarde des radiofilms sera assurée par l'Institut de Phonétique de Strasbourg (IPS) – Composante Parole et Cognition de l'E.A. 1339 et l'Institut de la Communication Parlée de Grenoble (ICP).

*Objectif /2/* : Le traitement des données, ainsi que le développement d'outils seront pris en charge par l'IPS, l'ICP et le Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA) – Groupe Parole.

*Objectif /3/* : La valorisation des données par l'intégration dans une base se fera par l'IPS, l'ICP et la Maison Inter-universitaire des Sciences de l'Homme d'Alsace (MISHA).

*Objectif /4/* : L'analyse des données en vue de l'évaluation de modèles physiques, ainsi que l'étude de l'inversion et de la coarticulation représenteront les principaux axes de recherche, en toile de fond du

projet, qui préoccupent les 3 laboratoires partenaires travaillant sur la production et la perception de la parole, à savoir l'IPS, l'ICP et le LORIA.

Durant le projet, les 3 laboratoires de parole partenaires et la MISHA procéderont à l'organisation d'ateliers et de manifestations scientifiques, ainsi qu'à des publications communes.

À l'issue du projet, nous mettrons à la disposition de la communauté scientifique :

/1/ un ensemble de données *multilingues* et multimédia, *unique au monde*, sur la production et la perception de la parole, comprenant images cinéradiographiques du conduit vocal, signal acoustique et tracés sagittaux des contours du conduit vocal, le tout synchronisé, qui sera accessible au sein d'une plateforme de traitement ;

/2/ des outils et des logiciels d'exploitation adaptés, permettant d'extraire les informations linguistiques de ces données ;

/3/ et, bien entendu, les résultats de nos productions scientifiques, sous forme d'ouvrages, d'articles, de publications électroniques, etc.

### **Problématique, originalité et l'état de l'art**

Aujourd'hui, les données cinéradiographiques sont encore d'une grande utilité. Ce sont les seules à offrir actuellement, en même temps, une résolution spatio-temporelle correcte sur l'ensemble du conduit vocal dans le plan sagittal. Elles sont à la base de l'élaboration de modèles géométriques ((Maeda, 1989 ; Laprie & Berger, 1995) et sont d'une grande utilité pour l'étude de la coordination spatio-temporelle des articulateurs de la parole (Wood, 1979 a & b ; Sock, 1998 ; Vaxelaire & Sock, 1999), ainsi que pour l'étude de l'inversion (Bailly *et al.* , 1992 ; Laprie & Ouni, 2002 ; Laprie *et al.*, 2003). La législation interdit l'acquisition de nouvelles données de ce type pour des questions d'éthique (Loi Hurriet 1988, révisée en 2000). Or de nombreuses données existent dans différents laboratoires. Il nous semble de la plus haute importance d'en faciliter l'accès à l'ensemble de notre communauté scientifique<sup>4</sup>.

Munhall et ses collègues (Munhall *et al.*, 1995) ont fait un magnifique travail de sauvegarde et de distribution de données cinéradiographiques, réalisées en Amérique du Nord, et essentiellement à l'Université Laval de Québec.

En France, nous bénéficions d'une situation tout à fait exceptionnelle. L'Institut de Phonétique de Strasbourg a en effet accumulé, depuis la fin des années 50, sous l'impulsion de Georges Straka, puis de Péla Simon, plus de 50 enregistrements cinéradiographiques, et cela sur un ensemble très large de langues.

Nous avons donc entrepris, en collaboration avec l'Institut de la Communication Parlée de Grenoble, un travail de mise en forme de ces données avec les objectifs suivants :

- 1 • assurer la sauvegarde des données (actuellement sur films 35mm et bandes audio) par stockage sur un support vidéo de haute qualité, et par numérisation et stockage sur DVDROM ou disque dur externe ;
- 2 • apporter une valeur ajoutée par le biais de tracés sagittaux réalisés par des experts phonéticiens, et montrant les limites du conduit vocal ;
- 3 • faciliter l'accès et le traitement de ces données par leur intégration dans une base de données, et la distribution de cette base.

### *Résultats obtenus*

Quatre films d'une durée de quelques minutes, et associé chacun à environ 550 tracés radiographiques, ont été ainsi traités et entrés dans la base de données. Les corpus de ces films sont centrés sur les questions suivantes :

*L'effet de jointure en français* (2 films de François Wioland) ;

*Les consonnes occlusives du français* (film de Jean-Pierre Zerling) ; et

---

<sup>4</sup> Avis du Ministère de la Recherche, Contractualisation 2001 – 2004 de l'E.A. IPS : « [...] Le projet, pertinent est porté par une petite équipe de bonne réputation. Il s'agit de valoriser 33 radiofilms cinéradiographiques réalisés sous rayons X [...] en les organisant en plateforme multimédia. Le projet est d'un réel intérêt pour la communauté scientifique. [...] »

*Les nasales du français* (film de Bernard Flament).

Nous avons pu commencer à développer une base de données cinéradiographiques (Arnal *et al.*, 2000) dont l'accès est simple et peu onéreux. Par rapport aux autres bases de données similaires existant par ailleurs, notre travail a la particularité d'associer aux données brutes des *tracés sagittaux*. C'est un point extrêmement positif, car la lecture de radiographies du conduit vocal est loin d'être aisée pour des personnes non expertes.

Il reste encore de nombreux films, de grande qualité, dans les armoires de l'IPS. Nous aurons donc à cœur de poursuivre ce travail que le <sup>5</sup>CNRS a permis d'initier.

La suite du projet a également été soutenue par un Programme <sup>6</sup>ACI du Ministère de la Recherche et des Nouvelles Technologies.

Cependant, le travail systématique de sauvegarde de ces données, ainsi que la valeur ajoutée que constituent la réalisation des tracés dans le plan sagittal, et l'intégration de l'ensemble dans une base de données sont extrêmement coûteux, aussi bien en équipement qu'en main d'œuvre, et nécessitent également une grande expertise dans le domaine.

Il nous est donc indispensable de trouver de nouvelles sources de <sup>7</sup>financement important permettant de poursuivre ce travail de sauvegarde, de numérisation, d'étiquetage et de valorisation de ces données cinéradiographiques.

En outre, les interactions multiples entre les laboratoires partenaires, dans le cadre de <sup>8</sup>contrats communs, ou dans celui de collaborations de recherche plus informelles, permettent d'envisager de développer, à très court terme, de premiers outils d'extraction automatique des contours sagittaux du conduit vocal, à partir des films radiologiques. Le développement de ces outils se fera parallèlement à la réalisation manuelle d'un certain nombre de tracés de ces contours par des experts, et on les exploitera comme base de référence. De tels traitements automatiques faciliteront l'exploitation systématique des données cinéradiographiques, que ce soit pour le développement de modèles géométriques du conduit vocal, pour la mise au point de méthodes d'inversion du signal de parole (de l'acoustique vers les commandes articulatoires), ou pour l'étude des phénomènes coarticulatoires.

Notre projet s'inscrit ainsi dans la continuité de notre démarche générale de recherche sur la production et la perception de la parole, dans la tradition transdisciplinaire propre aux Sciences de la Parole : de l'exploitation de systèmes d'acquisition numériques et analogiques hautement performants, à l'élaboration de modèles des contraintes physiques et spatio-temporelles, en passant par la mise au point d'outils pour le traitement automatique des données articulatoires, le tout pour une meilleure connaissance du code linguistique des différentes langues étudiées.

## **B-2 – Description du projet et résultats attendus : (12 pages maximum)**

### *De l'utilité des données cinéradiographiques*

---

<sup>5</sup> Le projet, qui réunissait l'ICP et l'IPS, avait été soutenu par le *Programme Ingénierie des Langues du CNRS, Dépts. SHS/SPI* « Production, validation et mise à disposition de données et d'outils linguistiques », 1997 - 2000.

<sup>6</sup> Programme « Action Concertée Incitative, ACI- TTT Données et Mesures, du Ministère de la Recherche et des Nouvelles Technologies 2003 – 2006. Projet entre l'IPS et l'ICP.

<sup>7</sup> Avis du Ministère de la Recherche, Contractualisation 2005 – 2008 de l'E.A. IPS : « Le projet de cette EA s'inscrit globalement dans la conservation de données relevant du patrimoine linguistique mais aussi dans le renouvellement de leur traitement et de leurs possibilités exploratoires. Le CNRS avait permis au projet de débiter : il semble utile de pouvoir le mener complètement à son terme afin que l'ensemble des données bénéficient du traitement. L'importance accordée à la gestique, trop souvent délaissée pour des raisons notamment matérielles et à cause des difficultés de saisie, est ici centrale. »

<sup>8</sup> a) Programme scientifique MISHA UMS 2552, 2005 – 2008 « Anticiper pour construire des modalités sémiotiques » qui réunit l'IPS et l'ICP, entre autres laboratoires de recherche. b) Contrat Européen Sixth Framework Programme, Priority 2 FET Open « Audiovisual to Articulatory Speech Inversion – ASPI, attribué au LORIA, l'IPS étant en sous-traitance.

L'avancée des travaux sur l'étude et la *modélisation* du processus humain de la production et de la perception de la parole passe nécessairement par l'étude du contrôle de l' « espace » et du « temps » en parole. Depuis le premier grand ouvrage de Fant (1960), dans le domaine de la production de la parole, l'intérêt pour l'étude des représentations phonétiques, abordées par la *radiocinématographie* ne s'est pas arrêté avec les percées théoriques et techniques accomplies dans ce domaine (cf. par ex. Heinz & Stevens, 1964 ; Perkell, 1969 ; Wood, 1979a & b, Maeda, 1988 ; Perrier *et al.* 2003, pour une application rentable dans le domaine de la modélisation des représentations phonétiques).

Cet intérêt s'est poursuivi et a été réactualisé, dans le cadre de la *phonétique générale et expérimentale* par Bothorel, Simon, Wioland & Zerling (1986), Maeda (1989) et Perrier *et al.* (1992), entre autres. Le problème central est celui d'établir le lien entre le niveau des représentations phonologiques, celui du code linguistique, et les signaux physiques moteur, articuloire et/ou acoustique. La modélisation articuloire, acoustique et perceptive, guidée par des théories et des méthodologies adaptées, correspond à un des moyens mis en œuvre pour apporter des éléments de réponse à ce problème.

Les données cinéradiographiques ont longtemps représenté la source de connaissances la plus complète sur les processus articuloires, à l'origine de la production des sons de la parole. Associant sons et images du conduit vocal dans le plan sagittal, elle livre des renseignements audiovisuels bruts et essentiels sur la géométrie du conduit vocal (le « contrôle de l'espace ») et le mouvement des diverses structures articuloires (le « contrôle du temps »), synchrones avec le signal acoustique. Ces données renseignent non seulement sur la position individuelle des différents articulateurs de la parole, mais aussi sur la forme globale des cavités supraglottiques responsables, en tant que résonateurs, de l'amplification et de la modulation du ton laryngien (activité des cordes vocales). Depuis la fin des années 1980, l'imagerie par Résonance Magnétique (IRM) offre des perspectives supérieures à la cinéradiographie pour tout ce qui concerne la description spatiale du conduit vocal, en particulier parce qu'il est possible de visualiser, par cette technique, le conduit vocal dans sa globalité tri-dimensionnelle. Cependant, avec une cadence maximale de l'ordre de 10 images par seconde, l'IRM n'offre pas encore le degré de résolution temporelle offert par la cinéradiographie (50 images/seconde).

Les données cinéradiographiques constituent donc, à ce jour, une source d'information encore inégalée sur les relations entre l'évolution temporelle des cavités supraglottiques et celle du signal acoustique de parole. Nous savons que les relations entre le domaine articuloire (positionnement des articulateurs et forme des cavités supraglottiques) et le domaine acoustique sont complexes : elles sont caractérisées par une forte non-linéarité (Stevens, 1972) et par une non-biunivocité, différentes dispositions articuloires pouvant être à l'origine d'une même réalisation acoustique (Atal *et al.*, 1978). Cette caractéristique non linéaire et non-biunivoque constitue un des obstacles majeurs auxquels nous nous heurtons, lorsque nous cherchons à trouver les corrélats articuloires d'une séquence de parole, à partir du seul signal acoustique. L'analyse simultanée de données articuloires et de données acoustiques permet l'observation de régularités et donc l'élaboration de *contraintes* qui, pour un son donné, réduisent l'espace possible des formes articuloires associées, en précisant les points critiques pour l'émergence des objets phonologiques - cognitifs (Baer *et al.*, 1988 ; Boë *et al.*, 1992 ; Perrier *et al.*, 1992). Les régularités observées, en passant d'un niveau d'analyse à l'autre, permettent en effet de restreindre les systèmes comportant des degrés de liberté en excès, en mettant en relief les coordinations fortement contraintes. Nous pensons contribuer de cette manière aux recherches préoccupées par la notion d'*inversion* en parole qui, dans le sens mathématique du terme, est un problème mal-posé (*ill posed*), puisqu'il peut exister plusieurs solutions à un problème unique. La mise au jour de régularités articuloire-acoustiques prend donc toute sa place dans des recherches *multilingues* en production-perception de la parole.

Il importe aussi de tester la flexibilité de l'appareil vocal humain par la *perturbation naturelle* du système de production de la parole. La perturbation de ce système permet, en effet, d'analyser les différentes stratégies cognitivo-motrices, utilisées par les sujets parlants pour maintenir la distinctivité phonétique des catégories phonologiques dans des conditions difficiles, en mettant en jeu divers phénomènes de *compensation*. Certaines de nos données cinéradiographiques permettent d'évaluer l'impact d'une des perturbations naturelles des configurations du conduit vocal en mouvement : l'augmentation de la *vitesse d'élocution* (Vaxelaire & Sock, 2000).

En outre, il sera possible aussi bien dans le cadre du projet que pour les chercheurs qui auront accès à ces données d'étudier la *coarticulation* (anticipatoire et rétentrice), sous toutes ses formes, et cela dans une perspective multilingue permettant de prendre en compte les contraintes intrinsèques du système phonologique de chaque langue, grâce à la possibilité d'analyser les différences et les

ressemblances de stratégies coarticulatoires, selon la grande *variété de langues* disponible dans la base. L'importance de la *comparaison interlangue* des structures articulatoires et acoustiques n'est plus à démontrer. Il devient ainsi envisageable de « naviguer » entre les différentes réalités linguistiques pour déterminer les facteurs spécifiques à une langue donnée, et ceux partagés par les langues en général, ces derniers reflétant les contraintes bio-mécaniques du système de production-perception humain.

L'aspect multilingue du corpus, comprenant des *langues dites rares* (cf. *infra*), permet en conséquence, sans aucun doute, l'extension et la généralisation des acquis actuels, pour des langues connues, dans le domaine de l'étude de la production-perception de la parole.

Il est clair que le cadre théorique qu'offre la diversité linguistique de la base est suffisamment large pour motiver un spectre très étendu des travaux actuels en production et en perception de la parole.

## Sauvegarde et numérisation des données cinéradiographiques : attendus scientifiques

### *Nature des radiofilms concernés par le projet scientifique*

Les données cinéradiographiques disponibles à l'Institut de Phonétique de Strasbourg revêtent un caractère unique au monde, tant par leur nombre (près de 50 films dont *une vingtaine est de très bonne qualité*) que par le spectre linguistique qu'elles couvrent : plus d'une dizaine de langues de familles linguistiques très variées ont été en effet analysées par ce biais.

Les continents, les pays et les langues concernés par ces données sont les suivants :

#### Afrique

- Algérie (3 variantes dialectales de l'arabe)
- Maroc (arabe classique et 2 variantes dialectales)
- Côte d'Ivoire (ogni-sanvi)
- Ghana (akan)
- Sénégal (ndut-sereer ; wolof)
- Tanzanie (kiswhahili)
- Zambie (lozi)

#### Amérique Latine

- Brésil (portugais)

#### Asie

- Syrie (arabe classique)
- Turquie (2 variantes dialectales du turc)
- Corée du Sud (coréens : 2 films)
- Thaïlande (thaï : 3 films)

#### Europe

- Allemagne (allemand)
- Bulgarie (bulgare)
- Espagne (espagnol)
- France (français - environ 20 films ; breton)
- Hongrie (hongrois)
- Islande (islandais)

Tous ces films ont été partiellement traités à la main par des experts phonéticiens qui, pour un certain nombre d'images radiographiques, ont dessiné un tracé précis des contours du conduit vocal dans le plan médian (tracé sagittal) et, dans certains cas, un tracé du pavillon labial. Toutefois, à raison d'une moyenne de 3000 images par films, on comprend aisément que toutes les données n'aient pas pu être traitées. Le développement de *techniques d'extraction automatique* des contours du conduit vocal est donc un aspect incontournable et central à notre projet. Nous y reviendrons plus bas.

En outre, les tracés sagittaux existants sont stockés sur des supports papiers ; ils sont difficilement accessibles à l'ensemble de la communauté scientifique travaillant sur la parole et susceptible de les utiliser.

## **Méthodologie**

*Les étapes du travail seront les suivantes :*

/1/ Sélection des films sur la base de leur qualité et de leur corpus

/2/ Repérage sur chacun des films des séquences pertinentes par rapport au corpus

Certains corpus incluent en effet les séquences, certes intéressantes, dans des phrases porteuses, mais pour lesquelles la réalisation de tracés radiographiques n'est pas indispensable.

/3/ Inventaire des tracés déjà existants pour ces séquences

Ces films ont été réalisés pour des travaux de thèse et ont donc été, le plus souvent, déjà partiellement exploités. Cependant, les tracés ne couvraient jamais la totalité des séquences, et, de plus, un certain nombre d'entre eux ne nous étaient pas accessibles. Il sera donc indispensable de réaliser un grand nombre de nouveaux tracés, tout en s'assurant de la validité de ceux qui existent déjà, et de la cohérence entre nouveaux et anciens tracés.

/4/ Numérisation des tracés

Cela se fait sous forme d'une numérisation simple par scanner. Le tracé permet de distinguer différentes sous-parties du contour, pertinentes du point de vue articulatoire, telles que les lèvres, le palais, le velum, la langue ou la mandibule. Les informations ainsi fournies par ce tracé seront exploitables pour toute analyse articulatoire de ces données, telle que, par exemple, une analyse statistique visant à l'élaboration d'un modèle géométrique du conduit vocal.

/5/ Transfert des films du standard cinématographique 35 mm vers le standard vidéo BetacamSP

Il s'agit d'un standard professionnel de haute qualité, assurant une préservation optimale de la précision des images. Nous disposons de deux types de films, qui diffèrent par la vitesse d'acquisition des images ; celle-ci était soit de 64 images/seconde, soit de 50 images/seconde. La vidéo restitue les images à 25 images/seconde, sous forme de trames entrelacées à 50 trames/sec. La technique que nous avons choisie, toujours dans le souci de préserver la qualité des enregistrements originaux, consiste à utiliser une image vidéo complète pour chaque image originale (préservation de la définition spatiale) et à conserver toutes les images (préservation de la définition temporelle). Cela entraîne évidemment un ralentissement de la vidéo par un facteur 2.56 (cas des enregistrements à 64 images/seconde) ou 2 (cas des enregistrements à 50 images/seconde). Ce ralentissement n'a évidemment aucune incidence sur les données, mais devra être pris en compte dans toute analyse temporelle ou fréquentielle ultérieure, qui reposerait sur le fichier vidéo.

/6/ Synchronisation des images et du son

À l'origine, les images et le son ont été enregistrés sur deux supports différents. L'enregistrement de tops de "synchro image" sur une des pistes de l'enregistrement audio a permis de conserver des traces de leur synchronisation originale. Cela, associé à l'expertise des phonéticiens, permettra une post-synchronisation de qualité. Le son est évidemment ralenti par le même facteur que le signal vidéo.

/7/ Numérisation des films vidéo

Pour cela, nous avons acquis une carte de numérisation vidéo de très haute qualité qui traite à la volée les signaux en provenance du magnétoscope, et génère des images numériques au standard MJPEG. Ensuite, nous transformerons ces données, soit en fichiers vidéo au standard QuickTime, soit en fichiers au standard JPEG (images statiques). Le standard QuickTime à 25 images/sec a été choisi, car il existe aussi bien sur PC que sur Macintosh et peut être obtenu gratuitement sur le Web. La relative imprécision de ce standard n'est pas un handicap, car les séquences vidéo n'ont pas vocation à être

l'objet central de l'analyse articulatoire. Nous les concevons plutôt comme une aide à l'analyse des images statiques au standard JPEG, beaucoup plus précises sur le plan spatial, et échantillonnées à 50 images/secondes.

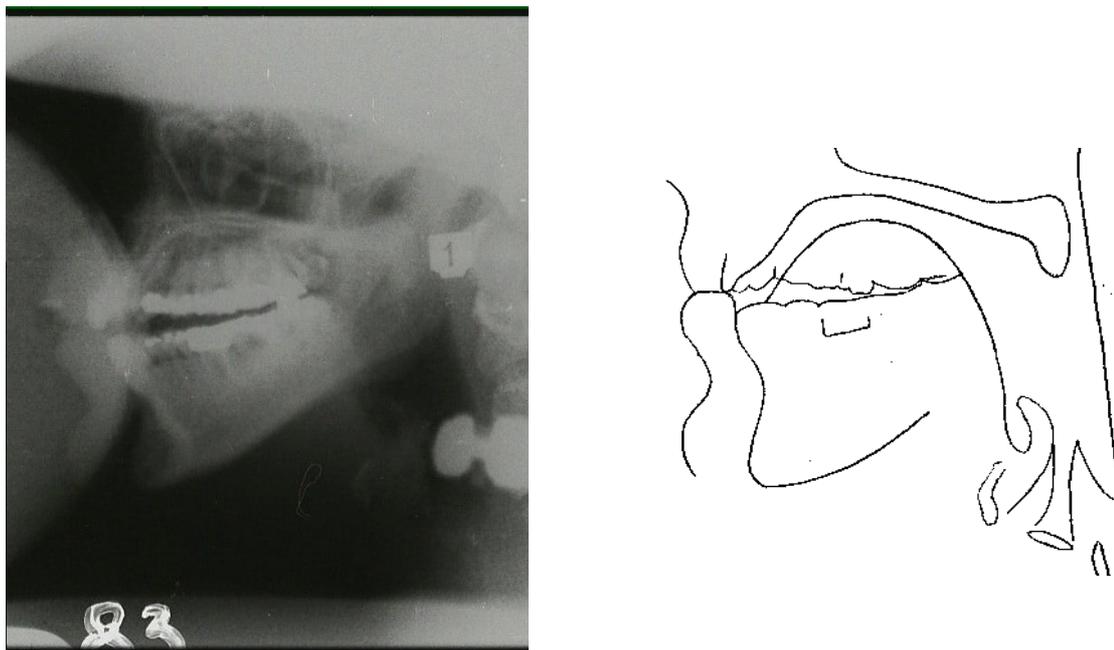


Figure 1: Radiographie de [m] dans [mi].(Figure de gauche).  
Tracé des sagittaux pour la même image (Figure de droite)

### Développement d'un outil d'extraction des mouvements du conduit vocal à partir de données cinéradiographiques

Comme nous l'avons signalé plus haut, l'extraction de données géométriques à partir de films radiologiques est généralement réalisée manuellement.

Cependant, à raison de 50 images par seconde, la quantité de tracés à effectuer est trop importante pour envisager de la traiter complètement de cette manière. C'est pourquoi, chaque film n'est associé aujourd'hui qu'à un nombre limité de tracés sagittaux, ceux qui étaient au centre des objectifs de recherche de son auteur à l'époque de la prise d'images. Dans le cadre du présent projet, nous devons donc envisager le développement d'un outil d'extraction automatique des tracés sagittaux à partir du film radiographique, qui pourrait s'appuyer sur les tracés déjà existants.

L'extraction automatique des contours de la langue a déjà été envisagée par Laprie et Berger (1996) pour exploiter au mieux ces grandes bases. Mais jusqu'à présent, seuls les travaux de Thimm et Luettin (1999) ont abouti au traitement complet d'un film (base ATR, Laval43).

En vue d'améliorer cette situation, nous mettons en place, dans le cadre de ce projet, une méthode semi-automatique qui sera applicable *film par film*, et qui combinera le marquage manuel et la reconstruction automatique du mouvement. Cette technique repose sur une adaptation de l'algorithme de rétro-marquage, dont le principe est d'associer des paramètres implicites et extraits du signal vidéo à des paramètres géométriques contrôlés et définis a posteriori, plutôt que d'extraire directement des données géométriques. Pour estimer les mouvements de langue, la méthode se décompose en 3 étapes :

/1/ le traitement manuel d'un nombre restreint d'images clefs qui permet de définir des paramètres géométriques (ici le contour de la langue) ;

/2/ une étape automatique d'indexation de la base à partir de ces mêmes images clefs réduites et cadrées, qui a pour but d'associer à chacune des images de la base le marquage géométrique et ;

/3/ des traitements postérieurs de régularisation. A noter que le rétro-marquage peut être rendu entièrement automatique, lorsque les informations géométriques sont extractibles dans les images clefs. Mais dans le cas de la langue, cette tâche très difficile, même pour l'expert humain, est dévolue au marquage manuel dans des conditions de facilitation que nous décrivons par la suite.

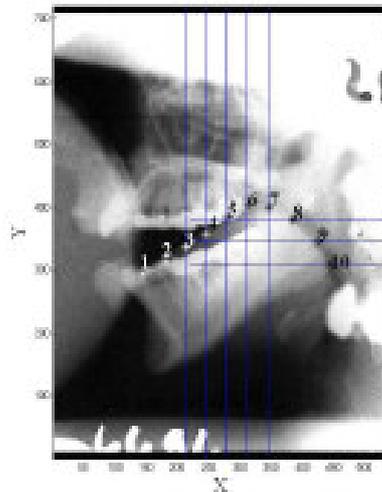
## Résultats obtenus

A l'heure actuelle, cette méthode a aisément été évaluée sur quelques films radiographiques et adaptée pour tirer profit des particularités de ces bases.

### Illustration de la méthode : Extraction des mouvements de la langue sur le film Wioland77 de notre base (Fontecave et Berthommier, 2006)

Le film Wioland77, enregistré en 1977, comprend 5673 images du conduit vocal provenant de 65 séquences vidéos (65 phrases prononcées par une locutrice française), enregistrées à 64 images par seconde.

L'étape manuelle consiste à décrire, pour 100 images clefs choisies aléatoirement, la position du contour de chaque articulateur cible (les lèvres, le corps de langue, l'apex de la langue, la mandibule, le velum, le palais dur, ou la paroi postérieure du pharynx) à l'aide d'un nombre limité de points choisis de telle sorte qu'il n'y ait pas de données manquantes. Par exemple, pour le contour de la langue, 10 points ont été définis, parmi lesquels 8 déterminent le contour du corps de la langue par intersection avec des lignes verticales et horizontales (base et dos), et 2 l'apex (Fig. 2).



**Figure 2 : Excepté pour la pointe, les points sont marqués à l'intersection entre le contour de la langue et les lignes verticales ou horizontales.**

Ensuite, pour chaque image de la séquence, l'index de l'image clef la plus proche est assigné par le biais d'une mesure de similarité fondée sur la distance Euclidienne entre les coefficients DCT (*Discrete Cosinus Transform*) basses fréquences calculés pour les deux images sur la partie où se situe l'articulateur cible. Après indexation, on aboutit ainsi à un premier marquage de l'ensemble des images du film, qui attribue à chacune d'entre elles la description géométrique de l'image clef associée. Cette description géométrique est ensuite affinée par des traitements postérieurs fondés sur du filtrage temporel et du moyennage des configurations voisines obtenues par multi-indexation, ainsi que sur un lissage par fonctions splines appliqué sur les points estimés du contour de l'articulateur. La figure 3 montre un exemple du résultat obtenu pour le contour de la langue.

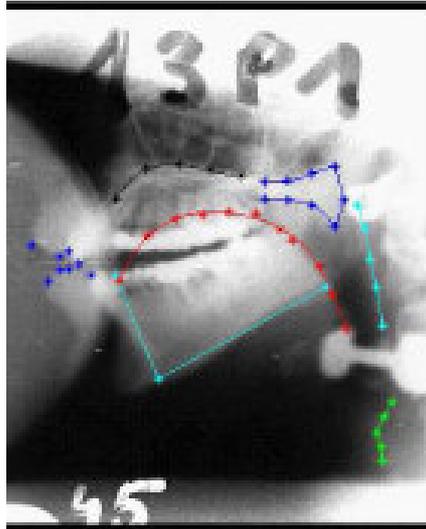


Figure 3 : Exemple des résultats obtenus pour le contour des lèvres (à gauche en bleu), du palais et du velum (en haut en bleu), du corps et de l'apex de la langue (en rouge) et de la paroi postérieure du pharynx (en bleu cyanure)

Une première évaluation objective du traitement proposé a été effectuée pour le contour du corps de la langue à l'aide d'un deuxième jeu de 100 images tests marquées manuellement. L'erreur quadratique moyenne a été d'abord calculée pour chacun des points déterminant le contour, puis nous en avons calculé la moyenne sur l'ensemble de ces points. L'erreur résultante est égale à 11 pixels (à comparer avec 350 pixels de longueur totale), ce qui équivaut à environ 3 mm selon une calibration approximative, l'information d'échelle n'étant pas directement disponible sur le film. Cette erreur reste encore très importante et elle n'est pas acceptable dans une perspective d'exploitation de nos données pour l'étude de la production de la parole. Nos travaux sur ce point au cours du projet viseront donc à atteindre une erreur moyenne de l'ordre du millimètre (à titre de comparaison, l'EMMA, articulographe électromagnétique permettant la mesure des déplacements de points situés sur les lèvres, la mandibule et sur la langue dans les parties palatale et alvéopalatale, permet une précision de l'ordre du demi millimètre).

## VALORISATION

### /1/ Intégration dans une Base de Données

Pour permettre une *large distribution de ces données*, et pour faciliter leur exploitation, nous avons posé, *a priori*, les contraintes suivantes pour la conception du système de gestion de la base de données :

/a/ travailler avec un SGBD relationnel tournant sous des machines standard, accessibles dans tout laboratoire de parole, de type PC ou Macintosh

/b/ permettre la recherche dans la base de données à l'aide de requêtes portant sur les caractéristiques phonétiques des sons (voyelle/consonne, voisé/non voisé, ouvert/fermé...), sur l'écriture phonétique des sons, de manière isolée ou en contexte (n'excédant pas 5 phonèmes), et cela de manière totalement transparente pour l'utilisateur.

La recherche dans la base de données pourra se faire par phonèmes, par phrase entière, par vidéo, en entrant le nom de la vidéo qui nous intéresse, ou par locuteur, en ne demandant à voir que les films d'un seul locuteur, par exemple. Les informations ainsi obtenues seront affichées de manière succincte et signalétique. Pour l'obtention d'une information complète, un système par accès limité et abonnement avec login et mot de passe sera mis en place.

**Nous utiliserons le SGBD relationnel Flora (société Ever) géré par la MISHA qui mettra en place et assurera le suivi de la base de données. La consultation se fera via un navigateur web, de type Firefox.**

**Nous créerons un manuel d'utilisateur en langage HTML qui expliquera la procédure de consultation de la base. Ce manuel comportera des renseignements sur le codage des phonèmes, des instructions sur la recherche dans la base de données, la visualisation des résultats, ainsi que des remarques à propos de l'étiquetage des films et des cas particuliers (traitement de la coarticulation, étiquetage des radiofilms, recherches sur plusieurs phonèmes...).**

/2/ Information et culture scientifique et technique

/a/ Ateliers, colloques, etc. Les résultats de ces rencontres (comptes rendus, actes...) seront publiés sur le serveur de publications électroniques de la MISHA.

/b/ Mise au point d'un outil multimédia pour l'information et la vulgarisation scientifique, auprès du grand public lors de manifestations scientifiques : le Jardin des Sciences, La Science en Fête, Les Journées des Universités, etc.

Trois axes de recherche : coarticulation, inversion et modélisation

Pour ce qui concerne l'exploitation pour l'étude de la production de la parole, l'analyse des données cinéradiographiques offre essentiellement 3 axes de recherche :

### **Axe 1 Caractérisation des phénomènes de coarticulation**

En ce qui concerne la caractérisation des phénomènes de coarticulation, un des intérêts potentiels de la base de données de l'Institut de Phonétique de Strasbourg est son aspect multilingue. Nous envisageons, en conséquence, un travail qui étudierait la coarticulation (et précisément la coarticulation anticipatoire) et sa variabilité, en fonction de la densité des systèmes phonologiques. Cela peut permettre d'intégrer relativement facilement une composante interaction production-perception, puisqu'il serait possible de relier la variabilité des phénomènes de coarticulation, selon les langues, à des contraintes perceptives.

### **Axe 2 : Inversion des données acoustiques**

L'élaboration de modèles de contrôle pourra se faire par inversion des données acoustiques. Pour y parvenir, il s'agira de procéder à la définition de contraintes coarticulatoires, à partir d'images du conduit vocal, extraites de la base. La dimension multilingue du corpus permettra de confronter diverses stratégies coarticulatoires, imposées largement par des contraintes phonologiques spécifiques aux langues. Le développement et l'évaluation des méthodes d'inversion pourront ainsi être enrichis par la nature diversifiée du corpus.

### **Axe 3 : Evaluation de modèles physiques**

La compréhension des mécanismes de contrôle de la production de la parole, c'est-à-dire du passage de la chaîne phonémique au déplacement des articulateurs du conduit vocal et à la production du signal acoustique, passe selon nous par le développement et l'évaluation de modèles du contrôle moteur (modèle de la génération des commandes musculaires, à partir de la chaîne de phonèmes) et de modèles physiques de l'appareil de production de la parole (génération du mouvement des articulateurs, puis du son à partir des commandes musculaires). C'est la comparaison de données mesurées sur des locuteurs, dans des conditions bien contrôlées avec les prédictions proposées par ces modèles, qui permet d'évaluer, de manière quantitative et objective, les différentes hypothèses proposées dans la littérature sur le contrôle de la production de la parole (existence de cibles en parole, nature de ces cibles, modèle de coarticulation, d'anticipation.....). De ce point de vue, les données cinéradiographiques présentent un intérêt tout particulier, puisqu'elles offrent une vue complète du conduit vocal, dans le plan sagittal, et cela à une cadence relativement importante (50 Hz au minimum). Cet objectif d'évaluation des modèles physiques est donc une des applications les plus naturelles et les plus attendues de notre travail de formatage et de mise à disposition de nos données.

## Axe 1 : Caractérisation des phénomènes de coarticulation

Nous savons que lors de toute coordination précise des articulateurs dans un espace donné, le chevauchement gestuel (*gestural overlap*) ou la coarticulation est une « règle » générale. Cependant, la production de certains sons requiert un taux de chevauchement plus prononcé que pour d'autres, un comportement coarticulatoire pouvant être particulièrement étendu dans des contextes spécifiques et suivant les locuteurs et les langues. Les recherches que nous menons dans le domaine de la coarticulation, et plus précisément dans celui de l'anticipation (Sock & Vaxelaire, 2004), focalisent sur une catégorie de ces gestes coarticulés : celle des gestes anticipatoires. Elles examinent divers aspects du contrôle et de la coordination de ces gestes pendant la production de la parole avec, en toile de fond, le *couplage structurel* constant entre efficacités sensorielles et gestes articulatoires.

Les gestes anticipatoires – compris ici simplement comme l'expansion ou l'extension de certains attributs des mouvements articulatoires à des segments adjacents ou avoisinants – sont souvent considérés comme apportant une contribution essentielle à la production de la parole. En effet, des données récentes indiquent que l'initiation précoce de certains gestes est nécessaire pour une perception adéquate du son, et c'est ainsi que les auditeurs exploiteraient des indices précoces liés à ces éléments anticipatoires dans la chaîne parlée (Abry & Lallouache, 1995 ; Sock *et al.*, 1999). Dans cette perspective, où le comportement anticipatoire n'est conçu, en réalité, que comme l'extension de certaines caractéristiques d'un son ou d'une configuration visuelle, à son entourage immédiat ou moins immédiat, la catégorisation de la notion du temps en *présent*, *passé* et *futur* ne peut plus être pertinente, puisque toute dimension dite anticipatoire ou rétentrice du son fait, au bout du compte, partie intégrante de l'identité articulatoire-acoustique du phénomène émergent lui-même (Hecker *et al.*, 2001 ; 2002 ; Hirsch *et al.*, 2003 ; Sock, 2003 ; Asci *et al.*, 2003).

La partie anticipatoire des gestes comporte ses portions audibles et inaudibles (Löfqvist, 1990). Signalons, toutefois, que les gestes inaudibles peuvent servir d'indices visuels de la parole, ou peuvent simplement servir, en initiant les mouvements aussi tôt que possible, à améliorer la précision des effecteurs finaux, et en conséquence leur efficacité perceptives (Vaxelaire *et al.*, 1999).

L'utilisation de tels indices, ou du décalage naturel des gestes articulatoires en avance sur le signal acoustique, a été démontrée pour le français, dans le domaine de la perception visuelle, par Cathiard (1994), et par Roy (2004) pour des sujets malentendants et sourds profonds, avec des résultats parfois divergents. Dans le domaine de la perception de la parole, et sur le plan acoustique-auditif, les résultats dont on dispose (Benguérel & Adelman, 1976) ne prennent, malheureusement, pas en compte la relation entre le niveau articulatoire et ses efficacités acoustiques.

Les recherches que nous proposons de mener, dans le cadre de ce projet, en production et en perception de la parole, suivent de très près celle conduite par Lubker et Lindgren (1982) sur le suédois.

Elles reposeront sur les données cinéradiographiques et acoustiques de la base en soulevant des questions précises :

/1/ Quelle est la particularité et l'extension temporelle de l'anticipation de certains gestes vocaliques et consonantiques pour une langue spécifique ?

/2/ Quelles sont les stratégies anticipatoires généralisables à travers des langues non-apparentées de la base ?

/3/ Est-ce que ces gestes anticipatoires contribuent à la perception auditive précoce d'une voyelle ou d'une consonne ?

/4/ Quel est le domaine de l'effet perceptif de ces gestes ?

/5/ De quelle manière la variation de la vitesse d'élocution et la stratégie individuelle du locuteur pourraient-elles influencer sur l'extension perceptive de ces gestes anticipatoires ?

Nous avons vu plus haut que l'augmentation de la vitesse d'élocution était en effet un moyen naturel de perturber le timing des gestes articulatoires. Il serait donc judicieux de savoir dans quelle mesure une perturbation motrice pourrait affecter l'organisation spatio-temporelle des gestes et leur efficacité perceptive.

### **Méthodologie : le paradigme du *gating***

Afin de vérifier les effets perceptifs des gestes anticipatoires, des tests de perception, reposant sur le paradigme du *gating* (dévoilement progressif d'un signal ou d'une séquence d'images cinéradiographiques) seront conduits pour des séquences choisies. Dans cette méthode, l'on sélectionne une phrase ou une séquence cible, puis des segments sont tronqués de ce signal acoustique ou des images correspondant à la séquence, en partant, soit du début acoustique, soit de la première image de l'élément cible, en arrière dans le signal ou dans les images de la phrase porteuse. Ensuite, des sujets doivent identifier la voyelle ou la consonne manquante. Notons que les sujets se trouvant dans de telles conditions sont en situation de tension ou de visée ouverte, étant donné que l'événement acoustique ou visuel présumé n'est pas encore livré, dans sa réalité effective, à leur conscience.

Nous avons pu apporter, grâce à une série d'investigations, quelques réponses aux questions (citées *supra*) que nous nous posons dans le domaine de la coarticulation, mais cela uniquement pour le français (Sock & Vaxelaire, 2004).

#### **Nous en présentons quelques unes ici :**

/1/ Quelle est l'extension temporelle de l'anticipation des gestes vocaliques et consonantiques ?

Réponse : Cette extension est variable qu'il s'agisse d'un geste vocalique ou d'un geste consonantique. Elle ne devrait pas être exprimée en seules durées absolues, étant donné la grande élasticité du signal de parole, mais plutôt en termes articulatoire-acoustiques événementiels.

/2/ Est-ce que ces gestes anticipatoires contribuent à la perception auditive précoce d'une voyelle ou d'une consonne arrondie en français ?

Réponse : Si les gestes vocaliques anticipatoires contribuent à la reconnaissance précoce d'une voyelle protruse, les gestes consonantiques anticipatoires, eux, sont inaudibles. En revanche, ils servent à améliorer la perception auditive de la consonne labialisée au moment de son émergence acoustique.

/3/ Quel est le domaine de l'effet perceptif de ces gestes ?

Réponse : Ce domaine d'efficacité sensorielle est fonction des événements moteurs localisés dans la façade anticipatoire des gestes vocaliques. En général, nos données montrent que son étendue dépend du moment de l'arrivée de l'événement articulatoire du « relâchement » et de l'émergence de l'événement cinématique du « pic de vitesse ». A partir de ces dates, la perception précoce de la voyelle va en s'améliorant, à mesure que d'autres événements articulatoire-acoustiques, tels que l'ouverture minimale des lèvres et le pic de protrusion apparaissent.

## Axe 2 : Inversion des données acoustiques

L'inversion acoustique articulatoire consiste à retrouver l'évolution de la forme du conduit vocal à partir du signal de parole. L'inversion repose sur le paradigme d'analyse par synthèse. Cela signifie que l'on utilise un modèle articulatoire pour synthétiser la forme géométrique du conduit vocal et le spectre de la parole produite. Les spectres générés sont comparés aux spectres de la parole à inverser.

Un élément central consiste à définir des contraintes qui soient autant restrictives que réalistes d'un point de vue phonétique, afin d'éliminer des solutions fausses. De telles contraintes sont habituellement obtenues à partir d'images du conduit vocal, permettant d'élaborer un modèle approximatif de la production de la parole et/ou à partir d'images du visage du locuteur, cela pour pouvoir extraire de l'information des articulateurs visibles, tel que les auditeurs le font.

Mise à part des problèmes théoriques, un des principaux défis dans ce domaine est le manque cruel de données articulatoires montrant et le conduit vocal du locuteur, et son visage. La sauvegarde et le traitement des données cinéradiographiques est donc une opportunité scientifique et technologique réelle pour aider au développement et à l'évaluation des méthodes d'inversion.

Outre les nombreuses applications potentielles dans le domaine de la reconnaissance automatique de la parole, du pilotage de têtes parlantes ou encore d'aides à l'apprentissage de langues étrangères, l'inversion permet aussi d'explorer le comportement acoustique d'un modèle articulatoire et les mécanismes de production de la parole vis-à-vis de l'interdépendance entre articulateurs.

L'exploration d'un modèle articulatoire permet en particulier de retrouver les points d'articulation des voyelles et donc découvrir les caractéristiques articulatoires propres à chaque locuteur.

Notre méthode d'inversion comporte trois étapes :

/1/ La première étape consiste à générer un grand nombre de solutions potentielles : pour cela, nous utilisons une table articulatoire (ou *codebook*), qui associe des vecteurs articulatoires (à 7 dimensions, correspondant aux 7 paramètres du modèle de Maeda) à leurs correspondants acoustiques (dans notre cas, le triplet des fréquences des 3 premiers formants). Un vecteur acoustique étant donné, il existe a priori une infinité de vecteurs articulatoires permettant de l'obtenir, nous n'avons donc pas l'ambition de générer toutes les solutions inverses possibles. Cependant, il est nécessaire, pour avoir une inversion de qualité, que les échantillons retenus soient suffisamment représentatifs pour contenir des solutions proches de la solution réelle.

/2/ La deuxième étape de notre méthode consiste en la reconstruction d'une trajectoire articulatoire qui soit suffisamment régulière au cours du temps. Nous utilisons pour cela un algorithme de programmation dynamique qui minimise une fonction de coût représentant la "distance" couverte par les articulateurs.

/3/ La dernière étape consiste en l'amélioration de la fidélité acoustique et de la régularité articulatoire de la solution obtenue à l'étape précédente en utilisant un algorithme de régularisation variationnelle.

### Construction du codebook articulatoire

La force de notre méthode d'inversion réside dans la résolution acoustique quasi uniforme du codebook. Cette propriété est garantie par la façon dont est construite la table : on explore l'espace récursivement en évaluant à chaque étape la linéarité locale de la relation articulatoire acoustique. Si la relation n'est pas suffisamment linéaire, on subdivise l'espace.

Plus précisément, les paramètres articulatoires du modèle de Maeda varient entre  $-3\sigma$  et  $+3\sigma$ , où  $\sigma$  est l'écart type, l'espace articulatoire peut-être vu comme un hypercube à 7 dimensions (de rayon  $6\sigma$ ). L'échantillonnage de l'espace articulatoire se fait en cherchant des points qui délimitent les zones linéaires ; un hypercube étant donné, on évalue sa linéarité en considérant tous les segments reliant deux sommets de l'hypercube : les valeurs des vecteurs acoustiques obtenues par synthèse au niveau des sommets sont linéairement interpolées au milieu du segment, et le vecteur correspondant est comparé à la valeur du vecteur obtenu par synthèse au milieu du segment. Si la différence entre les deux est inférieure à un certain seuil prédéfini, alors la relation est considérée comme linéaire pour le segment.

Si la relation est linéaire pour tous les segments, alors l'hypercube est considéré linéaire, et on sauvegarde l'hypercube dans le codebook. Sinon, on subdivise l'hypercube, et on applique récursivement les tests de linéarité dans tous les sous-hypercubes. Comme nous ne testons que la linéarité des segments reliant deux sommets, nous n'avons aucune garantie sur le comportement de la relation à l'intérieur de l'hypercube ; mais la simplicité du test est dictée par le temps de calcul important que prend la synthèse de l'image d'un vecteur articulatoire. Expérimentalement, nous avons montré (OUNI & LAPRIE 2001) qu'en utilisant un seuil de 0.3 bark par formant lors du test, l'erreur moyenne sur les formants, pour les points générés par la méthode d'inversion, était inférieure à 10Hz.

En pratique, pour des raisons d'explosion combinatoire, on est obligé de limiter le niveau de subdivisions (en dimension 7, faire une subdivision oblige à explorer  $2^7$  nouveaux hypercubes, donc augmenter d'un niveau la subdivision multiplie a priori le temps de calcul par 128).

On arrête de subdiviser quand le test de linéarité est satisfait, ou quand la taille d'un côté de l'hypercube passe au-dessous d'un certain seuil. Dans les deux cas, on sauve les cubes dans le codebook. Par ailleurs, on subdivise également si un ou plusieurs points d'un cube ne donnent pas une fonction d'aire de voyelle réaliste, c'est-à-dire si l'aire à la constriction est nulle ou trop faible. Ces cubes sont situés à la frontière de l'espace articulatoire ; comme pour le test de linéarité, on arrête la subdivision quand la taille du cube est inférieure à un certain seuil (différent du précédent), et dans ce cas on rejette le cube.

### **Exploration de l'espace nul de la relation articulatoire acoustique**

Pour chaque vecteur acoustique représenté par les trois premières fréquences formantiques, le processus d'inversion consiste en la recherche de tous les hypercubes qui peuvent générer le triplet de formants observé. Il faut ensuite trouver un ensemble de solutions dans chacun de ces cubes. Comme l'inversion consiste à trouver 7 paramètres à partir de 3, l'espace des solutions a *a priori* 4 degrés de liberté. La relation articulatoire acoustique (notée R) est supposée être localement linéaire au niveau du centre  $P_0$  de l'hypercube (c'est-à-dire que l'application  $P - P_0 \rightarrow R(P) - R(P_0)$  est supposée être une application linéaire). Trouver l'ensemble des solutions n'est pas un problème trivial, car il s'agit de trouver l'intersection d'un espace à 4 dimensions (l'espace nul de la relation précédente, c'est-à-dire l'ensemble des antécédents de 0 pour l'application linéaire) et d'un hypercube à 7 dimensions, ce que l'on ne sait pas faire de manière formelle. Une première approximation de l'intersection est obtenue par programmation linéaire. Puis l'espace nul est échantillonné, et l'appartenance à l'intersection de chacun des points est testée (Ouni & Laprie, 2001).

### **Adjonction de contraintes pour l'inversion**

La principale difficulté de l'inversion est qu'une infinité de formes du conduit vocal donnent le même spectre de parole. L'utilisation d'un modèle articulatoire régissant les modes de déformation du conduit vocal est une première réponse. Mais cela ne suffit pas à réduire toute l'indétermination. Il faut donc ajouter des contraintes qui réduisent encore le nombre de formes possibles. Nous avons étudié deux types de contraintes. Les premières sont les contraintes qui expriment des connaissances classiques de phonétique concernant l'articulation des sons. Ces contraintes donnent pour chaque voyelle des domaines de valeurs acceptables pour chacun des paramètres. Elles capturent des dépendances entre paramètres qui ne sont pas faciles à mettre en évidence autrement et permettent d'éliminer des gestes très compliqués à réaliser pour un locuteur humain (Potard & Laprie, 2005). Par ailleurs, nous commençons à introduire des contraintes qui portent sur les articulateurs visibles (Wrobel-Dautcourt *et al.*, 2005).

Axe 3 : Evaluation de modèles physiques

### ***De l'intérêt de modéliser les processus physiques de la production de la parole pour comprendre le contrôle de la production de séquence de parole, et le lien entre phonétique et phonologie***

L'étude de l'interface entre phonétique et phonologie repose sur l'analyse de données articulatoires et/ou acoustique qui est interprétée à la lumière de théories phonologiques. De la théorie

des traits (Chomsky & Halle, 1968) jusqu'à la phonologie non-linéaire (Clements, 1992), en passant par la phonologie articulatoire (Browman & Goldstein, 1986), les théories phonologiques sont nombreuses, et toutes trouvent un certain nombre d'arguments en leur faveur, dans les corpus de données que leurs auteurs ou leurs partisans analysent. Ces démarches expérimentales importantes et nécessaires souffrent cependant d'une limitation méthodologique, dont il ne faut pas ignorer l'existence : les caractéristiques articulatoires et acoustiques de la parole ne sont pas le résultat de l'influence directe et unique de la spécification phonologique. Le système périphérique de production de la parole a, en effet, des caractéristiques physiques qui lui sont propres et qui influent, parfois de manière déterminante, sur la façon dont la mandibule, la langue, les lèvres, le velum, ou les cordes vocales se déplacent au cours du temps, ainsi que sur la façon dont ces articulateurs interagissent les uns avec les autres pour produire les sons de la parole. Nos systèmes de mesure physiologique, cinématique, acoustique ou dynamique, aussi sophistiqués soient-ils, ne permettent donc pas d'observer les conséquences directes des commandes envoyées par le système nerveux central, mais seulement celles de l'interaction entre ces commandes et le système physique de production de la parole. Les incidences potentielles de ces caractéristiques physiques sur les données mesurées et sur l'interprétation des données qui en découle sont bien illustrées par les 2 exemples suivants, largement débattus dans la littérature.

En premier lieu, nous évoquerons le phénomène de « réduction vocalique » mis en évidence la première fois par Lindblom (1963). En mesurant les fréquences F1 et F2 des deux maxima du spectre (les *formants*) de la voyelle [u], prononcée par un locuteur de langue anglaise à l'intérieur de la séquence [dud] et à des débits d'élocution différents, Lindblom (1963) a observé une très large distribution des valeurs F2, allant d'une fréquence très basse, prototypique pour la voyelle /u/, jusqu'à des valeurs sensiblement plus élevées, classiquement associées aux voyelles centrales. Cette observation posait la question de savoir si une telle variabilité acoustique était compatible avec une représentation phonologique unique. Pour répondre, Lindblom (1963) a émis l'hypothèse que les caractéristiques dynamiques des transitions entre consonne et voyelle étaient similaires à celles d'un système masse-ressort. Ainsi, dans un cadre relativement simple de modélisation des caractéristiques physiques de l'appareil de production de la parole, et en jouant simplement sur la durée de la séquence [dud], Lindblom (1963) a pu montrer que le domaine de variation de F2 mesuré était tout à fait compatible avec l'hypothèse d'une cible acoustique invariante pour la voyelle : lorsque la vitesse d'élocution augmente, la durée de la séquence [dud] diminue de telle sorte que la cible visée n'a pas le temps d'être atteinte ; on parle alors de *ratage de cible* (*target undershoot*). On voit donc qu'une prise en compte très simple (et probablement même simpliste) des propriétés physiques du conduit vocal a permis d'aller au-delà de la simple observation expérimentale et d'interpréter les données en un sens qui permet de soutenir l'hypothèse que la vitesse d'élocution n'influerait pas sur la représentation phonologique d'un même son élémentaire, situé dans un environnement phonétique constant (pour une étude plus récente de ce phénomène, voir Loevenbruck & Perrier, 1993 ; Loevenbruck *et al.*, 1995 ; Perrier *et al.*, 1996).

La prise en compte des propriétés dynamiques des gestes articulatoires de la parole constitue un des éléments de base de la *Phonologie Articulatoire*, élaborée par Browman & Goldstein (1986) : "*In our view, then, a phonetic representation is a characterization of how a physical system (e.g. a vocal tract) changes over time.*" Les propriétés dynamiques des articulateurs de la parole influent sur le déroulement temporel des gestes articulatoires, dont l'organisation temporelle n'est plus l'image directe de celle de la séquence phonologique. Ainsi, selon Browman et Goldstein (1990), on n'a plus un enchaînement strictement séquentiel des mouvements dans le temps ; à cause de leur dynamique intrinsèque, les gestes articulatoires, dirigés vers des objectifs successifs, caractérisés par une séquence de gestes articulatoires appelée "*constellation gestuelle*" (la spécification phonologique), pourraient ainsi se superposer dans le temps, pour former des "*agrégats gestuels*". En s'appuyant sur cette hypothèse, Browman et Goldstein (1990) ont pu expliquer des données articulatoires mettant en évidence la réduction, et même dans certains cas la disparition, d'une syllabe dans une séquence parlée. Cette démarche, visant à évaluer en quoi les caractéristiques dynamiques de la parole pourraient influencer la forme des gestes articulatoires, a été reprise de manière très approfondie par Saltzman et Munhall (1989). Ces auteurs ont ainsi pu montrer que, au cours de la séquence [kl] en anglais, la variabilité de l'évolution temporelle des contacts palataux, mesurée par Hardcastle (1985) lorsque la vitesse d'élocution varie, pourrait être expliquée en faisant l'hypothèse que chaque consonne serait associée à un geste articulatoire invariant (la spécification phonologique), et que la variation de la vitesse d'élocution induirait une superposition temporelle variable de ces deux gestes. Saltzman et Munhall (1989) ont aussi exploité ce concept de superposition gestuelle pour analyser les mouvements laryngés,

et la coordination entre les cordes vocales et les articulateurs du conduit vocal. Löfqvist et Yoshioka (1980) et Munhall et Löfqvist (1992) ont en effet observé par transillumination que, au cours de la production des séquences consonantiques /t#k/ et /s#t/, lorsque la vitesse d'élocution était lente, il existait deux gestes successifs d'ouverture des cordes vocales, un pour chaque consonne, alors que ces deux gestes successifs se réduisaient à un seul lorsque la vitesse devenait rapide. Saltzman et Munhall (1989) ont pu reproduire ce phénomène, à l'aide de simulations dans lesquelles chacun des gestes glottiques était représenté comme la sortie d'un système dynamique du second ordre, et où l'augmentation du débit glottique correspondait à une superposition temporelle, toujours plus grande pour les deux gestes.

L'intérêt que représente une démarche approfondie de modélisation physique des mécanismes sous-jacents à la production de la parole est donc patent, et c'est ce à quoi travaillent les membres de l'équipe "Acoustique, Aéroacoustique, Biomécanique et Contrôle" de l'ICP depuis de nombreuses années. Les phénomènes modélisés vont de la génération des commandes musculaires, à partir d'une séquence phonologique donnée, jusqu'à la génération du son, en passant par une modélisation biomécanique des articulateurs de la parole (Pelorson *et al.*, 1996 ; Payan & Perrier, 1997 ; Perrier *et al.*, 2003 ; Vilain *et al.*, 2003 ; Gérard *et al.*, 2006 ; Perrier, 2006).

### **Utiliser les données cinéroradiographiques pour valider les modèles**

Parmi les modèles évoqués ci-dessus, nous nous intéresserons plus spécifiquement à l'évaluation modèles biomécaniques et à leur contrôle, selon différentes théories de la coarticulation (cf. ci-dessus). Les modèles biomécaniques actuellement étudiés à l'ICP sont des modèles de la langue qui représentent cet articulateur comme une structure à éléments finis déformable. Ces modèles peuvent être bi-dimensionnels (Payan & Perrier, 1997 ; Perrier *et al.*, 2003), c'est-à-dire représentant la langue dans sa projection dans le plan médio-sagittal de la tête, ou bien tridimensionnels (Gérard *et al.*, 2006). Entre ces deux types de modèles, la complexité de la modélisation est différente, et le degré de réalisme de la description plus ou moins grand, mais les principes de base sont les mêmes : la langue est discrétisée en petits éléments déformables, dont les caractéristiques élastiques sont spécifiées, et les muscles agissant sur la forme et la position de cet articulateur sont représentés à l'intérieur de cette structure, par un certain nombre d'éléments spécifiques, en conformité avec des données anatomiques publiées dans la littérature. Ces muscles sont commandés par un modèle de contrôle classique (quoique controversé) dans le domaine de la recherche sur le contrôle moteur, la théorie du Point d'équilibre de Feldman (1986).

L'exploitation des données cinéroradiographiques pour valider ce modèle intervient à deux niveaux. Le premier de ces niveaux consiste à vérifier que (1) les déformations et (2) les vitesses de déformations que prédisent les modèles dans le plan medio-sagittal sont réalistes pour des niveaux de force (ou d'activation) musculaire réalistes pour la production de la parole (quelques newtons). Cette phase permet de valider le réalisme de la description physique que nous proposons, depuis les signaux d'activation musculaire, jusqu'à la forme et la position de la langue. Elle nécessite d'avoir une description complète de la forme de la langue et non une description discrète limitée à quelques points sur le contour lingual supérieur. Seules la cinéroradiographie et l'Imagerie par Résonance Magnétique offrent cette description complète. On comprend donc tout l'intérêt de notre base de données de ce point de vue.

Le deuxième niveau, une fois validée la description physique, consistera à tester différents modèles de contrôle de la production de la parole à partir de la chaîne phonémique pour générer des séquences de parole. Cela renvoie à l'ensemble des modèles de coarticulation qui ont été évoqués plus haut dans la partie consacrée à l'axe 1. Le principe de l'évaluation ici sera de générer avec nos modèles physiques, pour une séquence de phonèmes donnée, différentes séquences de mouvements selon le modèle de coarticulation choisi. La comparaison systématique des déformations linguales, ainsi prédites par le modèle avec les déformations mesurées sur les cinéroradiographies, permettra d'avancer de manière quantitative sur l'évaluation de ces différents modèles. Là encore, la richesse de la description linguale proposée par la cinéroradiographie, et sa bonne précision temporelle sont des atouts importants par rapport aux autres techniques actuellement disponibles. La richesse linguistique des différents corpus disponibles dans notre base de données sera aussi un élément très important de cette étape d'évaluation, puisqu'elle permettra de mieux distinguer ce qui pourrait être commun à toutes les langues (et que l'on pourrait donc attribuer aux caractéristiques intrinsèques du système physique de production de la parole) de ce qui est plus spécifique à chaque langage (que l'on pourrait donc relier à ses caractéristiques phonologiques intrinsèques).

## **Calendrier et modalités du travail :**

### **Année 1 – 2006/2007 :**

#### **[Tâches assurées à l'IPS]**

- /1/ Sélection des films : 10 films
- /2/ Repérage du corpus pertinent
- /3/ Première phase de la sauvegarde des radiofilms (actuellement sur films 35mm et bandes audio) par stockage sur un support vidéo de haute qualité, et par numérisation et stockage sur DVDROM ou disque dur externe
- /4/ Synchronisation des images et du son
- /5/ Inventaire des tracés manuels existants
- /6/ Réalisation manuelle de nouveaux tracés
- /7/ Numérisation des tracés manuels

Livraison : films synchronisés et tracés manuels numérisés à l'ICP

#### **[Tâches assurées à l'ICP]**

- /1/ Numérisation des films
- /2/ Etiquetage des films
- /3/ Développement d'un outil d'extraction automatique des contours
- /4/ Extraction automatique des contours
- /5/ Développement de la plateforme logicielle
  
- /6/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'élaboration et l'évaluation de modèles physiques

Livraison : films numérisés, tracés manuels numérisés et contours extraits automatiquement à l'IPS – la MISHA et au LORIA

#### **[Tâches assurées au LORIA]**

- /1/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'inversion acoustique articulatoire
- /2/ Elaboration d'un modèle de contrôle

#### **[Tâches assurées à l'IPS et à la MISHA]**

- /1/ Développement du système de gestion de la base de données
- /2/ Intégration des données dans la base
- /3/ Evaluation du système de gestion de la base

#### **Réunions de travail**

- /1/ Deux réunions, dont une à l'IPS et l'autre à l'ICP : courant 2007
- /2/ Réunion de tous les partenaires du projet à Strasbourg : fin 2007
- /3/ Rédaction d'un rapport partiel sur l'avancement du projet

### **Année 2 - 2007/2008 :**

#### **[Tâches assurées à l'IPS]**

- /1/ Deuxième phase de la sauvegarde des radiofilms
- /2/ Synchronisation des images et du son

- /3/ Réalisation manuelle de nouveaux tracés
- /4/ Numérisation des tracés manuels
- /5/ Mesures et élaboration de contraintes articulatoires et acoustiques de la coarticulation

Livraison : films synchronisés et tracés manuels numérisés à l'ICP

#### **[Tâches assurées à l'ICP]**

- /1/ Numérisation des films
- /2/ Etiquetage des films
- /3/ Affinement de l'outil d'extraction automatique des contours
- /4/ Extraction automatique des contours
- /5/ Développement de la plateforme logicielle
- /6/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'élaboration et l'évaluation de modèles physiques

Livraison : films numérisés, tracés manuels numérisés et contours extraits automatiquement à l'IPS – la MISHA et au LORIA

#### **[Tâches assurées au LORIA]**

- /1/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'inversion acoustique articulatoire
- /2/ Affinement du modèle de contrôle

#### **[Tâches assurées à l'IPS et à la MISHA]**

- /1/ Développement du système de gestion de la base de données
- /2/ Intégration des données et des outils dans la base
- /3/ Evaluation du système de gestion de la base

#### **Réunions de travail, atelier et publications**

- /1/ Deux réunions, dont une à l'IPS et l'autre au LORIA : courant 2008
- /2/ Organisation d'un colloque-atelier international sur les données cinéradiographiques et les bases de données articulatoires, à Strasbourg : 2008
- /3/ Accès de la communauté scientifique à la base
- /4/ Publications communes sur : la base, la coarticulation, l'inversion et l'évaluation de modèles physiques
- /5/ Rédaction d'un rapport partiel sur l'avancement du projet

### **Année 3 - 2008/2009 :**

#### **[Tâches assurées à l'IPS]**

- /1/ Troisième et dernière phase de la sauvegarde des radiofilms
- /2/ Synchronisation des images et du son
- /3/ Réalisation manuelle de nouveaux tracés
- /4/ Numérisation des tracés manuels
- /5/ Mesures et élaboration de contraintes articulatoires et acoustiques de la coarticulation

Livraison : films synchronisés et tracés manuels numérisés à l'ICP

#### **[Tâches assurées à l'ICP]**

- /1/ Numérisation des films
- /2/ Etiquetage des films
- /3/ Affinement de l'outil d'extraction automatique des contours
- /4/ Extraction automatique des contours

- /5/ Développement de la plateforme logicielle
- /6/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'élaboration et l'évaluation de modèles physiques

Livraison : films numérisés, tracés manuels numérisés et contours extraits automatiquement à l'IPS – la MISHA et au LORIA

#### **[Tâches assurées au LORIA]**

- /1/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'inversion acoustique articulaire
- /2/ Affinement du modèle de contrôle

#### **[Tâches assurées à l'IPS et à la MISHA]**

- /1/ Développement du système de gestion de la base de données
- /2/ Intégration des données et des outils dans la base
- /3/ Evaluation du système de gestion de la base

### **Réunions de travail et publications**

- /1/ Deux réunions, dont une à l'IPS et l'autre à l'ICP : courant 2009
- /2/ Accès de la communauté scientifique à la base
- /3/ Publications communes sur : la coarticulation, l'inversion et l'évaluation de modèles physiques
- /4/ Rédaction d'un rapport partiel sur l'avancement du projet

### **Année 4 - 2009/2010 :**

#### **[Tâches assurées à l'IPS]**

- /1/ Réalisation manuelle de nouveaux tracés
- /2/ Numérisation des tracés manuels
- /3/ Mesures et élaboration de contraintes articulatoires et acoustiques de la coarticulation

Livraison : films synchronisés et tracés manuels numérisés à l'ICP

#### **[Tâches assurées à l'ICP]**

- /1/ Etiquetage des films
- /2/ Affinement de l'outil d'extraction automatique des contours
- /3/ Extraction automatique des contours
- /4/ Développement de la plateforme logicielle
- /5/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'élaboration et l'évaluation de modèles physiques

Livraison : films numérisés, tracés manuels numérisés et contours extraits automatiquement à l'IPS – la MISHA et au LORIA

#### **[Tâches assurées au LORIA]**

- /1/ Exploitation des images cinéradiographiques, des tracés et des contours pour l'inversion acoustique articulaire
- /2/ Affinement du modèle de contrôle

#### **[Tâches assurées à l'IPS et à la MISHA]**

- /1/ Développement du système de gestion de la base de données
- /2/ Intégration des données et des outils dans la base

/3/ Evaluation du système de gestion de la base

## Réunions de travail, atelier et publications

- /1/ Deux réunions, dont une à l'IPS et l'autre au LORIA : courant 2010
- /2/ Accès de la communauté scientifique à la base
- /3/ Organisation d'un atelier-colloque international (lieu à déterminer) pour clôturer le projet
- /4/ Publications communes sur : la coarticulation, l'inversion et l'évaluation de modèles physiques
- /5/ Rapport final sur les productions complètes du projet

## B-3 – Bibliographie

### Publications sélectives relatives au projet de recherche

*NB : Le nom d'un participant au projet est mis en exergue.*

Abry, C., & Lallouache, T. (1995). Modeling lip constriction anticipatory behaviour for rounding, in French with the MEM (Movement Expansion Model). Proceedings of the 13<sup>th</sup> International Congress of Phonetic Sciences, Vol. 4 (pp. 152-155)

Arnal A. Badin P. **Brock G.** Connan P.-Y. Florig E. N. Perez, **P. Perrier**, P. Simon, **R. Sock**, L. Varin, **B. Vaxelaire** & **J.-P. Zerling**. (2000) Une base de données cinéradiographiques du français. XXIIIèmes Journées d'Etude sur la Parole, pages 425-428.

Asci A. **Vaxelaire B.** Hecker V. Hirsch F. Guedet M. (2003) Anticipatory and carryover coarticulation in Turkish. Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona, 447 - 450.

Atal B.S. Chang J.J. Mathews M.V. Tukey J.W. (1978) Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. Journal of the Acoustical Society of America, vol. 63, 1535 – 1555.

Badin P. Gabioud B. Beutemps D. Lallouache M.T. Bailly G. Maeda S. **Zerling J.-P. Brock G.** (1995) Cineradiography of VCV sequences : articulatory-acoustic data for a speech production model. In Proceedings of the 15<sup>th</sup> International Conference on Acoustics. Vol. IV, 349-352. Trondheim, Norway.

Badin, P., Bailly, G., Réveret, L., Baciou, M., Segebarth, C., & **Savariaux, C.** (2002). Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images. Journal of Phonetics, 30, 533-553.

Baer T., Gore J., Boyce S. & Nye P. (1991) Analysis of vocal tract shape and dimensions using magnetic resonance imaging: vowels. Journal of the Acoustical Society of America, 90, 799-828.

Bailly G., Abry C., Boë L.J., Laboissière R., **Perrier P.** & Schwartz J.L. (1992). Inversion and speech recognition. Dans J. Vandewalle, R. Boite, M. Moonen & A Oosterlink (Eds.), *SIGNAL PROCESSING VI: Theories and Applications* (pp. 159-164). Elsevier Science Publishers B.V

Benguérel A.P. Adelman S. (1976) Perception of coarticulated lip rounding. *Phonetica* 33, 113-126.

- Berthommier F. (2004) Characterization and extraction of mouth opening parameters available for audiovisual speech enhancement. In Proc. Int. Conf. on Acoustics, Speech and Signal Processing, volume 3, pages 789- 792.
- Bothorel, G. **Brock** & Maillard-Salin, 1980. Mouvement des lèvres, déplacement du larynx et variation de la longueur du conduit vocal. Séminaire international sur la la bialité. Lannion, février 1980.
- Bothorel, C. **Brock**, G. Maillard-Salin, 1980. Contribution à l'étude des rapports entre les mouvements de l'os hyoïde et le déplacement du larynx. TIPS 12, 225-269.
- Bothorel A. Simon P. **Wioland F. Zerling J.-P.** (1986) Cinéradiographie des voyelles et consonnes du français. Recueil de documents synchronisés pour quatre sujets : vues latérales du conduit vocal, vues frontales de l'orifice labial, données acoustiques. Institut de Phonétique de Strasbourg.
- Bothorel A. Vaxelaire B. Sock R. Zerling J.-P. (1995) Consonantal intervocalic coarticulation and timing constraints. In Dynamic constraints and motor controls (Ed. G. BAILLY). Speech Maps - Mapping of Action and Perception in Speech. Sound-to-Gesture Inversion in Speech - ESPRIT Basic Research Project N° 6975. Vols. IV & V, 17-25.**
- Brock G.** (1977) Méthode de synchronisation graphique image/son pour l'exploitation des films radiologiques. Présentation de l'appareillage réalisé à l'Institut de Phonétique de Strasbourg. Travaux de l'Institut de Phonétique de Strasbourg 9, 221-232.
- Brock G.**, 1977. Méthode de synchronisation graphique images-son pour l'exploitation des films radiologiques. Présentation de l'appareillage réalisé à l'Institut de Phonétique de Strasbourg. TIPS 9, 221-232.
- Brock**, F. **Wioland.** & P; Simon, 1978. La chaîne de segmentation et l'analyse en temps réel. Présentation et application. TIPS 10, 133-164.
- Brock G.**, Najafizadeth H., Simon P., & Wolff F., 1984. Numérotation et codage synchrones graphique et phonique pour l'exploitation des films radiologiques. TIPS 16, 181-190.
- Brock G.**, Najafizadeth H., Simon P., & Wolff F., 1987. Numérotation synchronisée radiologique. Radiologie, 381-383.
- Brock G.**, 1989. Optimisation d'une station analogique de traitement du signal acoustique. Mélanges Péla SIMON, 147-155.
- Browman, C.P., & Goldstein, L.M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252.
- Browman C.P. & Goldstein L.M. (1990). Gestural Specification Using Dynamically-Defined Articulatory Structures. *J. Phonetics*, 18, 299-320.
- Cathiard, M.A. (1994). La perception visuelle de l'anticipation des gestes vocaliques. Doctorat UPMF, Grenoble.
- Chen J. Berger M.-O. **Laprie Y.** (2004) An Effective Lip Tracking Algorithm for Acoustic-to-Articulatory Inversion. – In : 5th International Workshop on Image Analysis for Multimedia -WIAMIS'2004, Lisbon, Portugal.
- Clements, N. 1992. Phonological Primes: Features or Gestures? *Phonetica* 49,181-193.
- Elisei, F., M. Odisio, G. Bailly and P. Badin (2001). Creating and controlling video-realistic talking heads. Proceedings of the Auditory-Visual Speech Processing Workshop (pp. 90-97). Scheelsminde, Denmark.
- Engwall, O. (2003). Combining MRI, EMA & EPG in a three-dimensional tongue model. *Speech Communication*, 41(2/3), 303-329.
- Fant G. (1960) Acoustic Theory of Speech Production. Mouton, La Hague, The Netherlands.

- Feldman A.G. (1986). Once more on the Equilibrium-Point hypothesis (I model) for motor control. *Journal of Motor Behavior*, Vol. 18, 1, 17-54.
- Flament B. (1984) Recherche sur la mise en relief en français. Approche théorique et essai de caractérisation phonétique à partir de données de la mingographie et de la radiocinématographie. Doctorat d'Etat, Institut de Phonétique - Université des Sciences Humaines de Strasbourg.
- Fontecave J. **Berthommier F.** (2005) Quasi-automatic extraction method of tongue movement from a large existing speech cineradiographic database. In Proc. Eur. Conf. on Speech Communication and Technology, pages 1081-1084.
- Fontecave J. & **Berthommier F.** (2006). Extraction des mouvements du conduit vocal à partir de données cinéradiographiques. *Actes des XXVIèmes Journées d'Étude sur la Parole*, Association Francophone de la Communication Parlée, Dinard – Juin 2006 (A paraître)
- Gérard, J.-M., **Perrier, P.** & Payan, Y. (2006, in press). 3D biomechanical tongue modelling to study speech production. In *Speech Production: Models, Phonetic Processes, and Techniques* (J. Harrington & M. Tabain, editors), New York: Psychology Press.
- Hecker V. **Vaxelaire B. Sock R.** Cathiard M-A (2000), Les effets perceptifs de gestes anticipatoires en français ", *Travaux de l'Institut de Phonétique de Strasbourg*, n° 29, 1-29.
- Hecker V. **Vaxelaire B.** Cathiard M. **Savariaux C. Sock R.** (2001) How lip protrusion expansion influences auditory perceptual extent. Probing into the Movement Expansion Model. Orage 2001 - Oralité et Gestualité : Communication Multimodale. Congrès International à Aix-en-Provence, juin 2001., Editions L'Harmattan, Paris, 450 - 456.
- Hecker-Ferbach V. (2002) La perception auditive de l'anticipation des gestes vocaliques en français. Doctorat des Sciences du Langage. Institut de Phonétique de Strasbourg – E.A. 3403. Université Marc Bloch (dir. **R. Sock**).
- Heinz J. M. Stevens K. N. (1964) On the derivation of area functions and acoustic spectra from cineradiographic films of speech. *Journal of the Acoustical Society of America* 36, 1037.
- Hirsch, F. **Sock, R.**, Connan, P.Y., & **Brock, G.** (2003). Auditory effects of anticipatory rounding in relation with vowel height in French. *Actes du 15<sup>ème</sup> Congrès International des Sciences Phonétiques* (pp. 1445-1448). Barcelona, Espagne.
- Ladefoged, P., & Maddieson, I (1996). *The Sounds of the World's Languages*. Oxford: Blackwell Publishers.
- Laprie Y.** & Berger M.-O. (1996) Extraction of Tongue Contours in X-Ray Images with Minimal User Interaction. *Proceedings of ICSLP'96* (vol 1. pp.268-271).
- Laprie Y. Ouni S.** (2002)\_ Introduction of constraints in an acoustic-to-articulatory inversion. – In : 7th International Conference on Spoken Language Processing - ICSLP 2002, Denver, USA. – septembre 2002.
- Laprie Y. Ouni S.** Maeda S. Potard B. (2003) Inversion experiments based on a descriptive articulatory model. – In : 6th International Seminar on Speech Production. – Sydney, Australia, décembre 2003.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773-1781.
- Løevenbruck, H., & **Perrier, P.** (1993). Vocalic reduction: prediction of acoustic and articulatory variabilities with invariant motor commands. *Actes de EUROSPEECH 93* (pp. 85-88). Berlin.
- Løevenbruck, H., **Perrier, P.**, & Ostry, D.J. (1995). Equilibrium Point Hypothesis and articulatory targets in speech : discussion from simulations of empirical data using a biomechanical model of the jaw. *Actes du XIII<sup>ème</sup> Congrès International des Sciences Phonétiques* (Vol. 2, pp. 462-465). Stockholm, Suède.
- Löfqvist. A. & H. Yoshioka. Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America* 68(3), 792-801

- Löfqvist A. (1990) Speech as audible gestures. In *Speech Production and Speech Modeling*. Hardcastle W.J. & Marchal A. (Eds), Dordrecht: Cluwer, 289 – 322.
- Lubker J.F. Lindgren R. (1982) The perceptual effects of anticipatory coarticulation. In P. Hurme (Ed.), *Papers in Speech Research*, Institute of Finnish Language and Communication, University of Jyväskylä, 252-271.
- Maeda S. (1988) Improved articulatory model. *Journal of the Acoustical Society of America*, vol. 65, S22.
- Maeda S. (1989a) Articulation compensatoire des voyelles : analyse de données cinéradiographiques avec un modèle linéaire. In *Mélanges offerts à Péla Simon*. Publications de l'Institut de Phonétique de Strasbourg, vol. 2, 545 – 562.
- Maeda S. (1989b) Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal tract shapes using an articulatory model. In Hardcastle W. & Marchal A. (Eds.) *Speech Production and Modelling* (pp. 131-149). Kluwer: Academic Publishers.
- Munhall, K., & Löfqvist, A. (1992). Gestural aggregation in speech: Laryngeal gestures. *Journal of Phonetics*, 20, 111-126
- Munhall, K.G., Vatikiotis-Bateson, E., & Tohkura, Y. (1995) X-ray Film database for speech research. *Journal of the Acoustical Society of America*. 98, 1222-1224.
- Ouni S. Laprie Y.** (2000) Improving acoustic-to-articulatory inversion by using hypercube codebooks. In : *International Conf. on Spoken Language Processing - ICSLP2000*, Beijing, Chine, II, pp. 178–181.
- Ouni S. Laprie Y.** (2001) Exploring the null space of the acoustic-to-articulatory Inversion using a hypercube codebook. In : *Proceedings of the 7th European Conference on Speech Communication and Technology - EUROSPEECH'01*, Aalborg, Denmark.
- Ouni S. Laprie Y.** (2003) A study of the French Vowels Through The Main Constriction of the Vocal Tract Using an Acoustic-to-articulatory inversion method ». – In : *15th International Congress of Phonetic Sciences 2003 - ICPHS'2003*, Barcelone, Espagne. – aug 2003.
- Payan, Y., & **Perrier, P.** (1997). Synthesis of V-V Sequences with a 2D Biomechanical Tongue Model Controlled by the Equilibrium Point Hypothesis. *Speech Communication*, 22, (2/3), 185-205.
- Pelorsson, X., Hirschberg, A., Wijnands, A.P.J., Bailliet, H., Vescovi C., & Castelli E. (1996). Description of the flow through the vocal cords during phonation. Application to voiced sounds synthesis. *Acta Acustica*, 82, 358-361.
- Perkell J. S. (1969) *Physiology of speech production: results and implications of a quantitative cineradiographic study*. MIT Press.
- Perkell, J.S., & Klatt, D.H. (1986). *Invariance & Variability in speech processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Perrier P.** Boë L.-J. **Sock R.** (1992) Vocal tract area function estimation from midsagittal dimensions with CT scans and a vocal tract cast: Modelling the transition with two sets of coefficients. *J. Speech Hearing Res.* 35, 53-67.
- Perrier, P.**, Løevenbruck, H., & Payan, Y. (1996). Control of tongue movements in speech: The Equilibrium Point hypothesis perspective. *Journal of Phonetics*, 24, 53-75.
- Perrier P.**, Laboissière R., Abry C. & Maeda S. (1997) *Speech Production: Models and Data* (Editorial). *Speech Communication*, 22 (2/3), 89-92

**Perrier, P.**, Payan, Y., Perkell, J.S., Zandipour, M., Pelorson, X., Coisy, V. & Matthies, M. (2000) An attempt to simulate fluid-walls interactions during velar stops. In Proceedings of the 5th Seminar on Speech Production : Models and Data & CREST Workshop on Models of Speech Production : Motor Planning and Articulatory Modelling, vol., pp. 149-152. Kloster Seeon, Germany.

**Perrier P.**, Payan Y., Zandipour M. & Perkell J. (2003) Influences of tongue biomechanics on speech movements during the production of velar stop consonants: A modeling study. *Journal of the Acoustical Society of America*, 114(3), 1582-1599.

**Perrier, P.** (2006, in press). About speech motor control complexity. In *Speech Production: Models, Phonetic Processes, and Techniques* (J. Harrington & M. Tabain, editors), New York: Psychology Press.

Potard B. **Laprie Y.** (2005) Using phonetic constraints in acoustic-to-articulatory inversion. In : Interspeech, Lisboa.

Roy (2004) Etude de la perception des gestes anticipatoires d'arrondissement par les sourds et les malentendants. Doctorat des Sciences du Langage. Institut de Phonétique de Strasbourg – E.A. 3403. Université Marc Bloch (dir. **R. Sock**)

Saltzman E.L. & Munhall, K.G. (1989). A Dynamical Approach to Gesture Patterning in Speech Production. *Ecological Psychology*, 1, 1615-1623.

Simon P. (1967) Les consonnes Françaises. Mouvements et positions articuloires à lumière de la radiocinématographie. Paris: Klincksieck.

Simon, G. **Brock** & M.H. Han, 1978. Description et utilisation d'un équipement à rayons X pour l'étude de certains aspects articuloires. Application au coréen in *Articulatory Modeling and Phonetics 1978* (pp. 223-242).

**Sock R.** (1995) Löfqvist A. Some Timing Constraints in the Production of Bilabial Stops. *Journal of Phonetics* 23, 129-138.

**Sock R.** (1998) Organisation temporelle en production de la parole. Emergence de catégories sensori-motrices phonétiques. Doctorat d'Etat ès Sciences Humaines. Université Stendhal - Grenoble III (479 p.). Presses du Septentrion.

**Sock R.** (2001) La Théorie de la Viabilité en production-perception de la parole. In KELLER D. DURAFOUR J.-P. BONNOT J.-F. & SOCK R. (Eds.), *Psychologie et Sciences Humaines*, Mardaga, Liège, 285 – 316.

**Sock R.** Cathiard M.A. Hecker V. (1999) The perceptual effects of anticipatory vocalic gestures in French. 14th International Congress of Phonetic Sciences, San Francisco, 1-7 August, 1999, Vol. 3, 2057-2060.

**Sock R. Vaxelaire B.** (2001) Peut-on travailler sans représentations en production-perception de la parole ? In Par monts et par vaux. Itinéraires linguistiques et grammaticaux. Mélanges de linguistique générale et française offerts au Professeur Martin RIEGEL. In BURIDANT C. KLEIBER G. & PELLAT J.-C. (Eds.), PEETERS, Liège, Paris, 379-391.

**Sock R. Vaxelaire B.** Hecker V. Hirsch F. (2002) What relationship between protrusion anticipation and auditory perception ? 7<sup>th</sup> International Conference on Spoken Language Processing (ICSLP), septembre 16-20, Denver, Colorado, USA, 1677-1680.

- Sock R.** (2003) Comprendre les gestes anticipatoires audibles et inaudibles. In *Stratégies et parcours. De l'anticipation à la didactique du FLE. Sciences Cognitives, Linguistique et Intelligence Artificielle (SCOLIA) 17*. Eds. 11 – 28.
- Sock R. Vaxelaire B.** (Eds.) *L'anticipation à l'horizon du présent*. 282 pages. Psychologie et Sciences Humaines, Mardaga, Liège, 2004.
- Perkell J.S. (1969) *Physiology of Speech Production*. Massachusetts Institute of Technology: Cambridge, Ma, USA.
- Stevens K.N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In David Jr. E.E. & Denes P.B. (Eds;), *Human Communication: A unified view* (pp. 51-66). New-York: Mc Graw Hill.
- Stevens K.N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.
- Straka G. (1965) *Album Phonétique*. Presses de l'Université Laval, Québec
- Thimm G. (1999) Segmentation of X-ray image sequences showing the vocal tract. IDIAP Research Report, IDIAP, Suisse.
- Thimm G. Luetin J. (1999) Extraction of articulators in Xray image sequences. In *Proc. Eur. Conf. on Speech Communication and Technology*, pages 157-160.
- Tiede M. & Vatikiotis-Bateson E. (1994) Extracting articulator movement parameters from a videodisc-based cineradiographic database. *Proceedings of ICSLP 94* (pp.45-48)
- Vaxelaire B. Sock R.** (2000) Gestural blending, place of articulation and speech rate. 6<sup>th</sup> International Conference on Spoken Language Processing. Beijing, China, oct. 16-20.
- Vaxelaire B. Sock R.** (1999) Perceiving anticipatory phonetic gestures in French. 6<sup>th</sup> European Conference on Speech Communication and Technology, Budapest, Hongrie, Sept. 5-10, 1999, Vol. 1, 165-168.
- Vaxelaire B. Sock R.** & Hecker V. (1999) The perceptual effects of anticipatory gestures in the production of French rounded vowels. Eurospeech'99. 6<sup>th</sup> European Conference on Speech Communication and Technology, 5-10 septembre 1999, Budapest, Hongrie, vol.1, 165-168.
- Vaxelaire B. Sock R.** (1996) A Cineradiographic and Acoustic Study of Velar gestures in French. First European Speech Communication Association (ESCA) Tutorial and Research Workshop, Autrans-France, 65-68.
- Vaxelaire B. Sock R.** (1997) Laryngeal movements and speech rate. An X-ray investigation. Eurospeech '97. 5<sup>th</sup> European Conference on Speech Communication and Technology, Rhodes, Greece. Vol. 2, 1039-1042.
- Vaxelaire B., Sock R., Bonnot J.F. & Keller D.** (1999) Anticipatory labial activity in the production of French rounded vowels. *Proceedings of ICPhS 99* (Vol. 1., pp. 53-56).
- Westbury J.R., Turner G. & Dembovski J. (1994) X-ray microbeam speech production database users' handbook. Waisman Center, Université du Wisconsin.
- Vilain, C., Pelorson, X., Hirschberg, A., Le Marrec, L., Op't Root, W. & Willems, J.F.H. (2003). Contribution to the physical modeling of the lips. Influence of the mechanical boundary conditions. *Acta Acustica*, **89**, 882-887.
- Wioland F.** (1985) *Faits de jointure en français. Implications aux niveaux articulatoire et acoustique. Incidences sur le plan des fonctions linguistiques*. Doctorat d'Etat, Institut de Phonétique - Université des Sciences Humaines de Strasbourg.

Wood S.A.J. (1979a) A radiographic examination of constriction location for vowels. *Journal of Phonetics*, 7, 25-43

Wood S.A.J. (1979b) A cinefluorographic study of the temporal organization of articulator gestures: Examples from Greenlandic. *Speech Communication*, 22, 207-225.

Wrobel-Dautcourt B. M. O. Berger, B. Potard, **Y. Laprie, S. Ouni** (2005) A low cost stereovision based system for acquisition of visible articulatory data. In : *Proceedings of International Conference on Auditory-Visual Speech Processing (AVSP'05)*, pp. 145–150.

**Zerling J.-P.** (1979) Articulation et coarticulation dans des groupes occlusive-voyelle en français. Etude cinéradiographique et acoustique : contribution à la modélisation du conduit vocal. Doctorat 3<sup>e</sup> Cycle, Institut de Phonétique, Université de Nancy II.

#### **B-4 – Les collaborations internationales prévues**

Des échanges soutenus se dérouleront entre les quatre membres partenaires du projet et des laboratoires étrangers. Ainsi, l'affinement des logiciels au sein de la plateforme de traitement multimédia suscitera une mise à jour de nos collaborations avec les partenaires étrangers, préoccupés par la valorisation et la distribution de bases de données cinéradiographiques :

- 1 • Munhall et ses collègues [Mun95] (à l'Université de Kingston au Canada) pour leur travail de sauvegarde et de distribution de données cinéradiographiques réalisées en Amérique du Nord, essentiellement à l'Université Laval de Québec ;

- 2 • Hoole pour son travail de sauvegarde et de traitement de données cinéradiographiques pathologiques et d'IRM, disponibles à l'Institut de Phonétique de Munich ;

- 3 • Tiede et Bateson [Tie94] à ATR (Advanced Technological Research à Kyoto, au Japon) et aux Laboratoires Haskins (New-Haven, Connecticut aux U.S.A).

Ces collaborations, que nous avons déjà avec ces laboratoires, seront re-ciblées dans le domaine de la valorisation de données, cela dès le début du projet. Les échanges se feront largement dans le cadre de *workshops* sur la valorisation de bases de données.

#### **B-5 – Justification scientifique des moyens demandés pour chaque équipe partenaire impliquée dans le projet.**

*On présentera ici une justification scientifique des moyens demandés pour chacun des partenaires impliqués dans le projet, en distinguant les demandes en équipement, fonctionnement, personnels. Pour les demandes d'équipement, préciser si les achats envisagés doivent être complétés par d'autres sources de crédits, le montant et l'origine des crédits complémentaires qui seront utilisés.*

Partenaires	Fonctionnement	Equipement	Missions	Post-doc	CDD	Prestations de service externe
Partenaire 1 IPS	*45100 euros	**39100 euros	***28300 euros	1 post-doc / 48 mois à <b>37017</b> euros / an	2 CDD / 48 mois – salaire mensuel = indice majoré = 338 (premier échelon assistant ingénieur)	****12000 euros  600 euros / film X 20 films
Partenaire 2 ICP	0 euros	7000 euros (2 PC)	1000 euros / an	1 post-doc / 12 mois à <b>37017</b> euros / an		
Partenaire 3 LORIA	0 euros	7000 euros (2 PC)	1000 euros / an	1 post-doc / 24 mois à <b>37017</b> euros / an		

### Partenaire 1 : IPS

#### **\*Fonctionnement (hors vacances) :**

Logiciels :

#### Année 1 :

- Achat MATLAB™ (multi-postes) et Logiciel de statistiques : 1,1 K euros
- Achat de Adobe Photoshop™ CS2 (Windows) French 1 K euros

#### Années 2, 3 et 4 :

- Maintenance des logiciels 9 K euros
- Organisation d'ateliers et de colloques 15 K euros

Autres frais de fonctionnement (consommables, téléphones, photocopies, etc.) 10 K euros

#### **\*\*Equipement :**

- 1 enregistreur / lecteur DVD et moniteur de contrôle : 7 K Euros
- 2 stations de travail (Mac) pour le traitement des données cinéradiographiques + scanner + imprimante couleur : 10 K euros
- 4 stations de travail (PC) pour le traitement des données cinéradiographiques (synchronisation images/son ; détection automatique des contours) + carte d'acquisition vidéo MATROX + carte Sound Blaster® Audigy™ 4 Pro: 14 K euros

- 1 station de travail (PC) pour la numérisation des croquis :

3,5 K euros

**\*\*\* Missions :**

Année 1 : 4000 euros

Année 2 : 8100 euros

Année 3 : 8100 euros

Année 4 : 8100 euros

**\*\*\*\* Prestations de service externe**

Pour conformation sur standard vidéo Betacam SP des synchronisations images/son ; réalisation de copies 35 mm sur standard vidéo Betacam SP et DV (Centre Image à Paris).

1 informaticien-phonéticien PostDoctorant (48 mois) : transfert des films 35 mm sur standard vidéo BetacamSP ; synchronisation des images et du son ; étiquetage des films ; extraction automatique de contours et intégration dans la base de données (valorisation de la base sous la direction de l'IPS et de la MISHA)

1 CDD (48 mois) expert phonéticien : synchronisation des images et du son ; étiquetage des films ; inventaire des tracés existants ; réalisation de croquis ; numérisation de croquis ; mesures et analyse des faits coarticulatoires

1 CDD (48 mois) expert phonéticien : synchronisation des images et du son ; étiquetage des films ; réalisation de croquis ; numérisation de croquis ; mesures et analyse des faits coarticulatoires.

**Partenaire 2 : ICP**

1 station de travail (PC) : mise en forme des données ; numérisation des films radiographiques ; développement de la plateforme logicielle pour le traitement des signaux de mouvements et des signaux acoustiques

1 station de travail (PC) : détection automatique de contours radiographiques

1 Postdoctorant (12 mois) : exploitation des données cinéradiographiques pour élaborer et évaluer des modèles de contrôle (évaluation de modèles de génération de séquences de parole sur des modèles physiques.

**Partenaire 3 : LORIA**

1 station de travail (PC) : développement de méthodes d'inversion

1 station de travail (PC) : élaboration d'un modèle de contrôle

1 Postdoctorant (24 mois) : exploitation des données cinéradiographiques pour l'inversion acoustique articulatoire ; élaboration d'un modèle de contrôle

### **Propositions d'experts et confidentialité**

Les membres du comité d'évaluation et du comité de pilotage sont astreints à la confidentialité.

- Possibilité de fournir une liste de 3 à 5 noms d'experts français ou étrangers (avec coordonnées complètes : adresse postale et adresse électronique) susceptibles d'évaluer le projet avec lesquels les équipes participant au projet n'ont ni conflit d'intérêt, ni collaborations en cours.
- Possibilité éventuelle de fournir une liste de 5 noms max. d'experts auxquels les participants au projet ne souhaitent pas que le projet soit envoyé s'il y a risque de conflits d'intérêts.

/1/

M. François PELLEGRINO  
Directeur du Laboratoire Dynamique Du Langage  
UMR 5596 CNRS - Université Lumière Lyon 2  
Tel. (+33/0)4 72 72 64 94  
DDL - ISH  
14, av. Berthelot  
69363 Lyon Cedex 7 / France  
[Francois.Pellegrino@univ-lyon2.fr](mailto:Francois.Pellegrino@univ-lyon2.fr)

/2/

M. Anders LÖFQVIST  
Senior Scientist, [Haskins Laboratories](#)  
Associate Editor for Speech Production, Journal of the [Acoustical Society of America](#)  
Haskins Laboratories  
300 George Street  
New Haven, CT 06511  
Haskins Phone: (203) 865-6163, ext. 232  
Yale Phone: (203) 764-9353  
Fax: (203) 865-8963  
[lofqvist@haskins.yale.edu](mailto:lofqvist@haskins.yale.edu)

/3/

Mme Alice TURK  
Reader  
Linguistics and English Language  
The University of Edinburgh  
Adam Ferguson Building  
40 George Square  
Edinburgh EH8 9LL  
Tel: +44 131 650 3483  
Fax: +44 131 650 3962  
[turk@ling.ed.ac.uk](mailto:turk@ling.ed.ac.uk)

## **Programmes SHS 2006**

### **C - Moyens financiers et humains demandés par chaque équipe partenaire du projet**

Chaque équipe partenaire remplira une fiche de demande d'aide selon les modèles proposés ci-dessous (laboratoire public ou fondation ; entreprise ou association) en fonction de son appartenance.

# Programmes SHS 2006

## Fiche de demande d'aide Laboratoire public / Fondation

Acronyme ou titre court du projet

Partenaire 1 - Coordinateur (nom, prénom) :

### Calcul de l'aide demandée à l'ANR et estimation du coût complet du projet pour le laboratoire du partenaire

*Avant de remplir ce tableau il vous faut décider quel sera votre établissement gestionnaire (cf notes 5 et 6 en bas de page)*

				Euros HT	Taux spécifiques à chaque établissement	
	Nbre Homme .mois	Coût Homme.mois (salaire chargé)	Nombre de personnes impliquées			
<b>Dépenses de personnel</b> <sup>(1)</sup> catégorie 1 catégorie 2 etc...				(P1)	Taux Env	$P = (P1) * \text{Taux Env}$
<b>Dépenses de personnel non permanent à recruter</b> <sup>(2)</sup> catégorie 1 catégorie 2 etc...				(Q1)	Taux Env	$Q = (Q1) * \text{Taux Env}$
<b>Equipements (&gt;4000 €)</b> détail § B-5				(R1)	Taux TVA non réc.	$R = (R1) * (1 + \text{Taux TVA non réc})$
<b>Petits matériels, consommables, fonctionnement, etc</b>				(S1)	Taux TVA non réc.	$S = (S1) * (1 + \text{Taux TVA non réc})$
<b>Frais de missions</b> si montant > 5% de la somme demandée, justification § B-5				(T1)	Taux TVA non réc.	$T = (T1) * (1 + \text{Taux TVA non réc})$
<b>Prestations de service externes, sous-contractant</b> <sup>(3)</sup>				(U1)	Taux TVA non réc.	$U = (U1) * (1 + \text{Taux TVA non réc})$
<b>Total des dépenses de fonctionnement</b>				$(X1) = S1 + T1 + U1$		$X = X1 * (1 + \text{Taux TVA non réc})$
<b>Frais généraux (assistance, encadrement, coût de structure) (max 4 % du coût total des dépenses)</b>						$Y = 4 \% (Q + R + X)$
<b>Assiette de l'aide</b> <sup>(4)</sup>						$Z = Q1 + R + X + Y$
<b>Aide demandée</b> $\leq ZZ$ <sup>(5)</sup>						<b>Aide demandée</b>

Coût complet du projet <sup>(6)</sup>

$$CC = P + Q + R + X + Y$$

(1) Il s'agit du personnel qui serait affecté au projet mais qui est présent dans le laboratoire ou l'entreprise indépendamment de la réussite de l'appel de l'agence. Salaire mensuel chargé (charges salariales et patronales).

Pour les enseignants-chercheurs ne compter que la part salariale correspondant à la part recherche (50% du salaire pour 100% de temps consacré à la recherche).

5 grandes catégories (CDD ou CDI) : Ingénieur, chercheur, enseignant chercheur, technicien, autres. Lorsque dans une même catégorie, plusieurs personnes de salaire différent sont mentionnées indiquer la valeur moyenne. Pour les laboratoires publics ou fondations, ces données ne servent qu'à calculer le coût complet du projet.

- (2) Personnel non statutaire directement affecté au projet exprimé en hommes mois. Les dépenses éligibles se limitent aux salaires et aux charges sociales. Exemple : post-doc (catégorie 1), ingénieur d'études (catégorie 2), etc.
- (3) Propriété intellectuelle, location de matériel, service, etc.
- (4) Assiette de l'aide signifie dépenses éligibles à une aide de l'agence.
- (5) L'aide demandée doit correspondre au montant HT augmenté éventuellement de la TVA non récupérable. La TVA non récupérable est actuellement, par exemple, de 88% pour le CNRS et l'INRA, de 94% pour l'Inserm et de 100% pour les universités. En conséquence pour une demande qui sera gérée par l'INRA, le taux de TVA non récupérable est  $0,88 \times 0,196 = 0,1725$ , ce qui conduit à inscrire dans la colonne de droite pour une demande HT de 10 000 euros,  $10000 \times (1 + 0,1725)$  soit 11 725 euros soit une demande d'aide de 11 725 euros si le partenaire veut disposer de 10 000 euros dans la réalisation de son projet.  
En cas d'aide accordée par un autre financeur sur les mêmes dépenses que celles listées dans le tableau, il peut y avoir une diminution de l'aide accordée par l'ANR pour rester conforme à la réglementation.
- (6) Pour le calcul en coût complet, il faut augmenter le salaire chargé d'un taux d'environnement, qui tient compte des conditions d'environnement des personnels (infrastructure, par exemple). Par exemple, ce taux est à l'heure actuelle de 1,8 pour l'Inserm et le CNRS.

# Programmes SHS 2006

## Fiche de demande d'aide Entreprise / Association

Acronyme ou titre court du projet

Partenaire n°                      Responsable scientifique (nom, prénom) :

Calcul de l'aide demandée à l'ANR et estimation du coût complet du projet pour le partenaire :

				Euros HT
	Nbre Homme. mois	Coût Homme. mois Salaire chargé	Nombre de personnes impliquées	
<b>Dépenses de personnel</b> <sup>(1)</sup> catégorie 1 catégorie 2 etc.				(P)
<b>Dépenses de personnel non permanent à recruter</b> <sup>(2)</sup> catégorie 1 catégorie 2 etc.				(Q)
<b>Amortissements des équipements (&gt;4000 €)</b> Nature et justification de la dépense				(R)
<b>Petits matériels, consommables, fonctionnement, etc.</b>				(S)
<b>Frais de missions</b> si montant >5% de la somme demandée, justification de la dépense				(T)
<b>Prestations de service externes</b> <sup>(3)</sup> , sous-contractant				(U)
<b>Prestation de service interne à l'entreprise ou à l'organisme</b>				(V)
<b>Total frais fonctionnement</b>				(x) =S+T+U+V
<b>Frais généraux (assistance, encadrement, coût de structure)</b> <sup>(4)</sup>				(Y)
<b>Coût complet du projet</b>				<b>CC= P+Q+R+X+Y</b>
<b>Assiette de l'aide</b> <sup>(5)</sup>				<b>CC</b>
<b>Aide demandée</b> <sup>(6)</sup> Se référer à la notice explicative				<b>Aide demandée</b>

- (1) Il s'agit du personnel qui serait affecté au projet mais qui est présent dans le laboratoire ou l'entreprise indépendamment de la réussite de l'appel de l'agence. Salaire mensuel chargé (charges salariales et patronales). Pour les enseignants-chercheurs ne compter que la part salariale correspondant à la part recherche (50% du salaire pour 100% de temps consacré à la recherche).  
5 grandes catégories (CDD ou CDI) : Ingénieur, chercheur, enseignant chercheur, technicien, autres. Lorsque dans une même catégorie plusieurs personnes de salaire différent sont mentionnées indiquer la valeur moyenne. Pour les laboratoires publics ou fondation, ces données ne servent qu'à calculer le coût complet du projet.
- (2) Personnel non statutaire directement affecté au projet exprimé en hommes mois. Les dépenses éligibles se limitent aux salaires et aux charges sociales. Exemple : post-doc (catégorie 1), ingénieur d'études (catégorie 2), etc.
- (3) Propriété intellectuelle, location de matériel, service, etc.
- (4) Pour les associations et TPE, les frais généraux peuvent être au maximum = 4% de R + 8% de (P+Q+S+T+U). Pour les sociétés civiles, les entreprises hors TPE, les GIE, les centres techniques, les frais généraux peuvent être au maximum de = 7% de (R+S+T+U) + 68% de (P+Q)

- (5) Assiette de l'aide signifie dépenses éligibles à une aide de l'agence
- (6) En cas d'aide accordée par un autre financeur sur les mêmes dépenses que celles listées dans le tableau, il peut y avoir une diminution de l'aide accordée par l'ANR pour rester conforme à la réglementation.

## Programmes SHS 2006

### D - Récapitulatif global de la demande financière pour le projet

Acronyme ou titre court du projet
-----------------------------------

#### **a-Estimation du coût complet de cette demande**

*(reporter les valeurs (CC) des fiches des différents partenaires)*

	<b>Coût complet</b>
<b>Coordinateur (Partenaire 1)</b>	
<b>Partenaire 2</b>	
<b>Partenaire n</b>	
...	
<b>Total</b>	

#### **b-Total de l'aide demandée**

*(reporter les valeurs (Aide demandée) des fiches des différents partenaires)*

	<b>Aide demandée</b>
<b>Coordinateur (Partenaire 1)</b>	
<b>Partenaire 2</b>	
<b>Partenaire n</b>	
...	
<b>Total</b>	

#### **c- Effort en personnel demandé**

*(reporter les valeurs des fiches des différents partenaires)*

	<b>en homme/mois</b>
<b>Coordinateur (Partenaire 1)</b>	
<b>Partenaire 2</b>	
...	
...	
<b>Total</b>	

#### **d- Dépenses de fonctionnement**

*(reporter les valeurs des fiches des différents partenaires)*

	<b>en euros</b>
<b>Coordinateur (Partenaire 1)</b>	
<b>Partenaire 2</b>	
...	
...	
<b>Total</b>	

**Contrats publics et privés sur les trois dernières années (effectués et en cours)**

Nom du membre participant à cette demande	% d'implication	Intitulé de l'appel à projets Source de financement Montant attribué	Titre du projet	Nom du coordinateur	Date début - Date fin
SOCK R. VAXELAIRE B. ZERLING J.-P. WIOLAND F. BROCK G. PERRIER P. SAVARIAUX C.	30% 20% 10% 10% 10% 10%	ACI « Cognitique » ACT 1b, Financement : Ministère de la Recherche Partenaires : IPS & ICP  182,93 kEuros	<b><i>Contrôle des gestes anticipatoires en production- perception de la parole</i></b>	SOCK Rudolph	2001 - 2003
SOCK R. VAXELAIRE B. ZERLING J.-P. WIOLAND F. BROCK G. PERRIER P. SAVARIAUX C.	30% 20% 10% 10% 10% 10%	ACI Terrains, Techniques et Théories en Sciences Humaines et Sociales – Mesures et Données Financement : Ministère de la Recherche Partenaires : IPS & ICP  8 kEuros et Contrat assorti d'une allocation de recherche pour 3 ans	<b><i>Evaluation d'une plateforme multimedia pour l'analyse et la valorisation de données cinéroradiographiques et d'IRM</i></b>	SOCK Rudolph	2003 - 2006
SOCK R. VAXELAIRE B. ZERLING J.-P. BROCK G. PERRIER P. SAVARIAUX C.	30% 20% 10% 10% 10% 10%	Contrat de recherche de la MISHA UMS 2552  Partenaires : IPS, ICP & Laboratoire de Psychologie et de Neurocognition de Grenoble (LPNC)  4,65 kEuros en 2005 3,75 kEuros en 2006	<b><i>Anticiper pour construire des modalités sémiotiques</i></b>	SOCK Rudolph	2005 - 2008

--	--	--	--	--	--

**Contrats publics et privés sur les trois dernières années (effectués et en cours)**

Nom du membre participant à cette demande	% d'implication	Intitulé de l'appel à projets Source de financement Montant attribué	Titre du projet	Nom du coordinateur	Date début - Date fin
P. Perrier	50%	P2R France Allemagne (Projet de recherche en réseaux France – Allemagne) (CNRS – MAE) 75 kEuros (à répartir sur les 3 labos partenaires : ICP, Laboratoire de Phonétique de Paris et LIMSJ)	<b><i>POPAART : Facteurs physiques et de contrôle moteur dans les propriétés acoustiques et articulatoires de la parole : données physiologiques, expérimentations physiques, simulations et interprétation phonétiques</i></b>	Pascal Perrier	Janvier 2005 - Décembre 2008
M. Chabanas P. Perrier	50% 10%	Cluster « Informatique, Signal, logiciels embarqués » (Région Rhône – Alpes) Pour l'ICP : 30 kEuros en 2007	<b><i>Gestes Médico-Chirurgicaux Assistés par Ordinateur – GMCAO</i></b>	Jocelyne Troccaz	2005 - 2008
M. Chabanas P. Perrier	50% 10%	Projet soutenu par les fédérations CNRS ELESA et IMAG et par l'INRIA Rhône Alpes (75 kEuros à répartir sur les 3 partenaires : ICP, Laboratoire des <i>Techniques de l'Imagerie, de la Modélisation et de la Cognition</i> de Grenoble, et le laboratoire <i>GRAVIR</i> de l'INRIA Rhône Alpes)	<b><i>MIDAS : Modèles Interactifs Déformables pour l'Aide à la Surgétique</i></b>	Matthieu Chabanas	2005 - 2006

P. Perrier	30 %	EMERGENCE 2005 (Région Rhône – Alpes) 25 kEuros (à répartir sur les 4 partenaires : ICP, Laboratoire des <i>Techniques de l'Imagerie, de la Modélisation et de la Cognition</i> de Grenoble, le <i>Laboratoire d'Etudes des Interfaces et des Biofilms en Odontologie</i> de Lyon, et le Service de <i>Chirurgie Maxillo Faciale</i> du CHU Grenoble	<b><i>Applications des modèles physiques des articulateurs du conduit vocal au traitement de pathologies affectant la zone buccale.</i></b>	Pascal Perrier	Septembre 2004 – Août 2005
F. Berthommier	50 %	Projet CNRS-STIC Cognition et Traitement de l'Information 20 kEuros	<b><i>Etude psychophysique et modélisation computationnelle des mécanismes de décodage acoustico-phonétique à partir de la parole dégradée spectralement et temporellement</i></b>	Frédéric Berthommier	2002 - 2004

**Contrats publics et privés sur les trois dernières années (effectués et en cours)**

Nom du membre participant à cette demande	% d'implication	Intitulé de l'appel à projets Source de financement Montant attribué	Titre du projet	Nom du coordinateur	Date début - Date fin
Y. Laprie	10%	RIAM (116000€)	LABIAO (Lecture labiale assistée par ordinateur)	Philippe Suignard	01/05 – 12-06
Y. Laprie	20%	IST-FET STREP (120000€)	ASPI (Audiovisual Speech Inversion)	Yves Laprie	11/05-10/08

### Demandes de contrats en cours d'évaluation <sup>9</sup>

Nom du membre participant à cette demande	% d'implication	Intitulé de l'appel à projets Source de financement Montant demandé	Titre du projet	Nom du coordinateur
SOCK R. VAXELAIRE B. ZERLING J.-P. WIOLAND F. BROCK G.	30% 15% 10% 10% 30%	INTERREG III A Rhin Supérieur Centre-Sud  13 partenaires européens : France, Allemagne & Suisse Coût du projet : 654,9 kEuros Montant demandé : 327,45 kEuros	<i>Banque de données et réseau d'archivage de porteurs de projets culturels du Rhin Supérieur – DANOK</i>	SOCK Rudolph
SOCK R. VAXELAIRE B. ZERLING J.-P. BROCK G.	10% 10% 10% 10%	ANR 2006 Corpus Et Outils De La Recherche En Sciences Humaines Et Sociales  Montant demandé : 20 kEuros	<i>Corpus et Outils de Recherche pour l'Etude de la Coordination de l'Articulation et de la Phonation</i>	VAISSIERE Jacqueline

### Demandes de contrats en cours d'évaluation <sup>10</sup>

Nom du membre participant à cette	% d'impli-	Intitulé de l'appel à projets	Titre du projet	Nom du coordinateur
-----------------------------------	------------	-------------------------------	-----------------	---------------------

<sup>9</sup> Mentionner ici les projets en cours d'évaluation soit au sein de programmes du GIP-ANR, soit auprès d'organisme de fondations, à l'union européenne, etc. que ce soit comme coordinateur ou comme partenaire. Pour chacun donnez le nom de l'appel à projets, le titre du projet et le nom du coordinateur.

<sup>10</sup> Mentionner ici les projets en cours d'évaluation soit au sein de programmes du GIP-ANR, soit auprès d'organisme de fondations, à l'union européenne, etc. que ce soit comme coordinateur ou comme partenaire. Pour chacun donnez le nom de l'appel à projets, le titre du projet et le nom du coordinateur.

demande	ation	Source de financement Montant demandé		
S. Ouni, Y. Laprie	40%, 20%	ANR TecSan	ALANES (Renforcements Perceptifs et Visualisations Acoustiques et Articulatoires des Contrastes Phonétiques pour l'Acquisition du Langage par des Enfants Malentendants)	Anne Bonneau