

Gérard Bailly

DR2, HDR

<http://www.icp.inpg.fr/~bailly>

Institut de la Communication Parlée – UMR CNRS 5009

INPG – Université Stendhal, 46, av. Félix Viallet – 38031 GRENOBLE CEDEX

30 juillet 2003

## Rapport d'activité

2001 – 2003

### Sommaire

1	Avant-propos.....	1
2	Bilan scientifique.....	2
2.1	De l'équipe.....	2
2.2	Thèmes de recherche prospectifs.....	3
2.2.1	Communication multimodale.....	3
2.2.2	La salle MICAL : les modèles en action.....	4
3	Partie administrative.....	5
3.1	Administration de la recherche.....	5
3.1.1	Responsabilités locales.....	5
3.1.2	Responsabilités internationales.....	5
3.1.3	Organisation de congrès.....	5
3.2	Publications (2002-2003).....	5
3.2.1	Revue à comité de lecture.....	5
3.2.2	Ouvrages édités.....	5
3.2.3	Conférences invitées.....	5
3.2.4	Conférences à comité de lecture.....	5
3.2.5	Chapitre dans les ouvrages.....	6
3.3	Vulgarisation.....	6
3.4	Enseignements.....	6
3.5	Responsabilités de contrats et subventions.....	6
3.6	Encadrement.....	6
3.6.1	Thèses.....	6
3.6.2	DEAs.....	7
3.6.3	Autres stages.....	7
3.7	Participations à des jurys de thèses.....	7

## 1 Avant-propos

Cette période coïncide avec de multiples prises de responsabilité au sein de structures locales (notamment ELESa) et internationales (notamment le Special Interest Group (SIG) Synthèse de l'ISCA). Comme directeur adjoint d'ELESa et comme co-animateur du secteur expert « Information, Communication et Cognition », j'ai contribué à la l'identification dans le bassin grenoblois d'un pôle de compétence sur les « Objets, Agents & Environnements Communicants ». Avec mes collègues Jean Caelen de l'IMAG et Philippe Mallein (MSH-Alpes), nous avons multiplié les initiatives (Projet inter-fédérations PRESENCE, projets soumis au Pôle Européen...) afin de faire identifier par les tutelles (fédérations, universités) la nécessité de faire émerger cette thématique transversale, complémentaire aux deux « poids lourds » verticaux grenoblois que sont le logiciel et les micro-nanotechnologies. Comme représentant du CNRS, j'ai co-organisé avec P. Senn de France-Télécoms R&D la deuxième édition – à présent internationale – de la Conférence sur les Objets Communicants (sOc'2003) à Grenoble, rendez-vous bi-annuel que nous avons le bon espoir de pérenniser par la constitution d'un comité d'organisation permanent. Nous poursuivons nos efforts de structuration de cette activité au sein du pôle grenoblois notamment par un travail de formalisation d'un projet scientifique pour un groupement d'intérêt scientifique (PEGASUS). Le financement du Plan Pluri-Formation « Objets, Agents & Environnements Communicants » déposé par trois universités va nous permettre de financer des plate-formes sur l'étude, la simulation et l'évaluation de systèmes d'interaction humaine et de lancer cet automne un appel d'offres à projets autour de cette thématique. Je reviens plus largement sur ce projet scientifique dans la section 2.2.1.

Cette période coïncide aussi avec un fort développement des activités contractuelles de l'équipe ... avec un effectif restreint (deux permanents, trois thésards et un contractuel). L'équipe a ainsi largement contribué au budget non consolidé du laboratoire en enchaînant trois projets RNRT (TempoValse, ARTUS et EvaSy) et un projet ROBEA

(HR+) dont je suis le responsable scientifique ainsi que deux projets financ s par le CNRS (Cognitive et Jeune  quipe) dont Denis Beaudemont est le responsable.

Finalement, je travaille depuis deux ans sur le projet  ditorial du livre « Audiovisual Speech Processing ». Nous avons collect  toutes les contributions (avec un volume actuel de 250 pages) et commen ons le processus final de r vision du manuscrit. Gr ce   la large couverture des th matiques et au renom des contributeurs, ce livre devrait  tre une r f rence incontournable des travaux en traitement audiovisuel de la parole.

## 2 Bilan scientifique

### 2.1 De l' quipe...

L' quipe « Machines Parlantes » que j'anime d veloppe des syst mes de communication multimodaux inspir s du vivant, ce qui engendre une forte synergie entre exp rimentation, analyse, mod lisation et  valuation – boucle de conception de « machines parlantes » que j'ai d j  d crit dans mon pr c dent rapport d'activit  et qui a fait l'objet du cadrage de nos recherches pour le pr sent quadriennal. Deux th mes de recherche se sont particuli rement d velopp s lors de ces deux derni res ann es :

- D veloppement du mod le prosodique SFC : le syst me d'apprentissage automatique d velopp  par Bleike Holm dans le cadre de sa th se constitue une  tape capitale dans l'impl mentation d'un mod le de g n ration de la prosodie compatible avec les propositions th oriques de V. Auberg . Nous avons r solu le probl me – initialement mal pos  – de l'apprentissage automatique de ce mod le depuis la simple observation de r alisations o  se trouvent suffisamment instanci es les fonctions discursives que l'on d sire  tudier. Cet apprentissage n'est pas hi rarchique mais comp titif : il permet donc de n'avoir aucun a priori sur l'importance relative des fonctions ni sur les formes des contours qui les v hiculent. Lors du stage de Stefan Raitd de l'universit  de Dresden et en collaboration avec Hansjoerg Mixdorff de l'universit  de Berlin, nous avons pu tester ce mod le sur l'allemand et comparer les performances du SFC et du mod le MGI sur un corpus bilingue de formules math matiques oralis es (un article est en cours de soumission   la revue *Computer, Speech & Language*). Nous avons soumis un PAI dans le cadre du programme Picasso avec l'universit  de Vigo afin de confronter le mod le   un mod le de synth se bas -corpus et   une langue   accent libre. Un projet est en cours de r daction avec Hansjoerg Mixdorff sur l'application compar e du SFC et du MGI   une langue   tons (i.e. chinois).

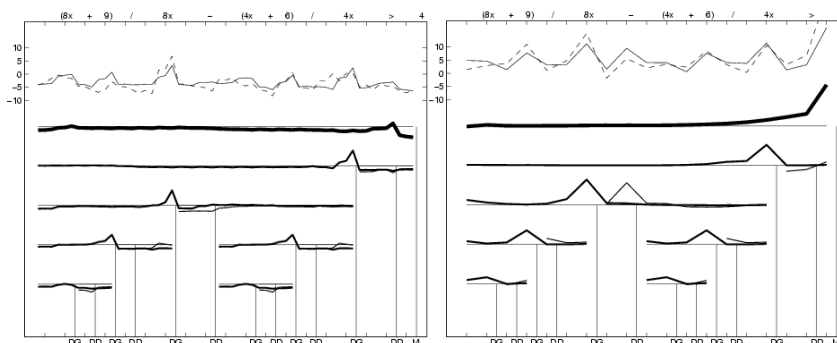


Figure 1 : R sultat de la d composition automatique de la prosodie d'une formule math matique  nonc e en allemand. A gauche, la m lodie et   droite, la courbe rythmique.

- Etude, mod lisation, synth se et  valuation du Langage Parl  Compl t  (LPC) : gr ce   de nombreuses sources de financement et une masse critique de chercheurs (th se de V. Attina et de G. Gibert, collaboration avec M.A. Cathiard de l' quipe Perception...), nous avons pris position dans le champ de l' tude du LPC. Gr ce   un dispositif exp rimental enti rement renouvel  par les contrats de l' quipe et notre collaboration dans le projet RNRT ARTUS avec Attitude Studio, nous avons collect  des donn es de capture de mouvement uniques et pr cieuses sur le LPC en action. L' tude d'une codeuse oraliste nous permet de b n ficier d'un cadre d'observation que nous ma trisons bien : le langage articul  et sa dimension audiovisuelle. La structure phonologique particuli re du LPC (typiquement Consonne-Voyelle) et le n cessaire contr le du mouvement de « nouveaux » segments de parole (i.e. le bras, la main et la t te) impose des contraintes de synchronisation avec la production vocale. Cet objet d' tude renouvelle radicalement les enjeux de la parole multimodale et donne de nouveaux th mes de recherche   l' quipe. Ainsi les mouvements de la t te, traditionnellement recrut s pour v hiculer la structuration discursive ou plus g n ralement des informations suprasegmentales, sont aussi utilis s pour porter le visage vers la main. Le LPC peut en fait  tre vu comme la r alisation de contractions main/visage en des lieux ( il, nez, joue, l vres, menton...) permettant de d sambigu ser la voyelle avec une cl  de doigts permettant de d sambigu ser la consonne. On retrouve ainsi des th matiques de contr le moteur, d' quifinalit  et de coarticulation, que nous avons d j  abord  en contr le moteur des articulateurs de la parole. Le LPC constitue aussi un d fi int ressant pour le

mod le SFC en offrant un nouveau jeu de param tres caract ristiques   g n rer (contr lant les mouvements de la t te) participant largement   l'encodage des structures segmentales et suprasegmentales.

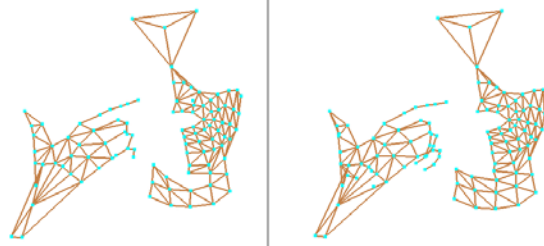
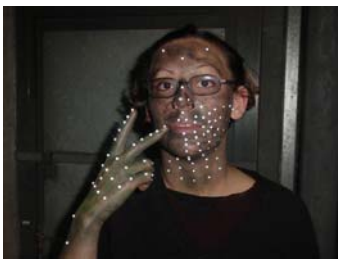


Figure 2 : *Capture du mouvement de la main et du visage d'une codeuse LPC.   gauche, une trame originale,   droite, sa reconstruction avec un mod le articul  par 29 param tres.*

L' tude et la caract risation du mouvement des segments du LPC (notamment celui de la main) repose aussi de mani re cruciale le probl me du mod le direct ou comment faire  merger des variables de contr le proximales (angle des doigts, mouvement du pouce, supination du poignet...) de l'observation du corps en mouvement. La Figure 2 illustre nos premiers r sultats de mod lisation non-lin aire. Ce probl me a fait l'objet du projet « V sale : Atlas articul  du visage et de la main » financ  par le BQR INPG et qui a permis de faire b n ficier Maxime B rar d'une bourse fl ch e pour une th se en co-encadrement entre trois laboratoires grenoblois (ICP, LIS et TIM-C).



**Figure 3 :** *Syst me de r alit  partag e avec interaction multimodale intuitive. L' cran est  quip  du syst me de suivi de regard de la soci t  Tobii. Le clone virtuel au centre de l' cran signale par des saccades oculaires qu'il conna t le centre d'int r t de l'utilisateur. D s que le temps de fixation sur un objet (r el ou virtuel) est sup rieur   300ms ou que l'utilisateur dit «  a ! », le clone donne oralement des renseignements sur ce dernier.*

## 2.2 Th mes de recherche prospectifs

### 2.2.1 Communication multimodale

Les th mes de recherche introduits plus haut suffisent largement   maintenir une activit  de recherche de niveau international. J'ai cependant initi  une r flexion au sein de l' quipe, du laboratoire et plus largement au sein du secteur expert d'ELESA sur les termes de l'interaction humaine avec un agent communicant.

En effet, si les th matiques « synth se de parole expressive » voire « synth se de parole  motionnelle » deviennent de plus en plus sensibles et mobilisent un nombre croissant de chercheurs, ces th mes de recherche ne sont souvent entrevues que comme une mani re d'accro tre la complexit  du substrat phon tique et des diverses fonctions de communication impliqu es dans une « vraie » interaction. Or cette interaction suppose (et supposait d j ... m me lorsqu'on se limite   la lecture de textes) un partenaire de communication... compl tement absent de l' change d'informations. Nos synth tiseurs parlent dans un espace vide et   destination d'inconnus... m me si les syst mes de dialogue homme-machine (ou plus r cemment personne-syst me) tente de g rer un espace de repr sentations communes (historique du dialogue, s mantique des actions...) mais comme si on connectait des entit s de traitement de l'information pures et d sincarn es...

Un des points de d part de cette r flexion a  t  nos efforts de mise en sc ne de clones parlants dans des applications de communication. Dans TempoValse, les usagers test s dans le laboratoire d'usage de France T l coms R&D ont tous soulign  l'importance du contact visuel : ainsi quand la cam ra et l' cran d'un terminal de visiophonie ne sont pas dans le m me axe de vue, les usagers ont  norm ment de g ne car l'objet d'attention du regard de l'autre semble hors de leur champ de vision.

L'autre source constante d'inspiration est le travail de l' quipe « Anthropologie Linguistique de la Parole » sur le suivi d'agents et le d veloppement du langage. Il y a appar t que les syst mes « O  » (Where) et « Qui » (Who) sont mis en

place de manière nettement plus précoce que le « Quoi » (What) et que ces systèmes exercent des fonctions largement préemptives sur notre analyse d'une scène multimodale. Avant donc de délivrer une information linguistique, il faut savoir où et à qui cette information doit être délivrée. Cette collaboration a débouché sur un texte commun « Biocybernétique de la communication parlée face à face » que nous avons présenté avec C. Abry au Pôle Européen des Universités de Grenoble. Au système ordonné WWW (Where, Who, What), j'ajouterais le « Comment » (How) cher aux paradigmes du « traitement affectif » (Affective Computing : <http://affect.media.mit.edu/>) et de « robotique incarnée » (Embodied robotics), tous deux mis en projet au MIT : *les signaux support de l'information linguistique sont aussi importants – sinon plus – que l'information elle-même*. En effet, le décodage correct et robuste de l'information nécessite un contrôle fin de multitudes de signaux susceptibles de maintenir l'attention partagée des interlocuteurs sur le discours. Ces signaux participent non seulement à l'organisation du discours, à l'engagement d'objets et d'agents dans celui-ci mais doivent contribuer aussi à donner des gages (au sens littéral d'engagement) d'attention et d'activité cognitive signifiant à l'autre que son message est compris et intéressant.

**Cette communication doit donc être ancrée sur le monde réel** : on parle de quelque chose à quelqu'un... dont le système de synthèse doit être conscient et doit se doter de capacités à construire des représentations. Pour ceci nous avons besoin d'engager des actions de recherche soit en propre soit en collaboration sur la caractérisation de l'environnement: vision par ordinateur (reconnaissance de locuteurs, identification d'objet...), traitement du signal (séparation parole/non parole, suivi locuteur, spatialisation du son ...). Nous avons pour ceci des collaborations avec plusieurs équipes proches :

- A l'ICP : « Perception » et « Anthropologie Linguistique de la Parole »
- Dans le bassin grenoblois : l'équipe « Prima » du laboratoire GRAVIR, les équipes « Images et Objets » du groupe GOTA et « Multimédia et fonctions associées » du groupe SIC du laboratoire LIS

Dans ce même cadre, j'ai lancé le projet « PRESENCE » qui sera présenté aux fédérations IMAG et ELESA à la rentrée. Ce projet consiste à développer des systèmes de réalité partagée où mondes physique et virtuel sont conscients l'un de l'autre et où les agents de communication respectifs (usagers, clones parlants) peuvent engager des conversations de manière intuitive engageant des objets physiques (objets de l'environnement, par ex. œuvres d'art dans un musée) ou virtuels (par ex. icônes). Une première maquette de type kiosque (intégrant un écran monté sur une tourelle mobile – afin d'assurer un face-à-face – et des microphones et des caméras – afin de combiner localisation acoustique et localisation visuelle ainsi que suivi du regard) est en cours de construction au sein de la salle MICAL, décrite ci-dessous. Elle sera rapidement couplée avec le système de deixis multimodale (cf. Figure 3) qui fut présenté en démo à sOc'2003.

### 2.2.2 La salle MICAL : les modèles en action

L'ICP est un laboratoire unique au niveau mondial rassemblant autant de compétences autour non pas d'une discipline mais d'un même objet de recherches éclairé par diverses disciplines. Le laboratoire est marqué d'une forte culture expérimentale : les théories sont confrontées aux données par l'intermédiaires de modèles, dont la théorie fixe la structure et que les données paramétrisent. Ces modèles sont donc souvent issus d'expériences *in vitro* ou *in vivo*, permettant de collecter un ensemble important de signaux (physiologiques, géométriques, acoustiques, aérodynamiques voire neurophysiologiques) sur un sujet exécutant une tâche donnée explicitement (par ex. par des consignes) ou implicitement (par ex. par la lecture d'un texte ou un rappel de souvenirs émotionnellement forts). Des expériences de perception viennent ensuite confirmer si le modèle issu des données possède bien les propriétés des données elle-même : d'acteur, le sujet devient alors spectateur attentif de ses actions ou de l'action des autres. Or ces deux phases expérimentales sont souvent déconnectées : la boucle perception-action est coupée : or il est certain que l'absence ou la présence de propriétés attendues dans un stimuli synthétique vont conditionner la réponse du sujet et ceci de manière dynamique. Pour explorer cette boucle d'interaction, il faut donc avoir la capacité d'implémenter des modèles temps-réel permettant :

- Soit de perturber la boucle perception-action (par ex. en manipulant des propriétés des stimuli produits par le sujet)
- Soit de simuler la boucle perception-action (par ex. par des expériences en Magicien d'Oz)
- Soit d'interagir directement avec un modèle complet (par ex. le système de deixis multimodale montré en Figure 3)

Nous avons pour ceci décidé en commun accord avec la direction de l'ICP de monter une salle d'expérimentation dédiée à l'interaction humaine sur le site GARE, dénommée salle MICAL<sup>1</sup>. Cette salle complète l'arsenal expérimental de l'ICP déjà doté d'une salle expérimental orientée « mesures » située sur le site CAMPUS.

---

<sup>1</sup> en l'honneur de l'abbé du même nom, qui en 1779 présenta à l'Académie Impériale des Sciences de Saint Pétersbourg deux têtes parlantes capables de prononcer un certain nombre de phrases. Placées sur un socle à l'intérieur d'un petit théâtre, un dialogue pouvait être engagé entre les deux têtes : "Le Roi donne la paix à l'Europe. La paix couronne le Roi de gloire. Et la paix fait le bonheur des peuples. O Roi adorable père de vos peuples, leur bonheur fait voir à l'Europe la gloire de votre trône." Un rapport de l'Académie des sciences et signé, entre autre, par Lavoisier et La Place, décrit en ces termes le mécanisme de création de la parole : "Les têtes recouvraient une boîte creuse, dont les différentes parties

Mandaté par le département, je travaille actuellement sur un projet de GIS intitulé « PEGASUS » visant à mettre en réseau les plate-formes expérimentales permettant d'étudier de nouvelles formes d'interaction humaine. MICAL sera la plate-forme de l'ICP.

### 3 Partie administrative

#### 3.1 Administration de la recherche

##### 3.1.1 Responsabilités locales

Directeur adjoint de la fédération ELESA depuis 2002

Membre élu du conseil de laboratoire de l'ICP depuis 1988

Responsable de l'équipe Synthèse depuis 1995 (devenue l'équipe Machines Parlantes en 1998)

##### 3.1.2 Responsabilités internationales

Secrétaire du Groupe Spécial d'Intérêt « Synthèse de la parole » (SynSIG) de l'ISCA depuis 2002

##### 3.1.3 Organisation de congrès

Co-organisateur de la Conférence sur les Objets Communicants (SOC) à Grenoble en 2003

Co-organisateur des Journées d'Études sur la Parole à Aussois en 2000

Organisation de séminaires et animation de groupes de travail multi-équipes: "Traitement parallèles distribués" avec J.L. Schwartz (1989-1992), Traitement de scènes audiovisuelles avec J. Luetin (2000)...

Relecteur du Journal of Acoustical Society of America, Journal of Phonetics, Speech Communication, IEEE Transactions on Speech and Audio Processing et Acta Acoustica

#### 3.2 Publications (2002-2003)

##### 3.2.1 Revues à comité de lecture

1. **Bailly, G., to appear.** Audiovisual speech synthesis. *International Journal of Speech Technology*
2. **Bailly, G. 2003.** "Close Shadowing Natural Versus Synthetic Speech." *International Journal of Speech Technology* 6(1): 11-19.
3. **Badin, P., Bailly G., Revéret L., Baciú 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(3): 533-553.

##### 3.2.2 Ouvrages édités

4. Vatikiotis-Bateson, E.; Bailly, G.; Perrier, P., In preparation. *Audiovisual speech processing*. Cambridge, MA, USA: MIT Press.

##### 3.2.3 Conférences invitées

1. **Bailly, G.; Holm, B., 2002.** Learning the hidden structure of speech: from communicative functions to prosody. In *Symposium on Prosody and Speech Processing*. Tokyo, Japan, 113-118.
2. **Bailly, G., 2002.** Audiovisual speech synthesis. From ground truth to models. In *International Conference on Speech and Language Processing*. Boulder - Colorado, 1453-1456.

##### 3.2.4 Conférences à comité de lecture

###### 3.2.4.1 Congrès internationaux

1. **Bérar, M.; Bailly, G.; Chabanas, M.; Elisei, F.; Odisio, M.; Pahan, Y., accepted.** Towards a generic talking head. In *6th International Seminar on Speech Production*. Sydney- Australia.
2. **Bailly, G.; Campbell, N.; Mobius, B., accepted.** ISCA Special Session: Hot Topics in Speech Synthesis. In *EuroSpeech*. Geneva, Switzerland.
3. **Odisio, M.; Bailly, G., accepted.** Shape and appearance models of talking faces for model-based tracking. In *Audio Visual Speech Processing*. St Jorioz, France.
4. **Bailly, G.; Elisei, F.; Odisio, M.; Pelé, D.; Grein-Cochard, K., 2003.** Objects and agents for MPEG-4 compliant scalable face-to-face telecommunication. In *Smart Object Conference*. Grenoble - France, 204-207.

---

étaient rattachées par des charnières et dans l'intérieur de laquelle l'auteur avait disposé des glottes artificielles de différentes formes sur des membranes tendues. L'air passant par ces glottes allait frapper les membranes qui rendaient des sons graves moyens ou aigus; et de leur combinaison résultait une espèce d'imitation très imparfaite de la voix humaine."

5. **Holm, B.; Bailly, G., 2002.** Learning the hidden structure of intonation: implementing various functions of prosody. In *Speech Prosody*. Aix-en-Provence, France, 399-402.
6. **Bailly, G.; Badin, P., 2002.** Seeing tongue movements from outside. In *International Conference on Speech and Language Processing*. Boulder - Colorado, 1913-1916.
7. **Bailly, G., 2002.** Audiovisual speech synthesis. From ground truth to models. In *International Conference on Speech and Language Processing*. Boulder - Colorado, 1453-1456.
8. **Bailly, G.; Gibert, G.; Odisio, M., 2002.** Evaluation of movement generation systems using the point-light technique. In *IEEE Workshop on Speech Synthesis*. Santa Monica, CA.

### 3.2.5 Chapitre dans les ouvrages

1. **Bailly, G.; Vatikiotis-Bateson, E.; Badin, P.; Revéret, L.; Yehia, H., In preparation.** Visible articulatory degrees of freedom of speech movements. In *Audiovisual speech processing*, E. Vatikiotis-Bateson; G. Bailly; P. Perrier, Editors. Cambridge, MA, USA: MIT Press.
2. **Bailly, G., 2002.** Towards more versatile signal generation systems. In *Improvements in Speech Synthesis*, E. Keller, et al., Editors. Chichester, England: J. Wiley & Sons, Ltd, 18-21.
3. **Bailly, G., 2002.** A parametric harmonic + noise model. In *Improvements in Speech Synthesis*, E. Keller, et al., Editors. Chichester, England: J. Wiley & Sons, Ltd, 22-38.
4. **Bailly, G., 2002.** The cost258 signal generation test array. In *Improvements in Speech Synthesis*, E. Keller, et al., Editors. Chichester, England: J. Wiley & Sons, Ltd, 39-51.

### 3.3 Vulgarisation

1. **FR3 Rhones-Alpes, 2003.** Reportage sur les démos sOc'2003 , 16 mai, 19:00-20:00.
2. **France Info, 2003.** L'intelligence ambiante , Paris, Chronique du 16 mai.
3. **01 net, 2003.** Les applications industrielles des objets communicants vont se multiplier, Paris, 23 Mai, 7:00.
4. **01 informatique, 2002.** *Les chercheurs truffent notre quotidien d'objets communicants* , Paris, 22 Novembre, p. 30-32.
5. **Le Monde, 2002.** *Des clones virtuels parlants pour mieux comprendre le langage* , Paris, 1<sup>er</sup> Mars 2002, p. 23.
6. **Pour La Science, 2002.** *Le vrai visage de la parole* , 291, Paris, Janvier, p. 14.

### 3.4 Enseignements

1. Cours d'option « Images et réalité virtuelle » de 3<sup>me</sup> année ENSIMAG depuis 2001 « Parole et langage: systèmes d'interaction verbale »
2. Cours de DEA Sciences Cognitives depuis 1998 « Synthèse de la parole »

### 3.5 Responsabilités de contrats et subventions

1. Responsable scientifique du partenaire ICP dans le projet RNRT Technolangue EvaSy « Evaluation des systèmes de synthèse de parole » depuis 2002
2. Responsable scientifique du partenaire ICP dans le projet RNRT ARTUS « Animation réaliste par tatouage audiovisuel à l'usage des sourds » depuis 2002
3. Responsable scientifique du partenaire ICP dans le projet ROBEA HR+ « Pour une interaction homme-robot autonome » depuis 2002
4. Responsable scientifique du partenaire ICP dans le projet RNRT TEMPO-VALSE « Visiophonie & animation labiale scalable » de 2000 à 2002

### 3.6 Encadrement

#### 3.6.1 Thèses

1. Co-directeur de thèse INPG de M. Bézar (depuis 2003)..... avec M. Desvignes (LIS) & Y. Payan (TIM-C) Modèles statistiques de la variabilité d'organes articulés du corps humain Thèse de troisième cycle, Spécialité Signal, Image, Parole & Télécommunications, Institut National Polytechnique de Grenoble, Grenoble - France.
2. Directeur de thèse INPG de G. Gibert (depuis 2002) Synthèse multimodale de la parole Thèse de troisième cycle, Spécialité Signal, Image, Parole & Télécommunications, Institut National Polytechnique de Grenoble, Grenoble - France.
3. Co-directeur de thèse INPG de P. Gacon (depuis 2002)..... avec P.Y. Coulon (LIS) Analyse d'images et modèles de formes pour la détection et la reconnaissance. Application aux visages. Thèse de troisième cycle, Spécialité Signal, Image, Parole & Télécommunications, Institut National Polytechnique de Grenoble, Grenoble - France.

4. Directeur de th se INPG de M. Odisio (depuis 2000)  
Analyse et synth se de sc nes audiovisuelles  
Th se de troisi me cycle, Sp cialit  Signal, Image, Parole & T l communications, Institut National Polytechnique de Grenoble, Grenoble - France.
5. Directeur de th se INPG de B. Holm (1998-2003)  
La prosodie de l' nonciation de formules math matiques  
Th se de troisi me cycle, Sp cialit  Sciences Cognitives, Institut National Polytechnique de Grenoble, Grenoble – France, 2003.

### **3.6.2 DEAs**

#### ***DEA Sciences Cognitives :***

1. M. B rar (2003)

#### ***3.6.2.1 DEA Signal-Image-Parole :***

1. G. Gibert (2002)
2. M. B rar (2002)

### **3.6.3 Autres stages**

#### ***3.6.3.1.1 Stages ing nieur***

1. C. Izoard, C. Renoux, I. Ghoza l , 2me ann e D partement T l coms (2003)
2. N. Lechat, 2me ann e ENSERG (2002)

### **3.7 Participations   des jurys de th ses**

1. B. Holm, Grenoble (juin 2003)
2. G. Breton, Rennes (octobre 2002)