

JEAN-LUC GAUVAIN

DÉCODER

LA PAROLE

Imaginez que vous écoutiez un discours en espagnol et qu'une machine vous le traduise en direct en français ou en arabe!

Ou que vous puissiez trouver dans un enregistrement d'émission de radio le passage exact où l'on parle de la Nouvelle-Zélande, sur laquelle vous recherchez des renseignements! Grâce aux recherches de Jean-Luc Gauvain, informaticien au Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (LIMSI) d'Orsay, ce seront peut-être très bientôt des applications réelles. Il faut dire que le lauréat de la médaille d'argent du CNRS s'y connaît « un peu » en informatique, puisqu'il a construit lui-même un ordinateur au cours de ses études. « Ce qui me plaisait, c'était de le concevoir, pas tant de m'en servir! » Et l'excellence de ses résultats se mesure à la très bonne position du système de reconnaissance de la parole dont il est l'architecte, système qui est l'un des meilleurs au niveau mondial.

« PARTICIPER AUX CAMPAGNES D'ÉVALUATION, C'EST UN PEU COMME PRENDRE LE TGV AU LIEU DE LE REGARDER PASSER. »

Après une maîtrise de sciences et techniques en télécommunications, Jean-Luc Gauvain suit un DEA d'électronique pendant lequel il se spécialise dans le traitement de l'information. C'est au cours de son service militaire, effectué comme scientifique du contingent au ministère de la Défense, qu'il vient visiter le LIMSI. La très bonne ambiance qu'il y trouve le convainc de le choisir comme laboratoire d'accueil pour sa thèse. Il s'attelle d'ores et déjà au développement d'algorithmes pour la reconnaissance de la parole.

Alors qu'il est recruté au CNRS en 1983, notre informaticien intègre tout naturellement le LIMSI.

Il s'intéresse à la plupart des problèmes du traitement automatique de la parole, développant des modèles probabilistes et des algorithmes toujours plus complexes, mais avec le souci constant de transfert et de valorisation. En 1990, il obtient une mise à disposition d'un an et demi et rejoint *AT&T Bell Laboratories*, aux États-Unis. L'occasion pour lui de réorienter ses travaux de recherche en mettant l'accent sur les aspects théoriques. Il s'intéresse en particulier à l'apprentissage adaptatif de modèles probabilistes, méthode utilisée aujourd'hui par la plupart des équipes du domaine.

Mais ce que Jean-Luc Gauvain rapporte surtout des États-Unis, c'est la conviction de l'importance des campagnes d'évaluation pour les technologies de la langue.

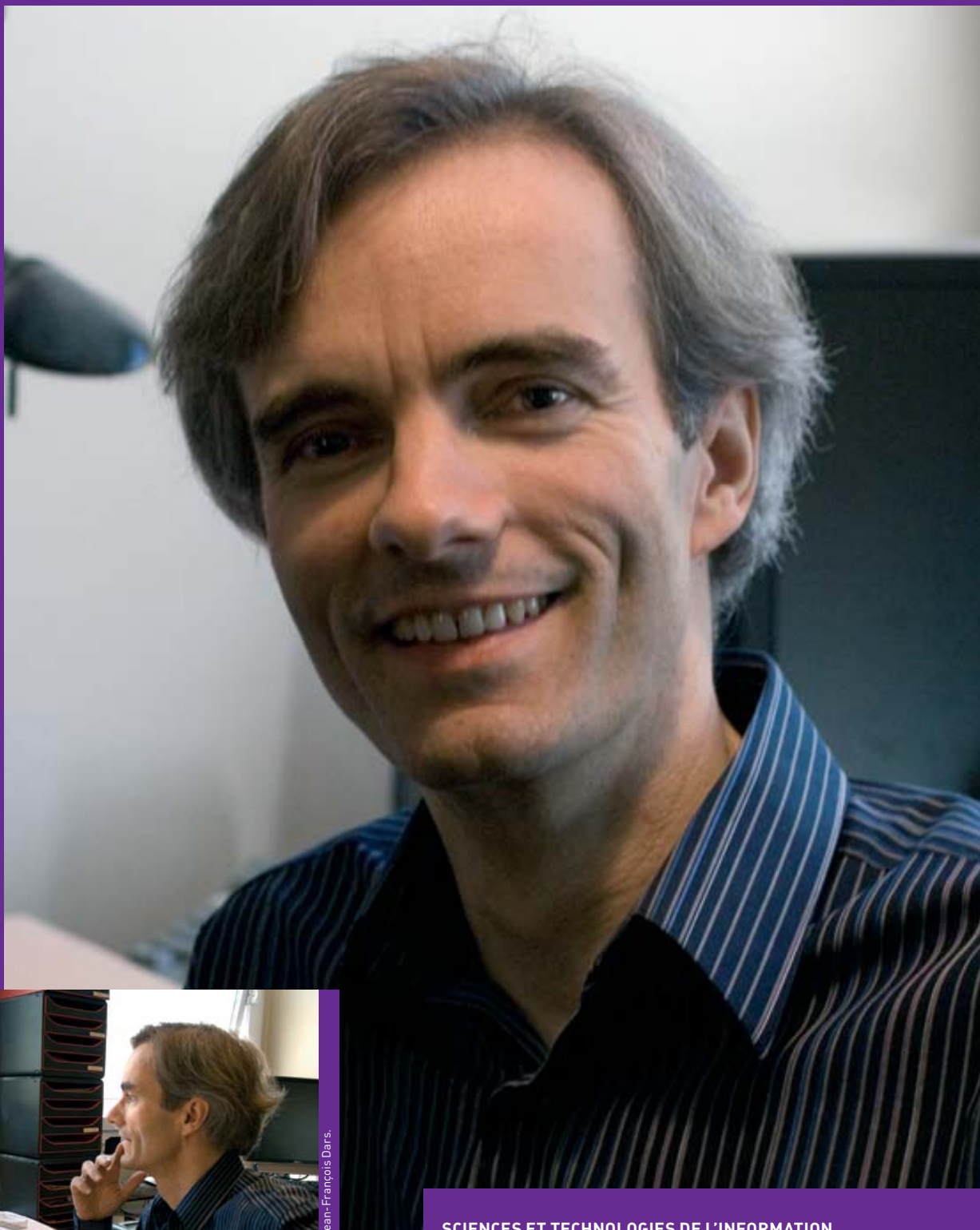
« C'est vraiment très efficace. Mais l'idée a parfois du mal à passer en Europe où les projets de recherche sont plus constitués sur la base de la complémentarité. Participer aux campagnes d'évaluation, c'est un peu comme prendre le TGV au lieu de le regarder passer. » Ce système, souligne Jean-Luc Gauvain, accroît grandement les échanges d'information et les progrès, en permettant une évaluation objective et contrastive de modèles et méthodes difficiles à évaluer avec les seuls moyens d'un laboratoire. « En 1992, lors de notre première participation à une campagne organisée par le NIST (*National Institute of Standards and Technology*), nous avons été classés premiers! » L'histoire s'est répétée depuis.

L'emploi de la première personne du pluriel caractérise l'esprit d'équipe de Jean-Luc.

« En travaillant sur le traitement automatique du langage parlé, nous devons faire appel aux compétences de nombreuses disciplines, le traitement du signal, l'acoustique, la phonétique, la linguistique... et donc être capables de fédérer nos complémentarités. » Le groupe « Traitement du langage parlé » que Jean-Luc dirige s'intéresse à l'ensemble des problèmes du traitement de la parole, qu'il s'agisse de la transcrire ou de la traduire, d'identifier les locuteurs et leurs émotions, ou de dialoguer avec l'ordinateur. « Ces recherches s'appuient sur de grands corpus oraux constitués de milliers d'heures de parole qui servent à l'étude et au développement de modèles statistiques. »

LE GROUPE « TRAITEMENT DU LANGAGE PARLÉ » QU'IL DIRIGE S'INTÉRESSE À L'ENSEMBLE DES PROBLÈMES DU TRAITEMENT DE LA PAROLE, QU'IL S'AGISSE DE LA TRANSCRIRE OU DE LA TRADUIRE, D'IDENTIFIER LES LOCUTEURS ET LEURS ÉMOTIONS, OU DE DIALOGUER AVEC L'ORDINATEUR.

Au passage, le chercheur signale que les systèmes de reconnaissance de la parole, à l'inverse de ce que peut prédire la théorie, fonctionnent mieux lorsque le locuteur est une locutrice. « Peut-être les femmes ont-elles un discours plus cohérent? »



© CNRS Photothèque - Jean-François Daris.



© CNRS Photothèque - Jean-François Daris.

**SCIENCES ET TECHNOLOGIES DE L'INFORMATION
ET DE L'INGÉNIERIE (ST2I)**

LABORATOIRE D'INFORMATIQUE POUR LA MÉCANIQUE
ET LES SCIENCES DE L'INGÉNIEUR (LIMSI)

CNRS
ORSAY

<http://www.limsi.fr/Scientifique/tlp/>

Plus récemment, notre médaillé, par ailleurs co-éditeur en chef du journal *Speech Communication* (Elsevier), s'intéresse aussi au problème de l'indexation par le contenu de documents audiovisuels. « Ce type de recherche nécessite des moyens matériels et humains importants, c'est la raison pour laquelle nombre de

mes activités de recherche sont liées à des ressources contractuelles offrant les moyens et un cadre applicatif propices à la valorisation de ces activités. » Les résultats de ces recherches devraient grandement faciliter l'accès à l'information dans des documents multimédias et multilingues.