

Workshop MIDL 2004

Identification des langues et des variétés dialectales par les humains et par les machines

— Appel à communications 2004 —

Organisation : par le consortium du projet MIDL (Modélisations pour l'IDentification des Langues) dans le cadre du programme interdisciplinaire STIC-SHS du CNRS : **LIMSI-CNRS, ILPGA/LPP Paris3, ENST et DGA,**

avec le soutien de l'AFCP.

Lieu et date : Paris, 29 - 30 novembre 2004

Présentation du colloque :

Ce colloque a comme ambition de rassembler des chercheurs de disciplines souvent éloignées autour d'un même objet d'étude : la discrimination et l'identification à partir du signal de parole des langues et de leurs variétés, soit par l'humain (nouveau-né, nourrisson, enfant et adulte), soit par la machine. Des éclairages de spécialités différentes permettront de mieux répondre aux questions suivantes (liste non exhaustive) :

- Sur quels indices acoustiques et autres connaissances non acoustiques se fonder pour identifier une langue ou une variété dialectale ?
- Quels sont les indices véhiculés par le signal acoustique et/ou perçus dans ce signal ?
- Quelles sont les connaissances innées et/ou acquises dès le plus jeune âge ?
- Comment identifier une langue forte et une langue faible chez les bilingues ?
- Parmi les indices utilisés par l'homme, quels sont ceux potentiellement utilisables par la machine ?
- Comment modéliser le signal acoustique et d'autres connaissances sur les langues en vue d'une identification automatique ?
- Quels sont les rapports entre perception et traitement automatique, y a-t-il des mécanismes communs ? Peuvent-ils s'éclairer mutuellement ?
- Le traitement automatique permet-il de valider ou éventuellement de remettre en question des connaissances partiellement établies ?
- Quelle est la part des indices segmentaux (voyelles, consonnes, coarticulation) et des indices suprasegmentaux (tons, rythme, intonation, etc.) ?

- Existe-t-il des techniques vocales différentes, des clichés mélodiques ou des habitudes articulatoires particulières permettant d'identifier langues et accents ?
- Comment mesurer les performances en identification/discrimination et en regroupement/ confusion des langues et de leurs variétés dialectales ?
- Le débit de parole et le style (spontané/lu, par exemple) influencent-ils les résultats ?

Thèmes du colloque :

La liste des thèmes ci-dessous est plus indicative qu'exhaustive. Ces thèmes sont à développer de préférence avec une perspective « identification des langues ».

- Langues et variétés de langue (dialectales, sociolinguistiques, ...)
- Humain et machine
- Typologie des langues et de leurs variétés
- Phonétique, phonologie, prosodie
- Production, perception
- Identification et qualité de voix
- Corpus et applications
- Acquisition et apprentissage : bébés, L2
- Cognition du bilinguisme

Langues du colloque : français et anglais

Calendrier des soumissions :

- 15 mai : soumission d'un abstract de 1000 mots
- 30 juin : notification d'acceptation
- 1^{er} novembre : soumission du papier final (4-6 pages)
- 29 - 30 novembre : déroulement du colloque

Soumission des résumés :

Les auteurs devront envoyer leur soumission sous forme d'un document attaché à un courrier électronique ayant pour titre « soumission MIDL 2004 » et contenant le titre de la communication, le nom, l'affiliation, l'adresse postale, l'adresse électronique, le numéro de téléphone et le fax de l'auteur principal. Les formats acceptés sont .txt, .rtf. ou .pdf.

Les résumés devront parvenir au comité d'organisation avant le 15 mai, sous forme électronique, à l'adresse <midl2004@limsi.fr>. Ils seront examinés par le comité de lecture.

Soumission des communications :

Les articles préalablement sélectionnés ne devront pas dépasser 6 pages en Times 10, sur deux colonnes, format A4.

Les détails sur le format définitif seront disponibles sur la page WEB du colloque.

Comité scientifique :

Martine Adda-Decker, LIMSI, Orsay
Régine André-Obrecht, IRIT, Toulouse
Fabien Antoine, CTA-DGA, Arcueil
Philippe Boula de Mareüil, LIMSI, Orsay
Maria Candea, Paris3 EA1483, Paris
René Carré, ENST, Paris
Christophe d'Alessandro, LIMSI, Orsay
Albert di Cristo, LPL, Aix-en-Provence
Edouard Geoffrois, CTA-DGA, Arcueil
Pierre Hallé, Paris3, Paris
Lori Lamel, LIMSI, Orsay
Jean-Sylvain Liénard, LIMSI, Orsay
Philippe Martin, PARIS VII, Paris
Christophe Pallier, INSERM U562, Orsay
François Pellegrino, DDL, Lyon
Franck Ramus, LSCP-ENS, Paris
Jacqueline Vaissière, ILPGA/LPP Paris3, Paris
Nathalie Vallée, ICP, Grenoble
Ioana Vasilescu, ENST, Paris

Comité d'organisation :

Martine Adda-Decker, LIMSI
Fabien Antoine, CTA-DGA
Philippe Boula de Mareüil, LIMSI
Maria Candea, Paris3 EA1483
Cedric Gendrot, ILPGA/LPP Paris3
Jean-Sylvain Liénard, LIMSI
Jacqueline Vaissière, ILPGA/LPP Paris3
Ioana Vasilescu, ENST

Contacts :

Martine Adda-Decker, madda@limsi.fr
Philippe Boula de Mareüil, mareuil@limsi.fr
Ioana Vasilescu, vasilesc@tsi.enst.fr.

Le colloque aura lieu au

Carré des Sciences (Ministère de la Recherche)

1, rue Descartes- 75005 Paris

et les communications seront présentées lors d'une séance plénière. Des sessions posters peuvent être rajoutées suivant le souhait des participants. Les articles seront publiés par les presses de l'ENST et un CD-ROM avec les articles complets sera distribué aux participants.

