

Les 16, 17 et 18 novembre 2009
A l'Université d'Avignon et des Pays de Vaucluse

8^{ème} CONFÉRENCE

RJCP 2009

Rencontres Jeunes Chercheurs en Parole

Organisée par de jeunes chercheurs
pour de jeunes chercheurs

<http://rjcp2009.univ-avignon.fr>

Programme et recueil des résumés



RPM2



Les Incubateurs
en Provence Alpes Côte d'Azur
du projet à la création d'entreprise innovante





La huitième édition des **Rencontres Jeunes Chercheurs en Parole**

Avignon, 16-18 Novembre 2009

Cette manifestation est organisée par les associations NP-LIA et AFCP ainsi que le LIA.



L'Association des Non-Permanents du Laboratoire Informatique d'Avignon,
Le Laboratoire Informatique d'Avignon, l'Association Francophone de la Communication Parlée
et l'Université d'Avignon et des Pays de Vaucluse.

Le comité scientifique est composé de :

Georges Linarès, président
Nicolas Audibert
Odile Bagou
Frédéric Béchet
Jean-François Bonastre
Nathalie Camelin
Eric Charton

Richard Dufour
Anika Falkert
Benoit Favre
Corinne Fredouille
Fabrice Lefèvre
Driss Matrouf
Marie-Jean Meurs

Vladimir Popescu
Albert Rillard
Solange Rossato
Isabelle Rousset
Christophe Servan
Pascal Vaillant

Le comité d'organisation est composé de :

Présidente Claire Petiteau
Vice-président Raphaël Rubino
Resp. Logistique Pierre Gotab
Grégory Senay
Caroline Buffoni
Essaid Sabir
Resp. Trésorerie Juliette Kahn
Florian Verdet
Rodrigo Acuna
Resp. Edition Electronique Stanislas Oger

Correspondants Scientifiques Christophe Servan
Rémi Lavalley
Resp. Communication Johanne Tremblay
Bessam Fallah
Frédéric Duvert
Resp. Partenaires Industriels Khalil Ibrahimi
Florian Pinault
Daniela Maftuleac
Resp. Web Mickaël Rouvier
Carlos Diego Rodrigues

ISBN 978-2-9534233-1-0



PROGRAMME

Lundi 16/11	Mardi 17/11	Mercredi 18/11
09h00	09h00	09h30
Accueil	Accueil	Accueil
Hall d'entrée	Hall d'entrée	Hall d'entrée
09h45	09h45	10h15
Discours d'introduction Philippe Michelon	O3. Phonétique 112, 222, 232, 242, 92	Conférence académique Solange Rossato
AT05	1w08	1w08
10h15	12h05	11h45
Pause	Déjeuner	Déjeuner
Hall d'entrée	Restaurant	Restaurant
10h30	13h20	13h40
Conférence conjointe Yves Jeanneret	O4. TAP 102, 172, 212, 272, 62	O5. Perception 122, 182, 202, 252
AT05	1w08	1w08
11h30	15h05	15h25
Déjeuner	Pause	Remise des Best Papers
Restaurant	Hall d'entrée	Discours de clôture
13h30	15h25	Hall d'entrée
O1. Production 132, 152, 72, 162	Forum	
0w19	0w16	
15h15	17h15	
Pause	Conférence conjointe Frédéric Beaugendre	
Hall d'entrée	AT05	
15h35	18h15	
O2. Pathologies 262, 282, 82	Table ronde	
0w19	0w16	
17h30	20h00	
Noms complets des sessions :	Diner de gala	
01 Production	Tapalocas	
02 Pathologies		
03 Phonétique		
04 Traitement automatique de la parole		
05 Perception		

CONFÉRENCES INVITÉES

Conférence conjointe MajecSTIC et RJCP

Lundi 16/11 de 10h30 à 11h30 Amphithéâtre AT05

Yves JEANNERET

Laboratoire Culture et Communication, Université d'Avignon et des Pays de Vaucluse

“Sciences de la communication, sciences de l’information et de la communication, sciences et technologies de l’information et de la communication : un sens peut en cacher un autre”

voir page iii

Conférence conjointe MajecSTIC et RJCP

Mardi 17/11 de 17h15 à 18h15 Amphithéâtre AT05

Frédéric BEAUGENDRE

Voice-Insight, Bruxelles

“Perspectives industrielles dans le domaine des technologies vocales”

voir page iv

Conférence académique RJCP

Mercredi 18/11 de 10h15 à 11h45 Salle 1w08

Solange ROSSATO

Laboratoire d'Informatique de Grenoble

“Phonétique acoustique et articulatoire : des résonances aux coefficients utilisés en TA”

voir page v

Conférence conjointe MajecSTIC et RJCP

Lundi 16/11 de 10h30 à 11h30 Amphithéâtre AT05

Président : Philippe Michelon

Sciences de la communication, sciences de l'information et de la communication,
sciences et technologies de l'information et de la communication : un sens peut en
cacher un autre

Yves JEANNERET

Laboratoire Culture et Communication, Université d'Avignon et des Pays de Vaucluse

Biographie

Professeur à l'Université d'Avignon et des Pays de Vaucluse, membre du Laboratoire Culture et Communication et directeur de recherche dans le programme de doctorat international Muséologie, médiation, patrimoine, Yves Jeanneret mène des recherches sur la circulation des savoirs, l'écriture, les transformations médiatiques et l'épistémologie de sciences sociales. Il est aussi membre du Collège des Etudes Doctorales et membre du Conseil Scientifique de l'Université d'Avignon et des Pays de Vaucluse. Yves Jeanneret s'intéresse aux thématiques de recherche suivantes :

- Analyse de la circulation des savoirs et des objets culturels dans la société : médiation, divulgation, communication scientifique et technique, médiatisation de la littérature, muséologie.
- Analyse des médiations écrites : sémiotique de l'écriture, modalités de circulation et de transformation des écrits, rôle de l'énonciation écrite, des supports de l'écriture et des pouvoirs éditoriaux dans la culture.
- Analyse des transformations médiatiques : étude du rapport entre médiatisation technique et médiation symbolique, étude des médias informatisés et du texte de réseau, analyse des usages et de leurs normes
- Épistémologie des sciences sociales : analyse de l'activité d'écriture et de textualisation dans la recherche, de la circulation des savoirs en sciences sociales, étude des critères et ressorts de l'analyse communicationnelle, étude des rapports entre sémiotique et communication.

Résumé

La communication, fondée sur une expérience subjective de participation à l'animation de recherche dans une discipline et à la reconnaissance d'un champ disciplinaire complexe au CNRS et à l'ANR, interrogera les relations qui peuvent s'établir entre constructions institutionnelles, schèmes d'intelligibilité et enjeux des recherches dans un domaine que certains définissent comme un secteur, d'autres un champ d'analyse et d'autres un objet théorique. De même que certains ne voyaient pas la nécessité de créer une sociologie à la fin du 19ème siècle parce que tout le monde peut penser le social, beaucoup estiment aujourd'hui que la communication ne mérite pas d'être instituée en discipline. Pour discuter cette question, il faut se débattre entre l'institution sociale, l'institution cognitive et la poétique des savoirs. On ne fera ici qu'esquisser cette réflexion ambitieuse et pourtant essentielle, dans une société qui revendique en permanence la communication sans réellement la prendre au sérieux comme objet conceptuel. Plus personne ne pense qu'observer la chute des corps suffit pour être physicien. Quand en viendra-t-on là pour les faits de communication dans la société ?

Conférence conjointe MajecSTIC et RJCP

Mardi 17/11 de 17h15 à 18h15 Amphithéâtre AT05

Président : Georges Linarès

Perspectives industrielles dans le domaine des technologies vocales

Frédéric BEAUGENDRE

Voice-Insight, Bruxelles

Biographie

Frédéric Beaugendre est spécialiste dans le domaine des technologies vocales depuis le début des années 1990. Après avoir réalisé une thèse de doctorat et un post-doctorat en synthèse vocale dans le milieu académique, il a pu exercer depuis 1997 les fonctions de chercheur et chef de projets au sein de trois entreprises spécialisées dans les différentes disciplines du domaine vocal et du traitement du langage naturel. Cette expérience industrielle a été acquise au sein d'entreprises développant les technologies de base (reconnaissance et synthèse vocale notamment) aussi bien que d'intégrateurs de ces technologies pour des applications industrielles et grand-public. Frédéric Beaugendre s'intéresse aux thématiques de recherche suivantes :

- Reconnaissance vocale : front-end, robustesse, adaptation, modélisation acoustique, langue tonales, modèles de langage
- Synthèse vocale : Synthèse par concaténation, par HMMs, prosodie, perception, évaluation

Résumé

La marché des technologies vocales se divise aujourd'hui en deux segments principaux que sont les applications dans le domaine de la téléphonie d'une part, et les applications multimedia, industrielles et embarquées d'autre part. Après un fourmillement d'activités dans le domaine pendant des années 1990, les analyses financières promettant une croissance importante des revenus à court terme sur ce marché, le nombre d'entreprises fournisseur de cette technologie a substantiellement diminué depuis le début des années 2000. Aujourd'hui, peu d'acteurs subsistent, même si paradoxalement la technologie n'a jamais atteint un tel niveau de maturité. Dans ce contexte, l'objectif de cet exposé est de faire un survol des technologies vocales et leur évolution récente, et d'en présenter les principales perspectives industrielles associées à court et moyen terme.

Conférence académique RJCP

Mercredi 18/11 de 10h15 à 11h45 Salle 1w08

Président : Georges Linarès

Phonétique acoustique et articulatoire : des résonnances aux coefficients utilisés en TA

Solange ROSSATO

Laboratoire d'Informatique de Grenoble

Biographie

Solange Rossato est Maître de Conférence en Sciences du Langage à l'Université Stendhal Grenoble III. Elle effectue sa recherche au Laboratoire d'Informatique de Grenoble, au sein de l'équipe GETALP ainsi qu'à GIPSA-Lab. Ses thématiques de recherche se situent dans le cadre de la phonétique acoustique et de l'interface entre phonétique et phonologie. Elle a travaillé plus spécifiquement sur la nasalisation et l'émergence des voyelles nasales ainsi que sur certaines caractéristiques spécifiques de la voix (affects, reconnaissance du locuteur...) en étudiant aussi bien les processus humains que le traitement automatique de la parole.

Résumé

Cette présentation s'attache à décrire les liens simples entre forme du conduit vocal et résonnances, en reliant les formants des voyelles aux cavités. L'interprétation des formants en termes de résonnances de cavités permet d'expliquer les effets acoustiques de l'arrondissement ou la différence des valeurs formantiques entre hommes, femmes et jeunes enfants. Ces modélisations facilitent la lecture de spectrogramme en français. La deuxième partie s'intéresse aux coefficients utilisés en Traitement Automatique tels que les LPCC, MFCC, PLP et aux théories phonétiques qui y sont rattachées.

Une partie pourra se faire sous forme de TD.

INDEX DES SESSIONS

- **Lundi 16/11 de 13h30 à 15h15**

- Session Orale - O1 - *Production* 1
Président : *Nicolas Audibert*

- **Lundi 16/11 de 15h35 à 17h30**

- Session Orale - O2 - *Pathologies* 2
Président : *Gilles Pouchoulin*

- **Mardi 17/11 de 9h45 à 12h05**

- Session Orale - O3 - *Phonétique* 2
Présidente : *Christine Meunier*

- **Mardi 17/11 de 13h20 à 15h05**

- Session Orale - O4 - *Traitement automatique de la parole* 3
Président : *Yannick Estève*

- **Mercredi 18/11 de 13h40 à 15h25**

- Session Orale - O5 - *Perception* 4
Présidente : *Corinne Fredouille*
-

PROGRAMME DÉTAILLÉ

LUNDI 16/11 DE 13H30 À 15H15

• Session Orale - O1

Production

Lundi 16/11 de 13h30 à 15h15, 0w19

132 : Etude articulatoire du mouvement d'étirement et d'ouverture des lèvres lors d'émotions et une attitude simulées.

Laurianne Georgeton

– ilpga

Résumé

Dans cette étude, nous nous sommes intéressée aux variations articulatoires des lèvres (étirement et ouverture aux lèvres) pour quatre voyelles du français /a/, /i/, /u/ et /y/ lors d'émotions (anxiété, dégoût, colère, joie et tristesse) et une attitude (tendresse) simulée. Nous avons utilisé des marqueurs placés autour des lèvres. Le mouvement des lèvres a été étudié grâce au Qualisys qui utilisent un système de caméra infra-rouge. Les coordonnées en 3D des marqueurs sont ensuite reconstruites puis étudiées. Cette étude a montré que les contrastes intrinsèques (donc attendus) des voyelles ne sont pas observables lors de la parole normale. Le geste d'étirement est maximisé lors de la réalisation de la joie et de la tendresse. Le geste d'ouverture aux lèvres est maximisé lors de la colère, du dégoût et de l'anxiété.

152 : Représentations cérébrales des articulateurs de la parole

Krystyna Grabski, Marc Sato, Jean-Luc Schwartz, Laurent Lamalle, Coriandre Vilain

– Gipsa-Lab
– INSERM

Résumé

In order to localize cerebral regions involved in articulatory control processes, ten subjects were examined using functional magnetic resonance imaging while executing lip, tongue and jaw movements. Although the three motor tasks activated a set of common brain areas classically involved in motor control, distinct movement representation sites were found in the motor cortex. These results support and extend previous brain imaging studies by demonstrating a sequential dorsoventral somatotopic organization of lips, jaw and tongue in the motor cortex.

72 : Utilisation d'une grille polaire adaptative pour la construction d'un modèle articulatoire de la langue

Julie Busset

– LORIA

Résumé

The construction of articulatory models from medical images of the vocal tract, especially X-ray images, relies on the application of an articulatory grid before deriving deformation modes via some factor analysis method. One difficulty faced with the classical semi-polar grid is that some tongue contours do not intersect the grid what gives rise to incomplete input vectors, and consequently poor tongue modeling in the front part of the mouth cavity which plays an important role in the articulation of many consonants. First, this paper describes preparation of data, i.e. drawing or tracking articulator contours, compensation of head movements and the construction of the adaptive polar grid. Then, the results of the principal component analysis are presented and compared with those obtained with the semi-polar grid.

162 : Peut-on utiliser la voix chantée pour améliorer la correction phonétique segmentale en langue étrangère ?

Sandra Cornaz, Nathalie Henrich, Antonio Romano, Nathalie Vallée

– GIPSA-Lab, DPC - Département Parole et Cognition - UMR 5216 CNRS/Université de Grenoble
– LPEAG, Laboratorio di Fonetica Sperimentale "Arturo Genre" di Torino

Résumé

Music may have a positive impact on learning processes. In linguistics, the positive role of music on perception of prosodic features has been pointed out, and a recent study shows that the segmentation of words in a foreign language would be facilitated by sung. In the present study, we aim at investigating whether singing-voice tasks could help to improve the learning of French phonemes. For comparison purpose, a traditional phonetic method was slightly modified to introduce singing-voice tasks. Native speakers of Italian were divided into two groups : one for the common phonetic teaching, and one for the phonetic teaching including singing-voice tasks. The results show that the subjects who were taught with additional singing-voice tasks learn faster than the others, produce better than the other one the anterior phonemes /y/ and /ø/ in the acoustical regions expected for these vowels in French, and the overlap of acoustic scatterings is less important.

LUNDI 16/11 DE 15H35 À 17H30

• Session Orale - O2

Pathologies

Lundi 16/11 de 15h35 à 17h30, 0w19

262 : Étude descriptive préliminaire de la voix de l'enfant implanté cochléaire à partir des mesures aérodynamiques

Harold Andrés Guerrero Lopez, Benoit Amy De La Breteque, Michel Mondain, Patrick Serrafero, Catherine Trotter, Melissa Barkat-Defradas

- CHU Gui de Chauliac, Montpellier
- Ecole Centrale de Lyon
- Praxiling UMR 5267 CNRS - Montpellier III
- UMR I3M - Université Montpellier II

Résumé

The purpose of this study was to describe the voice physiological characteristics of cochlear implanted children by voice aerodynamic measurements. Subjects were 6 girls and 14 boys prelingual or congenital profound deaf children. Voice aerodynamic measurements were obtained from the children by EVA'2 system : estimated subglottic pressure (PSGE), oral airflow mean (DAB), intensity, glottal efficiency, laryngeal efficiency. Although our findings are descriptive and not have been compared to other populations at the present study (such as children with normal hearing and hearing aids), we can assume that cochlear implanted children's voice physiological behavior is similar to the phonatory behavior of children with normal hearing.

282 : Méthodes objectives issues du traitement automatique de la parole pour la recherche de zones 'déviants' dans la parole dysarthrique

Pierre Clement, Corinne Fredouille

- Laboratoire Informatique d'Avignon

Résumé

Une déficience ou un dysfonctionnement d'une enzyme présente dans les lysosomes est à l'origine des maladies de surcharge lysosomale (ou maladies lysosomales). Parmi les nombreux symptômes pouvant être liés à ces maladies, les patients peuvent être atteints de dysarthrie. La dysarthrie se définit par un trouble de l'élocution dû à une lésion du système nerveux. A l'heure actuelle, l'évaluation du degré de sévérité de la dysarthrie se fait de façon perceptive par les cliniciens. Bien qu'il existe des critères perceptuels et visuels définis sur lesquels les cliniciens peuvent s'appuyer pour évaluer la dysarthrie, cette évaluation reste très dépendante du clinicien l'effectuant, et revêt par conséquent un caractère très subjectif. Pour cette raison, la mise en place de méthodes plus objectives de l'évaluation de la dysarthrie devient une nécessité. Cette mise en place doit reposer au préalable sur une meilleure connaissance et compréhension

des phénomènes acoustico-phonétiques liés à la parole dysarthrique. Cet article décrit les méthodologies objectives mise en place afin de rechercher des zones 'déviants' dans la parole dysarthrique. Cette analyse de la parole dysarthrique sera effectuée grâce à des outils de traitement automatique de la parole.

82 : L'analyse prédictive des données orales des sujets Alzheimer et des patients MCI peut-elle contribuer au diagnostic précoce de la maladie ?

Hye Ran Lee, Melissa Barkat-Defradas

- Laboratoire Praxiling, UMR5267-CNRS/ Université Montpellier 3

Résumé

L'objet de nos travaux est de déterminer – à travers l'analyse du discours oral de personnes âgées saines vs. à risque (patients MCI) vs. pathologiques (patients souffrant de démences de type Alzheimer légère à modérée) – des indices pré-morbides de la maladie. Parmi les indicateurs linguistiques, nous avons nous intéressé spécifiquement à la densité des idées (DI), c'est-à-dire la qualité informative des propositions langagières. Pour mesurer la DI, nous avons adopté l'analyse prédictive pratiquée par Denhière (1991). Nous avons également calculer la complexité des propositions de chaque échantillon. La DI moyenne obtenue par l'analyse prédictive des données orales varie significativement selon les groupes de sujets. La qualité informative du discours oral des patients atteints de la maladie d'Alzheimer est plus faible que celle des patients MCI. Aussi, ce résultat montre que la densité des idées est un indicateur sensible pour différencier les patients MCI des personnes âgées saines.

MARDI 17/11 DE 9H45 À 12H05

• Session Orale - O3

Phonétique

Mardi 17/11 de 9h45 à 12h05, 1w08

112 : Identification des consonnes finales du vietnamien par des locuteurs natifs

Thi-Thuy-Hien Tran, Nathalie Vallée

- Département Parole et Cognition de GIPSA-lab

Résumé

A great difficulty encountered by Vietnamese subjects, who learn French, is that consonant clusters, which do not exist in Vietnamese, are mispronounced. This problem persists even after several years of practicing, and even when the French clusters correspond to Vietnamese consonant sequences. The general aim of our project is to identify the factors which are the main cause of this problem. In this paper, we examine the perception of syllable-final stops (/p/, /t/, /k/, /m/, /n/, /g/) in Vietnamese by 20 native Northern-Vietnamese listeners. Our findings suggest

that specific acoustic characteristics and probably the lexical frequency of final consonants lead the subjects in their choice of responses.

222 : Caractérisation automatique des accents étrangers

Abdelkarim Mars

– Laboratoire d’informatique de grenoble

Résumé

Parmi les phénomènes qui affectent la manière dont nous parlons, l’accent est une des composantes principales de la variation observée. La prononciation d’un locuteur peut en effet nous renseigner sur son origine, géographique et sociale. La description des caractéristiques phonétiques qui sous-tendent les différences d’accent perçues constitue donc un intérêt scientifique particulier. De plus, la recherche dans le domaine des accents contribue à l’amélioration d’applications technologiques telles que la reconnaissance de la parole et l’indexation du locuteur. Ce papier propose une étude phonétique acoustique des accents étrangers en français. Afin d’analyser à grande échelle les variations liées à l’origine de locuteur, nous avons évalué l’apport des outils automatiques de codage acoustico-phonétique et d’alignement forcé.

232 : Une Base de données Etiquetée Formantiquement en Langue Arabe Standard

Imen Jemaa, Oussama Rekhis, Kais Ouni, Yves Laprie

– Equipe Parole, LORIA Nancy1, France
– Unité de Recherche Traitement du Signal, Traitement de l’image et Reconnaissance de Formes

Résumé

While formant frequencies are known to play a critical role in human speech perception and in computer speech processing, there has been a lack of standard databases needed for the quantitative evaluation of automatic formant extraction techniques especially in Arabic language. We report in this paper our recent effort to create a reference database of the first three formant tracks. The manually Formant labeling is carried out using the Winsnoori tool. Furthermore, we present in this paper an exploratory use of the database to quantitatively evaluate the automatic LPC method implemented in the popular open source Praat using the hand edited formant trajectories as reference.

242 : Construction d’un corpus robuste de différents dialectes arabes

Mohamed Belgacem

– Laboratoire LIDILEM

Résumé

Notre article s’intègre dans le cadre du projet intitulé ‘Oréodule’ : un système embarqué temps réel de reconnaissance, de traduction et de synthèse de la parole arabe. L’objet de notre intérêt dans cet article est la présentation

d’un corpus vocal de la parole arabe. Nous détaillerons les étapes de constitution de ce corpus et les difficultés rencontrées lors de son élaboration. Nous intégrerons également les différents résultats pratiques obtenus lors de chaque phase (tailles des enregistrements, volume total du notre corpus, etc.).

92 : Perception d’expressions multimodales du Feeling of Thinking (états mentaux et affectifs, intentions, attitudes) en interaction

Anne Vanpé, Véronique Aubergé

– GIPSA-lab, Département Parole et Cognition (ex-ICP), UMR 5216 CNRS/Université de Grenoble

Résumé

Human-Machine Interaction, as interaction between two humans, can be considered as a dynamic process where the human is continuously communicating, even when he is “expressively” listening (informative backchannel and feedback). The present study analyses the audio-visual non speech expressions for two subjects in spontaneous HMI corpora, following an ethology-based methodology. First results reveal a large panel of values expressed outside of turns (e.g. mental states, intentions, attitudes, emotions) that we have globally called Feeling of Thinking. We have shown the role of static vs. dynamic processing of visual information and we are now attempting to investigate some specific non speech “vocal events”. Their temporal distribution seems to be particularly relevant for the perception of Feeling of Thinking expressions.

MARDI 17/11 DE 13H20 À 15H05

• Session Orale - O4

Traitement automatique de la parole

Mardi 17/11 de 13h20 à 15h05, 1w08

102 : Traduction automatique de la parole arabe/anglais par segmentations multiples

Fethi Bougares

– Laboratoire d’Informatique de Grenoble

Résumé

La traduction de la parole est un thème de recherche récent, car il combine deux problèmes scientifiques complexes : la reconnaissance de la parole et la traduction automatique. Dans ce contexte, nous nous intéressons à la construction d’un système de traduction statistique pour la paire de langues arabe/anglais. Ces deux langues sont de structures éloignées, ce qui nécessite plus d’effort de préparation et de segmentation des données textuelles ou orales à traduire. Après avoir mis en lumière la relation entre l’analyse morphologique de l’arabe et la qualité de traduction, nous abordons les problèmes relatifs à l’ambiguïté segmentale de l’arabe avec la formulation et l’intégration de la multi-segmentation dans un système de traduction statistique.

172 : Architecture d'un Système de Vérification Automatique du Locuteur appuyée par la Détection du Genre

Hayet Djellali, Radia Amirouche, Mohamed Tayeb Laskri

- Université de Badji Mokhtar Annaba
- Université de Badji Mokhtar Annaba

Résumé

We propose a new approach in Automatic speaker verification ASV based on detection Gender (male, female). We determine with speaker voice his gender. Knowing that, the speaker could be an impostor with opposite gender that he claims. The aim of this work is to experiment if detection gender module can improve speaker verification decision when we compare it with baseline ASV system.

212 : Corrections spécifiques du français sur les systèmes de reconnaissance automatique de la parole

Richard Dufour, Yannick Estève, Paul Deléglise

- LIUM - Université du Maine

Résumé

Automatic speech recognition (ASR) systems are used in a large number of applications, in spite of the inevitable recognition errors. In this study we propose a pragmatic approach to automatically repair ASR outputs by taking into account linguistic and acoustic information, using formal rules or stochastic methods. The proposed strategy consists in developing a specific correction solution for each specific kind of errors. In this paper, we apply this strategy on two case studies specific to French language. We show that it is possible, on automatic transcriptions of French broadcast news, to decrease the error rate of a specific error by 11.4% in one of two the case studies, and 86.4% in the other one. These results are encouraging and show the interest of developing more specific solutions to cover a wider set of errors in a future work.

272 : Modélisation Stochastique du Dialogue par Structures Sémantiques

Florian Pinault

- CERI-LIA

Résumé

Dans le domaine de l'interaction Homme-Machine, les systèmes de dialogue à initiative mixte sont actuellement étudiés, afin de permettre aux utilisateurs de parler librement avec la machine. Cependant, les systèmes de dialogue en langue naturelle manquent souvent de la robustesse nécessaire pour assurer la satisfaction de l'utilisateur. Une solution consiste à utiliser une représentation sémantique riche du dialogue, ainsi qu'une modélisation statistique du cours du dialogue.

62 : Vers des Modèles Autonomes de Reconnaissance Automatique de la Parole Multilingue

Sethserey Sam

- Laboratoire d'Informatique de Grenoble (LIG)

Résumé

In multilingual automatic speech recognition, one interesting research challenge is how to deal with a multilingual speech utterance (the utterance that contains different speech languages and/or native or non-native speech) ? In order to overcome this problem, we focus our research on autonomous acoustic models (AM) and language models (LM). Autonomous means the multilingual AM and LM are automatically re-adapted themselves, in every given time slot (5s or 10s), before final decoding. The re-adaptation of AM and ML models could be done based on a module called Autonomous observer. In this article, we introduce the concept of autonomous AM and ML in multilingual ASR system (for automatic phone transcription purpose) and also the techniques to create an observer module

MERCREDI 18/11 DE 13H40 À 15H25

• Session Orale - O5

Perception

Mercredi 18/11 de 13h40 à 15h25, 1w08

122 : Perception de la variation linguistique : étude comparative entre l'aire de Lesbos (Grèce) et celle des « vallées vaudoises » du Piémont occidental (Italie)

Silvia Gally, Maria Goudi

- GIPSA-lab UMR 5216, DPC – SLD Université Stendhal - Grenoble 3

Résumé

Dans cet article nous proposons une étude de dialectologie perceptuelle (DP) qui met en parallèle des travaux effectués dans deux aires linguistiques bien distinctes : l'île de Lesbos, en Grèce, et une zone du Piémont occidental, en Italie. Les données traitées dans ces études sont issues d'enquêtes de terrain dans les deux aires respectives.

182 : HMMs and GMMs based methods in acoustic-to-articulatory speech inversion

Atef Ben Youssef, Viet-Ahn Tran, Pierre Badin, Gérard Bailly

- DPC / GIPSA-lab, UMR 5216, Grenoble

Résumé

Afin de récupérer les mouvements des articulateurs tels que les lèvres, la mâchoire ou la langue, nous avons développé et comparé deux méthodes d'inversion basées l'une sur les modèles de Markov cachés (HMMs) et l'autre sur les modèles de mélanges de gaussiennes (GMMs).

Les mouvements des articulateurs sont représentés par les coordonnées médiosagittale de bobines d'un articulographe électromagnétique (EMA) fixées sur les articulateurs. Dans la première méthode, des HMMs à deux flux, acoustique et articulatoire, sont entraînés à partir des signaux acoustique et articulatoire synchrones. Le HMM acoustique sert à reconnaître les phones, ainsi que leurs durées. Ces informations sont ensuite utilisées par le HMM articulatoire pour synthétiser les trajectoires articulatoires. Pour la deuxième méthode, un GMM s'associe entre traits acoustique et articulatoire est entraîné sur le même corpus suivant le critère de minimum d'erreur quadratique moyenne (MMSE) à partir des trames acoustiques d'empan temporel plus ou moins grand. Pour un corpus de données EMA mono-locuteur enregistré par un locuteur français, l'erreur RMS de reconstruction sur le corpus de test pour la méthode fondée sur les HMMs se situe entre 1.96 et 2.32 mm, tandis qu'elle se situe entre 2.46 et 2.95 mm pour la méthode basée sur les GMMs.

202 : Rôle de l'information visuelle dans l'accès au lexique mental

Mathilde Fort, Justine Chipot, Sonia Kandel, Christophe Savariaux, Elsa Spinelli

- GIPSA-Lab
- Laboratoire de Psychologie et de Neurocognition

Résumé

Cette étude vise à déterminer le rôle de l'information visuelle dans l'accès au lexique. Pour cela, nous avons utilisé un paradigme d'amorçage phonologique. Les participants devaient effectuer une tâche de décision lexicale sur une cible présentée en modalité auditive. Cette cible était toujours précédée par une syllabe en amorce : cette dernière pouvait être présentée en modalité audiovisuelle (AV), auditive (A), ou visuelle seule (V). L'analyse des résultats sur les mots cibles indique un effet d'amorçage pour toutes les modalités de présentations de la syllabe. En conséquence, notre étude suggère que l'information visuelle seule permet d'activer les représentations de mots contenues dans le lexique mental.

252 : Espace perceptuel de similarité : étude sur 17 langues

Marie Rimbault Joffard

Résumé

The goal of the present study was to devise a means of representing languages in a perceptual similarity space based on their overall sound structures. In experiment 1, native French listeners performed a free classification task in which they grouped 17 diverse languages based on their overall similarity. A similarity matrix of the grouping patterns was then submitted to clustering and multidimensional scaling analyses. In experiment 2, the same group of French listeners sorted the 17 languages in term of their distance to French. Taken together, the results of the two experiments provide the basis for estimating the distance between a given mother tongue and other languages and for understanding the role of the phonological filter.

INDEX PAR AUTEURS – RJCP

Amirouche Radia	4	Fort Mathilde	5	Rekhis Oussama	3
Amy De La Breteque Benoit	2	Fredouille Corinne	2	Rimbault Joffard Marie	5
Aubergé Véronique	3	Gally Silvia	4	Romano Antonio	1
Badin Pierre	4	Georgeton Laurianne	1	Sam Sethserey	4
Bailly Gérard	4	Goudi Maria	4	Sato Marc	1
Barkat-Defradas Melissa	2	Grabski Krystyna	1	Savariaux Christophe	5
Barkat-Defradas Melissa	2	Guerrero Lopez Harold Andrés	2	Schwartz Jean-Luc	1
Belgacem Mohamed	3	Henrich Nathalie	1	Serrafero Patrick	2
Ben Youssef Atef	4	Jemaa Imen	3	Spinelli Elsa	5
Bougares Fethi	3	Kandel Sonia	5	Tran Thi-Thuy-Hien	2
Busset Julie	1	Lamalle Laurent	1	Tran Viet-Ahn	4
Chipot Justine	5	Laprie Yves	3	Trottier Catherine	2
Clement Pierre	2	Laskri Mohamed Tayeb	4	Vallée Nathalie	1, 2
Cornaz Sandra	1	Lee Hye Ran	2	Vanpé Anne	3
Deléglise Paul	4	Mars Abdelkarim	3	Vilain Coriandre	1
Djellali Hayet	4	Mondain Michel	2		
Dufour Richard	4	Ouni Kais	3		
Estève Yannick	4	Pinault Florian	4		