



Journées d'Études sur la Parole
 Traitement Automatique des Langues Naturelles
 Rencontre des Étudiants Chercheurs en Informatique pour le
 Traitement Automatique des Langues

PARIS Inalco du 4 au 8 juillet 2016
 Organisé par les laboratoires franciliens

<https://jep-taln2016.limsi.fr>



Conférenciers invités:

Christian Chiarcos (Goethe-Universität, Frankfurt.)

Mark Liberman (University of Pennsylvania, Philadelphia)

Coordinateurs comités d'organisation

Nicolas Audibert et Sophie Rosset (JEP)

Laurence Danlos & Thierry Hamon (TALN)

Damien Nouvel & Ilaine Wang (RECITAL)

Philippe Boula de Mareuil, Sarra El Ayari & Cyril Grouin (Ateliers)



©2016 Association Francophone pour la Communication Parlée (AFCP) et
Association pour le Traitement Automatique des Langues (ATALA)

Préface

Pour la cinquième fois, après Nancy en 2002, Fès en 2004, Avignon en 2008 et Grenoble en 2012, l'AFCP (Association Francophone pour la Communication Parlée) et l'ATALA (Association pour le Traitement Automatique des Langues) organisent conjointement leurs principales conférences pour réunir en un seul lieu les communautés du traitement des langues écrites, parlées et signées.

Plus précisément, la conférence JEP-TALN-RECITAL 2016 réunit cette année la 31^e édition des Journées d'Étude sur la Parole (JEP 2016), la 23^e édition de la conférence sur le Traitement Automatique des Langues Naturelles (TALN 2016) et la 18^e édition des Rencontres des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2016).

Cet événement est organisé par une quinzaine de laboratoires franciliens (ALPAGE, CLILLAC-ARP, ERTIM, IRCAM, LaTTiCe, LFF, LIPN, LIMSI, LPP, LSCP, LTCI, MoDyCo, PLIDAM, SFL, STIH) et se déroulera dans les locaux de l'INALCO (13^e arrondissement de Paris).

JEP-TALN-RECITAL 2016 accueille deux conférenciers invités, Christian Chiarcos (Johann Wolfgang Goethe Universität Frankfurt a. M., Allemagne) et Mark Liberman (LDC & University of Pennsylvania, USA) ainsi que neuf ateliers :

- Celtic Language Technology Workshop (CLTW)
- Défi Fouille de Texte (DEFT)
- Enseignement des Langues et TAL (ELTAL)
- La Voix à la Barre
- Risque et TAL : détection, prévention et gestion
- Traitement automatique de la parole non standard
- Communautés en ligne : outils et applications en TAL (COLTAL)
- Traitement automatique des langues africaines (TALAf)
- Hackathon dans le domaine du TAL (HackaTAL)

Nous remercions tous les relecteurs et membres des différents comités de programme pour leur travail ainsi que nos sociétés savantes, l'AFCP et l'ATALA dont le CPERM (comité permanent) assure la continuité des éditions successives de TALN.

Pour les JEP, 99 articles ont été soumis, parmi lesquels 86 ont été sélectionnés, soit un taux de sélection de 87 %. 35 seront présentés en session orale et 51 lors de sessions posters.

Pour TALN, 41 articles longs ont été soumis, parmi lesquels 23 ont été sélectionnés, soit un taux de sélection de 56 %. 53 articles courts ont été soumis parmi lesquels 32 ont été sélectionnés pour un poster, soit un taux de sélection de 60 %.

Pour RECITAL, 11 articles ont été soumis, parmi lesquels 8 ont été sélectionnés, soit un taux de sélection de 73 %. Le faible nombre de soumissions reçues peut être, pour partie, lié au fait que de nombreux doctorants en première année de thèse soumettaient un article court co-signé par

leur(s) encadrant(s). Nous regrettons cette pratique qui prive les jeunes chercheurs des relectures pédagogiques, approfondies et signées. RECITAL est une conférence volontairement accessible, dans laquelle des travaux préliminaires peuvent être soumis et, pour certains, donner lieu à des présentations orales.

Nous rappelons que les conférences sont soucieuses du respect des règles de déontologie de la recherche et de la publication scientifique¹, qui s'imposent à tous. Les comités de programme sont donc particulièrement vigilants à cet égard.

Les actes de JEP-TALN-RECITAL-2016 sont en ligne sur le site <https://jep-taln2016.limsi.fr/actes>. Ils seront référencés par l'ACL (Association for Computational Linguistics) dans l'ACL Anthology (<http://www.aclweb.org/anthology/>).

Nous espérons que ces actes donneront à leurs lecteurs des idées fructueuses pour faire avancer la recherche et qu'ils serviront au rayonnement de la communauté francophone spécialiste du traitement des langues écrites, parlées et signées.

S. Rosset (LIMSI-CNRS)	N. Audibert (Univ. Paris 3, LPP-CNRS)	Présidents JEP
T. Hamon (Univ. Paris 13, LIMSI-CNRS)	L. Danlos (Univ. Paris-Diderot, ALPAGE-INRIA)	Présidents TALN
D. Nouvel (INALCO, ERTIM)	I. WANG (Université Paris 10, MoDyCo-CNRS)	Présidents RECITAL

1. Ces règles sont rappelées entre autres dans les chartes des thèses et par le COMETS, voir en particulier les pages 11 à 13 de ce document : http://www.cnrs.fr/comets/IMG/pdf/guide_promouvoir_une_recherche_inte_gre_et_responsable_8septembre2014.pdf.

Message des présidents de l’AFCP et de l’ATALA

C’est avec un plaisir toujours renouvelé que nous assistons à la tenue conjointe des Journées d’Études sur la Parole (JEP), de la Conférence sur le Traitement Automatique des Langues Naturelles (TALN) et des Rencontres des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL). Ces conférences, issues de communautés scientifiques voisines, se sont retrouvées pour la première fois en 2002 à Nancy. Après le succès de la seconde organisation conjointe à Fès en 2004, l’ATALA et l’AFCP ont décidé d’une organisation commune tous les 4 ans. Cette cinquième édition conjointe — la 31^{ème} édition des JEP, la 23^{ème} de TALN et la 18^{ème} de RECITAL — s’inscrit dans une série dont nous espérons la pérennité. L’organisation conjointe de ces événements se pare d’une aura particulière tant l’intersection entre communication parlée et traitement automatique des langues est grande. Le regard croisé sur l’écrit, l’oral et le signe, dans ses dimensions scientifiques, technologiques, applicatives et humaines, nous semble essentiel pour faire avancer l’état des connaissances sur ces différentes formes du langage, qui sont autant d’objets d’étude riches et complexes et pour renforcer la recherche francophone dans ces domaines.

Cet événement (car il faut avant tout voir cette réunion de JEP – TALN – RECITAL comme une entité à part entière et non comme trois manifestations dans un même lieu) est évidemment avant tout un événement scientifique pour échanger et faire le point sur l’état de l’art. Mais, de manière tout aussi importante, c’est également l’occasion de renforcer les liens sociaux entre chercheurs, entre communautés scientifiques, entre institutions académiques et industrielles et d’intégrer les étudiants et jeunes diplômés, qui souhaitent faire carrière dans nos domaines, au sein d’une dynamique collective pour leur faire profiter du capital de savoir et d’infrastructures accumulé au cours du temps. Préserver et développer des conférences francophones comme JEP, TALN et RECITAL est à notre avis vital pour la communauté, car ce sont d’abord des points d’entrée dans le monde de la publication scientifique où les jeunes chercheurs peuvent faire leur premiers pas, sans avoir à affronter la barrière de la langue. C’est aussi contribuer à la préservation de la diversité culturelle, tout en restant ouvert sur le monde. On peut même se prendre à rêver un jour de réunir dans un même lieu avec JEP, TALN et RECITAL les événements comparables qui existent dans les pays de langues romanes.

Construction d’une communauté scientifique, promotion d’une dynamique collective autour de la communication par le langage, et structuration des acteurs dans la sphère francophone nous apparaissent comme autant d’éléments importants pour résister à la dilution des thématiques scientifiques dans les applications et dans la course à l’impact sociétal dont nous sommes les témoins. Ce constat est particulièrement vrai dans nos disciplines où il est notamment exacerbé par la pluridisciplinarité inhérente à nos travaux. Nous savons tous à quel point mener des travaux mêlant plusieurs disciplines scientifiques est difficile, notamment en terme d’impact et de valorisation. De plus, la course actuelle au transfert industriel quasiment immédiat pour les disciplines scientifiques ayant des prolongements technologiques, et la quête de l’impact sociétal poussent à limiter la pluridisciplinarité à un simple assemblage de spécialités. Nous sommes persuadés que c’est par la discussion, l’échange et la fertilisation croisée entre les différentes disciplines en jeu dans l’étude de la communication parlée, écrite, ou signée, que nous pourrions faire émerger de nouveaux axes de recherche et permettre aux contributions fondamentales d’avoir la place qu’elles méritent dans le paysage de la recherche internationale. En particulier, le partage de ressources (données annotées mais aussi algorithmes,

protocoles d'évaluation, guide d'annotations, etc.) joue un rôle important dans la structuration de notre communauté et dans l'émergence de recherches pluridisciplinaires s'appuyant sur l'ensemble de nos champs d'investigation.

La tenue d'un événement joint est aussi l'occasion de s'interroger sur la manière dont nous abordons dans nos communautés respectives les changements, parfois révolutionnaires, qui surviennent dans notre pratique quotidienne de la recherche. Plusieurs évolutions sont venues récemment bouleverser nos disciplines en profondeur : les masses de données (big data) et son corollaire, l'informatique en nuage (cloud computing), ouvrent de nouvelles portes sur le traitement massif de données de parole et d'écrit ; l'apprentissage profond a investi le paysage scientifique, notamment avec l'avènement des modèles distributionnels ; et, moins récemment, des nouveaux médias et des médias sociaux ont émergés à l'échelle planétaire en un temps record, modifiant l'usage de la langue. Épiphénomènes ou modification profonde des fondements de nos domaines ? Le débat est ouvert et l'avenir le dira. En attendant, ces rencontres seront à n'en pas douter l'occasion de faire le point sur ces évolutions fortes, sur le plan scientifique bien sûr, mais aussi sur le plan de l'évolution générale de nos domaines. Le regard croisé de plusieurs disciplines, le partage d'expérience et de vision entre les différents domaines seront autant de richesses apportées à ce débat.

Il nous reste à clore ces quelques réflexions d'ouverture, en remerciant l'ensemble des personnes qui ont rendus possibles cet événement qui sera, à n'en pas douter, riche et passionnant. L'ATALA et l'AFCP tiennent tout d'abord à remercier les organisateurs des JEP, de TALN et de RECITAL pour l'énorme travail qu'ils ont réalisé afin d'assurer la qualité scientifique et la convivialité de l'événement. Nous constatons avec plaisir que tout a été mis en œuvre pour maximiser les échanges entre les différents acteurs du domaine, tous secteurs et tous domaines confondus. Nos remerciements vont également à l'ensemble des membres des comités de programme, sans qui une conférence ne serait qu'une longue litanie. Enfin, et surtout, les véritables garants de la qualité scientifique de la conférence sont les nombreuses personnes ayant participé à l'évaluation des soumissions, c'est-à-dire... la communauté scientifique ! Un grand merci aux relecteurs pour le temps qu'ils ont dédié à ce travail anonyme et pour la qualité du travail effectué qui se reflète dans le programme que vous aurez le plaisir de découvrir tout au long de la semaine.

Nous vous souhaitons une très bonne conférence, faite de découvertes scientifiques, d'échanges fructueux et de convivialité, en espérant que cette expérience vous donnera d'ores et déjà envie d'un prochain rendez-vous dans quatre ans.

Guillaume Gravier, président de l'Association Francophone de la Communication Parlée
Patrick Paroubek, président de l'Association pour le Traitement Automatique des Langues

Comité d'organisation de JEP-TALN-RECITAL 2016 :

Martine Adda (CNRS, LPP)
Nicolas Audibert (Univ. Sorbonne Nouvelle, LPP)
Marion Blondel (CNRS, SFL)
Philippe Boula de Mareüil (CNRS, LIMSI)
Annelies Braffort (CNRS, LIMSI)
Ioana Chitoran (Univ. Paris Diderot, CLILLAC-ARP)
Chloé Clavel (Telecom-ParisTech, LTCI)
Alejandrina Cristia (CNRS, LSCP)
Laurence Danlos (Univ. Paris Diderot, ALPAGE)
Elisabeth Delais-Roussarie (CNRS, LLF)
Laurence Devillers (Univ. Paris-Sorbonne, LIMSI)
Sarra El Ayari (CNRS, SFL)
Michael Filhol (CNRS, LIMSI)
Karën Fort (Univ. Paris-Sorbonne, STIH)
Cécile Fougeron (Univ. Sorbonne Nouvelle, LPP)
Cyril Grouin (CNRS, LIMSI)
Thierry Hamon (Univ. Paris 13, LIMSI)
Agata Jackiewicz (Univ. Paris-Sorbonne, STIH)
Sylvain Kahane (Univ. Paris Ouest Nanterre, MoDyCo)
Jovan Kostov (INALCO, PLIDAM)
Thomas Lavergne (Univ. Paris-Sud, LIMSI)
Anaïs Lefeuvre (Univ. Paris-Sorbonne, STIH)
Anne-Laure Ligozat (ENSIIE, LIMSI)
Jean-Luc Minel (Univ. Paris Ouest Nanterre, MoDyCo)
Adeline Nazarenko (Univ. Paris 13, LIPN)
Aurélie Névéol (CNRS, LIMSI)
Damien Nouvel (INALCO, ERTIM)
Nicolas Obin (Univ. Pierre et Marie Curie, IRCAM)
Patrick Paroubek (CNRS, LIMSI)
Axel Roebel (Univ. Pierre et Marie Curie, IRCAM)
Sophie Rosset (CNRS, LIMSI)
Djamé Seddah (Univ. Paris-Sorbonne, ALPAGE)
Xavier Tannier (Univ. Paris-Sud, LIMSI)
Isabelle Tellier (Univ. Sorbonne Nouvelle, LaTTiCe)
Nadi Tomeh (Univ. Paris 13, LIPN)
Mathieu Valette (INALCO, ERTIM)
Cyril Verrecchia (CNRS, LIMSI)
Anne Vilnat (Univ. Paris-Sud, LIMSI)
Coralie Vincent (CNRS, SFL)
Ilaine Wang (Univ. Paris Ouest Nanterre, MoDyCo)
Pierre Zweigenbaum (CNRS, LIMSI)

Comité de programme de JEP 2016 :

Président :

Guillaume Gravier (CNRS, IRISA)

Membres :

Gilles Adda (CNRS, LIMSI)
Nicolas Audibert (Univ. Sorbonne Nouvelle, LPP)
Melissa Barkat-Defradas (Univ. Paul Valéry Montpellier 3, PRAXILING)
Jean-François Bonastre (Univ. d'Avignon et des Pays de Vaucluse, LIA)
Nathalie Camelin (Univ. du Maine, LIUM)
Alejandrina Cristia (CNRS, LSCP)
Elisabeth Delais-Roussarie (Univ. Paris Diderot, LLF)
Véronique Delvaux (FNRS/Univ. de Mons, Service de Métrologie et Sciences du Langage, Belgique)
Camille Fauth (Univ. de Strasbourg, LiLPa)
Benoit Favre (Aix-Marseille Univ., LIF)
Emmanuel Ferragne (Univ. Paris Diderot, CLILLAC-ARP)
Corinne Fredouille (Univ. d'Avignon et des Pays de Vaucluse, LIA)
Nathalie Henrich Bernardoni (CNRS, GIPSA-lab)
Frédéric Isel (Univ. Paris Ouest Nanterre, MoDyCo)
Juliette Kahn (LNE)
David Langlois (Univ. de Lorraine, LORIA)
Yohann Meynadier (Aix-Marseille Univ., LPL)
Claude Montacie (Univ. Paris-Sorbonne, STIH)
Nicolas Obin (Univ. Pierre et Marie Curie, IRCAM)
Thomas Pellegrini (Univ. Paul Sabatier Toulouse, IRIT)
Sophie Rosset (CNRS, LIMSI)
Sophie Wauquier (Univ. Paris 8, SFL)

Relecteurs additionnels :

Martine Adda-Decker (CNRS, LPP)
Moez Ajili (Université d'Avignon et des Pays de Vaucluse, LIA)
Vincent Arnaud (Université du Québec à Chicoutimi, Canada)
Grégoire Bachman (CNRS, LPP)
Pierre Badin (CNRS, GIPSA-lab)
Loïc Barrault (Université du Maine, LIUM)
Claude Barras (Université Paris-Sud, LIMSI)
Nathalie Bedoin (Université Lyon 2, DDL)
Frédéric Berthommier (Université Grenoble Alpes, GIPSA-lab)
Louis-Jean Boë (Université Grenoble Alpes, GIPSA-lab)
Anne Bonneau (CNRS, LORIA)
Stéphanie Borel (Université Sorbonne Nouvelle, LPP)
Fethi Bougares (Université du Maine, LIUM)
Philippe Boula De Mareüil (CNRS, LIMSI)
Damien Bouvier (CNRS, IRCAM)
Annelies Braffort (CNRS, LIMSI)
Mélanie Canault (Université Lyon 1, DDL)
Meunier Christine (CNRS, LPL)
Ioana Chitoran (Université Paris Diderot, CLILLAC-ARP)
Vincent Colotte (Université de Lorraine, LORIA)
Lise Crevier Buchman (CNRS, LPP)
Sébastien Delecraz (Aix-Marseille Université, LIF)
Paul Deléglise (Université du Maine, LIUM)
Didier Demolin (Université Sorbonne Nouvelle, LPP)
Giovanni Depau (Université Grenoble Alpes, GIPSA-lab)
Laurence Devillers (Université Paris-Sorbonne, LIMSI)
Mariapaola D'Império (Aix-Marseille Université, LPL)
Christelle Dodane (Université Paul Valéry Montpellier 3, PRAXILING)
Marion Dohen (Université Grenoble Alpes, GIPSA-lab)
David Doukhan (INA)
Richard Dufour (Université d'Avignon et des Pays de Vaucluse, LIA)
Camille Dutrey (LNE)
Yannick Estève (Université du Maine, LIUM)
Camille Fauth (Université de Strasbourg, LiLPa)
Isabelle Ferrané (Université Paul Sabatier Toulouse, IRIT)
Dominique Fohr (CNRS, LORIA)
Cécile Fougeron (Université Sorbonne Nouvelle, LPP)
Olivier Galibert (LNE)
Jiayin Gao (Université Sorbonne Nouvelle, LPP)
Silvia Gally (Université Grenoble Alpes, GIPSA-lab)
Maëva Garnier (CNRS, GIPSA-lab)
Cédric Gendrot (Université Sorbonne Nouvelle, LPP)
Laurianne Georgeton (Aix-Marseille Université, LPL)
Arseniy Gorin (CNRS, LIMSI)
James German (Aix-Marseille Université, LPL)
Alain Ghio (CNRS, LPL)
Anne Guyot Talbot (Université Paris Diderot, CLILLAC-ARP)
Bernard Harmegnies (Université de Mons, MSL, Belgique)
Sophie Herment (Aix-Marseille Université, LPL)
Thomas Hueber (CNRS, GIPSA-lab)
Céline Hidalgo (Aix-Marseille Université, LPL)
Fabrice Hirsch (Université Paul Valéry Montpellier 3, PRAXILING)
Kathy Huet (Université de Mons, MSL, Belgique)
Stéphane Huet (Université d'Avignon et des Pays de Vaucluse, LIA)
Irina Illina (Université de Lorraine, LORIA)
Bassam Jabaian (Université d'Avignon et des Pays de Vaucluse, LIA)
Philippe Joly (Université Paul Sabatier Toulouse, IRIT)
Denis Jovet (INRIA, LORIA)
Takeki Kamiyama (Université Paris 8, DEPA, LPP)
Jennifer Krzonowski (CNRS, DDL)
Barbara Kuhnert (Université Sorbonne Nouvelle, LPP)
Anne Lacheret (Université Paris Ouest Nanterre, MoDyCo)
Muriel Lalain (CNRS, LPL)
Yves Laprie (CNRS, LORIA)
Anthony Larcher (I2R, A*STAR, Singapore)
Jérôme Lechien (Université de Mons, MSL, Belgique)
David Le Gac (Université de Rouen, Dysola)
Thierry Legou (CNRS, LPL)
Marco Liuni (IRCAM)
Marianne Louis (LPL)
Thibault Magallon (Aix-Marseille Université, LIF)
Sylvain Meignier (Université du Maine, LIUM)
Odile Mella (Université de Lorraine, LORIA)
Yohann Meynadier (Aix-Marseille Université, LPL)
Olivier Michalon (Aix-Marseille Université, LIF)
Alexis Michaud (CNRS, LACITO)
Julia Monnin (Université de la Nouvelle-Calédonie, CNEP)
Pascal Nocera (Université d'Avignon et des Pays de Vaucluse, LIA)
Francois Pellegrino (CNRS, DDL)
Pascal Perrier (Université Grenoble Alpes, GIPSA-lab)
Myriam Piccaluga (Université de Mons, MSL, Belgique)
Cristel Portes (Aix-Marseille Université, LPL)
Angélique Remacle (Université de Liège, ULV)
Rachid Ridouane (Université Sorbonne Nouvelle, LPP)
Albert Rilliard (CNRS, LIMSI)

Solange Rossato (Université Grenoble Alpes, LIG)
Amélie Rochet-Capellan (Université Grenoble Alpes, GIPSA-lab)
Marc Sato (CNRS, LPL)
Christophe Savariaux (Université Grenoble Alpes, GIPSA-lab)
Jean-Luc Schwartz (Université Grenoble Alpes, GIPSA-lab)
Christine Senac (Université Paul Sabatier Toulouse, IRIT)
Takaaki Shochi (Université Bordeaux Montaigne, CLLE ERSSàB)
Jérémy Tafforeau (Aix-Marseille Université, LIF)

Marie Tahon (IRISA)
Anne Tortel (Aix-Marseille Université, LPL)
Thi Thuy Hien Tran (Université Stendhal, GIPSA-lab)
Jérémy Trione (Aix-Marseille Université, LIF)
Gabor Turcsan (Aix-Marseille Université, LPL)
Nathalie Vallee (Université Grenoble Alpes, GIPSA-lab)
Ioana Vasilescu (CNRS, LIMSI)
Christophe Veaux (IRCAM)
Coriandre Vilain (Université Grenoble Alpes, GIPSA-lab)
Emmanuel Vincent (INRIA, LORIA)
Hiyon Yoo (Université Paris Diderot, LLF)

Comité de programme de TALN 2016 :

Présidents :

Laurence Danlos (Univ. Paris-Diderot, ALPAGE)
Thierry Hamon (Univ. Paris 13, LIMSI)

Membres du comité de programme :

Laurent Besacier (Univ. de Grenoble, LIG)
Vincent Claveau (CNRS, IRISA)
Olivier Ferret (CEA LIST)
Laurence Danlos (Univ. Paris Diderot, ALPAGE)
Gaël Dias (Univ. de Caen Basse-Normandie, GREYC)
Thierry Hamon (Univ. Paris 13, LIMSI)

Nabil Hathout (CNRS, CLLE)
Philippe Langlais (Univ. de Montréal, RALI, Canada)
Laurence Meurant (Univ. de Namur, LSFb, Belgique)
Emmanuel Morin (Univ. de Nantes, LINA)
Adeline Nazarenko (Univ. Paris 13, LIPN)
Pascale Sébillot (INSA de Rennes, IRISA)

Membres du comité de relecture :

Stergos Afantenos (Univ. Paul Sabatier, IRIT)
Salah Aït-Mokhtar (Xerox Research Centre Europe)
Maxime Amblard (Univ. de Lorraine, LORIA)
Jean-Yves Antoine (Univ. François Rabelais Tours/Blois, LI)
Delphine Battistelli (Univ. Paris Ouest Nanterre, MoDyCo, CNRS)
Frédéric Béchet (Aix-Marseille Univ., LIF)
Delphine Bernhard (Univ. de Strasbourg, LiLPa)
Romaric Besançon (CEA, LIST)
Brigitte Bigi (CNRS, LPL)
Philippe Blache (CNRS, LPL)
Hervé Blanchon (Univ. Pierre-Mendès-France, LIG)
Marion Blondel (Univ. Paris 8, SFL)
Florian Boudin (Univ. de Nantes, LINA)
Annelies Braffort (CNRS, LIMSI)
Nathalie Camelin (Univ. du Maine, LIUM)

Thierry Charnois (Univ. Paris 13, LIPN)
Guillaume Cleuziou (Univ. d'Orléans, LIFO)
Benoit Crabbé (Univ. Paris Diderot, ALPAGE)
Beatrice Daille (Univ. de Nantes, LINA)
Marco Dinarelli (CNRS, LaTTiCe)
Iris Eshkol (Univ. d'Orléans, LLL)
Yannick Estève (Univ. du Maine, LIUM)
Cécile Fabre (Univ. Toulouse - Jean Jaurès, CLLE)
Benoit Favre (Aix-Marseille Univ., LIF)
Karën Fort (Univ. Paris-Sorbonne, STIH)
Thomas François (Univ. Catholique de Louvain, CENTAL, Belgique)
Nathalie Friburger (Univ. François Rabelais, LI)
Éric Gaussier (Univ. Joseph Fourier, LIG)
Natalia Grabar (CNRS, STL)
Lamia Hadrich Belguith (Univ. de Sfax, MIRACL, Tunisie)

Nicolas Hernandez (Univ. de Nantes, LINA)
Stéphane Huet (Univ. d'Avignon et des Pays de Vaucluse, LIA)
Sylvain Kahane (Univ. Paris Ouest Nanterre, MoDyCo)
Olivier Kraif (Univ. Stendhal Grenoble 3, LIDILEM)
Mathieu Lafourcade (Univ. de Montpellier, LIRMM)
Guy Lapalme (Univ. de Montréal, Canada)
Joseph Le Roux (Univ. Paris 13, LIPN)
Jean-Marc Lecarpentier (Univ. de Caen Basse-Normandie, GREYC)
Anne-Laure Ligozat (ENSIIE, LIMSI)
Denis Maurel (Univ. François Rabelais Tours, LI)
Richard Moot (CNRS, LaBRI)
Véronique Moriceau (Univ. Paris-Sud, LIMSI)
Philippe Muller (Univ. Paul Sabatier Toulouse, IRIT)
Jian-Yun Nie (Univ. de Montréal, Canada)
Aurélié Névéol (CNRS, LIMSI)
Yannick Parmentier (Université d'Orléans, LIFO)
Thierry Poibeau (CNRS, LaTTiCe)

Andrei Popescu-Belis (IDIAP Research Institute, Suisse)
Jean-Philippe Prost (Univ. de Montpellier, LIRMM)
Solen Quiniou (Univ. de Nantes, LINA)
Christian Raymond (INSA de Rennes, IRISA)
Christian Retoré (Univ. de Montpellier, LIRMM)
Mathieu Roche (Cirad, TETIS)
Didier Schwab (Univ. Pierre-Mendès-France, LIG)
Djamé Seddah (Univ. Paris-Sorbonne, ALPAGE)
Kamel Smaïli (Univ. de Lorraine, LORIA)
Xavier Tannier (Univ. Paris-Sud, LIMSI)
Isabelle Tellier (Univ. Sorbonne Nouvelle, LaTTiCe)
Juan-Manuel Torres-Moreno (Univ. d'Avignon et des Pays de Vaucluse, LIA)
Christel Vrain (Univ. d'Orléans, LIFO)
Eric Wehrli (Univ. de Genève, Suisse)
Guillaume Wisniewski (Univ. Paris-Sud, LIMSI)
François Yvon (Univ. Paris-Sud, LIMSI)
Pierre Zweigenbaum (CNRS, LIMSI)

Relecteurs additionnels :

Chafik Aloulou (Univ. de Sfax, MIRACL, Tunisie)
Ibrahim Bounhas (Univ. de la Manouba, LISI, Tunisie)
Eric De La Clergerie (INRIA, ALPAGE)
Maher Jaoua (Univ. de Sfax, MIRACL, Tunisie)
Hugo Mougard (Univ. de Nantes, LINA)

Yuliya Korenchuk (Univ. de Strasbourg, LiLPa)
Ophélie Lacroix (CNRS, LIMSI)
Souha Mezghani (Univ. de Sfax, MIRACL, Tunisie)
Sarah Zenasni (Univ. de Montpellier, LIRMM)

Comité de programme de RECITAL 2016 :

Présidents :

Damien Nouvel (Inalco, ERTIM)
Ilaine Wang (Univ. Paris 10, MoDyCo)

Membres :

Maxime Amblard (Univ. de Lorraine, LORIA)
Jean-Yves Antoine (Univ. François Rabelais Tours/Blois, LI)
Patrice Bellot (Polytech Marseille - Aix-Marseille Univ., LSIS)
Houda Bouamor (Carnegie Mellon University, Qatar)
Florian Boudin (Univ. de Nantes, LINA)
Annelies Braffort (CNRS, LIMSI)
Sandra Bringay (Univ. Paul Valéry Montpellier 3, LIRMM)
Marie Candito (Univ. Paris Diderot, ALPAGE)
Gaël de Chalendar (CEA, LIST)
Marcel Cori (Univ. Paris Ouest Nanterre, MoDyCo)
Benjamin Duthil (Univ. de La Rochelle, L3I)

Maud Ehrmann (EPFL, Suisse)
Benoit Favre (Aix-Marseille Univ., LIF)
Karën Fort (Univ. Paris-Sorbonne, STIH)
Thomas François (Univ. Catholique de Louvain, CENTAL, Belgique)
Bruno Gaume (Univ. Toulouse - Jean Jaurès, CLLE)
Mathieu Lafourcade (Univ. de Montpellier, LIRMM)
Philippe Langlais (Univ. de Montréal, RALI, Canada)
Charlotte Lecluze (Univ. de Caen Basse-Normandie, GREYC)
Benjamin Lecouteux (Univ. Pierre-Mendès-France, LIG)
Anaïs Lefeuvre (Univ. Paris-Sorbonne, STIH)
Gaël Lejeune (Univ. de Caen Basse-Normandie, GREYC)

Laurence Longo (Univ. de Strasbourg, Lilpa)
Romain Loth (Institut des Systèmes Complexes)
Pierre Magistry (Kodex-Lab)
Jean-Marc Marty (Proxem)
Yann Mathet (Univ. de Caen Basse-Normandie, GREYC)
François Morlane-Hondère (CNRS, LIMSI)
Yayoi Nakamura-Delloye (INALCO, Centre d'Etudes
Japonaises)
Alexander Panchenko (TU Darmstadt, Allemagne)
Alexander Pak (Google)
Yannick Parmentier (Université d'Orléans, LIFO)
Gaël Patin (XiKO)
Laurent Prévot (Aix-Marseille Univ., LPL)

Jean-Philippe Prost (Univ. de Montpellier, LIRMM)
Christian Raymond (INSA de Rennes, IRISA)
Charlotte Roze (CommunicoTool)
Yves Scherrer (Univ. de Genève, LATL, Suisse)
Jérémy Tafforeau (Aix-Marseille Univ., LIF)
Isabelle Tellier (Univ. Sorbonne Nouvelle, LaTTiCe)
Nadi Tomeh (Univ. Paris 13, LIPN)
Juan-Manuel Torres-Moreno (Univ. d'Avignon et des
Pays de Vaucluse, LIA)
Yannick Toussaint (LORIA)
Mathieu Valette (INALCO, ERTIM)
Pierre Zweigenbaum (CNRS, LIMSI)

Coordinateurs ateliers :

Cyril Grouin (CNRS, LIMSI)
Sarra El Ayari (CNRS, SFL)
Philippe Boula de Mareuil (CNRS, LIMSI)

Soutiens

Partenaires institutionnels



Entreprises



JEP-TALN-RECITAL 2016 est soutenu par Google

Table des matières

Communications orales

Apprentissage d'analyseur en dépendances cross-lingue par projection partielle de dépendances

Ophélie Lacroix, Lauriane Aufrant, Guillaume Wisniewski, François Yvon 1

Bleu, contusion, ecchymose : tri automatique de synonymes en fonction de leur difficulté de lecture et compréhension

Thomas François, Mokhtar B. Billami, Núria Gala, Delphine Bernhard 15

Comparaison d'approches de classification automatique des actes de dialogue dans un corpus de conversations écrites en ligne sur différentes modalités

Soufian Salim, Nicolas Hernandez, Emmanuel Morin 29

Construire un lexique de sentiments par crowdsourcing et propagation

Mathieu Lafourcade, Nathalie Le Brun, Alain Joubert 43

Détection de concepts pertinents pour le résumé automatique de conversations par recombinaison de patrons

Jérémy Trione, Benoit Favre, Frédéric Béchet 57

Détection et classification non supervisées de relations sémantiques dans des articles scientifiques

Kata Gábor, Isabelle Tellier, Thierry Charnois, Haïfa Zargayouna, Davide Buscaldi 70

Etude de l'impact d'un lexique bilingue spécialisé sur la performance d'un moteur de traduction à base d'exemples

Nasredine Semmar, Othman Zennaki, Meriama Laib 84

Étude de réseaux de neurones récurrents pour l'étiquetage de séquences

Marco Dinarelli, Isabelle Tellier 98

Évaluation de l'apprentissage incrémental par analogie

Vincent Letard, Gabriel Illouz, Sophie Rosset 112

Évaluation des modèles sémantiques distributionnels : le cas de la dérivation syntaxique

Gabriel Bernier-Colborne, Patrick Drouin 125

Évaluation d'une nouvelle structuration thématique hiérarchique des textes dans un cadre de résumé automatique et de détection d'ancres au sein de vidéos

Anca Simon, Guillaume Gravier, Pascale Sébillot 139

<i>Exploitation de reformulations pour l'acquisition d'un vocabulaire expert/non expert</i> Edwide Antoine, Natalia Grabar	153
<i>Extension lexicale de définitions grâce à des corpus annotés en sens</i> Loïc Vial, Andon Tchechmedjiev, Didier Schwab	167
<i>Extraction de lexiques bilingues à partir de corpus comparables spécialisés à travers une langue pivot</i> Alexis Linard, Emmanuel Morin, Béatrice Daille	180
<i>Fouille de motifs et CRF pour la reconnaissance de symptômes dans les textes bio-médicaux</i> Pierre Holat, Nadi Tomeh, Thierry Charnois, Delphine Battistelli, Marie-Christine Jaulent, Jean-Philippe Métivier	194
<i>Une méthode non-supervisée pour la segmentation morphologique et l'apprentissage de morphotactique à l'aide de processus de Pitman-Yor</i> Kevin Löser, Alexandre Allauzen	207
<i>Modèles adaptatifs pour prédire automatiquement la compétence lexicale d'un apprenant de français langue étrangère</i> Anaïs Tack, Thomas François, Anne-Laure Ligozat, Cédrick Fairon	221
<i>Modélisation unifiée du document et de son domaine pour une indexation par termes-clés libre et contrôlée</i> Adrien Bougouin, Florian Boudin, Béatrice Daille	235
<i>Ne nous arrêtons pas en si bon chemin : améliorations de l'apprentissage global d'analyseurs en dépendances par transition</i> Lauriane Aufrant, Guillaume Wisniewski, François Yvon	248
<i>Prédiction automatique de fonctions pragmatiques dans les reformulations</i> Natalia Grabar, Iris Eshkol-Taravella	262
<i>Projection Interlingue d'Étiquettes pour l'Annotation Sémantique Non Supervisée</i> Othman Zennaki, Nasredine Semmar, Laurent Besacier	276
<i>Utilisation des relations d'une base de connaissances pour la désambiguïsation d'entités nommées</i> Romaric Besançon, Hani Daher, Olivier Ferret, Hervé Le Borgne	290
<i>Word2Vec vs DBnary ou comment (ré)concilier représentations distribuées et réseaux lexico-sémantiques ? Le cas de l'évaluation en traduction automatique</i> Christophe Servan, Zied Elloumi, Hervé Blanchon, Laurent Besacier	304

Communications sous forme de poster

<i>Amélioration de la traduction automatique d'un corpus annoté</i>	
Marwa Hadj Salah, Hervé Blanchon, Mounir Zrigui, Didier Schwab	318
<i>Analyse d'une tâche de substitution lexicale : quelles sont les sources de difficulté ?</i>	
Ludovic Tanguy, Cécile Fabre, Camille Mercier	325
<i>L'anti-correcteur : outil d'évaluation positive de l'orthographe et de la grammaire</i>	
Lydia-Mai Ho-Dac, Sophie Muller, Valentine Delbar	333
<i>Appariement d'articles en ligne et de vidéos : stratégies de sélection et méthodes d'évaluation</i>	
Adèle Désoyer, Delphine Battistelli, Jean-Luc Minel	342
<i>Approximate unsupervised summary optimisation for selections of ROUGE</i>	
Natalie Schluter, Héctor Martínez Alonso	349
<i>L'architecture d'un modèle hybride pour la normalisation de SMS</i>	
Eleni Kogkitsidou, Georges Antoniadis	355
<i>Une catégorisation de fins de lignes non-supervisée</i>	
Pierre Zweigenbaum, Cyril Grouin, Thomas Lavergne	364
<i>Classification automatique de dictées selon leur niveau de difficulté de compréhension et orthographique</i>	
Adeline Müller, Thomas François, Sophie Roekhaut, Cédric Fairon	372
<i>Combiner des modèles sémantiques distributionnels pour mieux détecter les termes évoquant le même cadre sémantique</i>	
Gabriel Bernier-Colborne, Patrick Drouin	381
<i>Comparing Named-Entity Recognizers in a Targeted Domain : Handcrafted Rules vs. Machine Learning</i>	
Ioannis Partalas, Cédric Lopez, Frédérique Segond	389
<i>Compilation de grammaire de propriétés pour l'analyse syntaxique par optimisation de contraintes</i>	
Jean-Philippe Prost, Rémi Coletta, Christophe Lecoutre	396
<i>Découverte de nouvelles entités et relations spatiales à partir d'un corpus de SMS</i>	
Sarah Zenasni, Eric Kergosien, Mathieu Roche, Maguelonne Teisseire	403
<i>Description de la juxtaposition en Langue des Signes Française à partir d'une grammaire récursive</i>	
Mohamed Hadjadj, Michael Filhol	411

<i>Détecter le besoin d'information dans des requêtes d'utilisateurs d'agents virtuels : sélection de données pertinentes</i>	
Octavia Efrain, Fabienne Moreau	419
<i>Estimer la notoriété d'un nom propre via Wikipedia</i>	
Mouna Elashter, Denis Maurel	428
<i>Étiquetage multilingue en parties du discours avec MElt</i>	
Benoît Sagot	435
<i>Extraction d'expressions-cibles de l'opinion : de l'anglais au français</i>	
Grégoire Jadi, Laura Monceaux, Vincent Claveau, Béatrice Daille	443
<i>Extraction d'opinions ambiguës dans des corpus d'avis clients</i>	
Joseph Lark, Emmanuel Morin, Sebastián Peña Saldarriaga	451
<i>Extraction de relations temporelles dans des dossiers électroniques patient</i>	
Julien Tourille, Olivier Ferret, Aurélie Névél, Xavier Tannier	459
<i>Impact de l'agglutination dans l'extraction de termes en arabe standard moderne</i>	
Wafa Neifar, Thierry Hamon, Pierre Zweigenbaum, Mariem Ellouze Khemakhem, Lamia Hadrich Belguith	467
<i>Inbenta Semantic Clustering : un outil de classification non-supervisée hybride</i>	
Quintana Manon, Planes Laurie	475
<i>Intégration de la similarité entre phrases comme critère pour le résumé multi-document</i>	
Maâli Mnasri, Gaël de Chalendar, Olivier Ferret	482
<i>Investigating gender adaptation for speech translation</i>	
Rachel Bawden, Guillaume Wisniewski, Hélène Maynard	490
<i>Médias traditionnels, médias sociaux : caractériser la réinformation</i>	
Cédric Maigrot, Ewa Kijak, Vincent Claveau	498
<i>Mise au point d'une méthode d'annotation morphosyntaxique fine du serbe</i>	
Aleksandra Miletic, Cécile Fabre, Dejan Stosic	506
<i>Patrons sémantiques pour l'extraction de relations entre termes - Application aux comptes rendus radiologiques</i>	
Lionel Ramadier, Mathieu Lafourcade	514
<i>Recherche de «périsegments» dans un contexte d'analyse conceptuelle assistée par ordinateur : le concept d'«esprit» chez Peirce</i>	
Davide Pulizzotto, José A. Lopez, Jean-François Chartier, Jean-Guy Meunier, Louis Chartrand, Francis Lareau, Tan Le Ngoc	522
<i>Segmentation automatique d'un texte en rhèmes</i>	
Constance Nin, Victor Pineau, Béatrice Daille, Solen Quiniou	531

<i>Système hybride pour la reconnaissance des entités nommées arabes à base des CRF</i> Emna Hkiri, Souheyl Mallat, Mounir Zrigui	539
<i>Vers un lexique ouvert des formes fléchies de l'alsacien : génération de flexions pour les verbes</i> Lucie Steiblé, Delphine Bernhard	547
<i>Vers une analyse des différences interlinguistiques entre les genres textuels : étude de cas basée sur les n-grammes et l'analyse factorielle des correspondances</i> Marie-Aude Lefer, Yves Bestgen, Natalia Grabar	555