

## Initiative Corpus de référence du français 2

28-29 mars 2013 Amphithéâtre Emeraude, ParisTech

46 rue Barrault 75013 Paris

L'Institut de Linguistique Française (CNRS, 2393 FR) organise pour la seconde fois deux journées de réflexion et de débats sur le thème « Initiative Corpus de référence du français ». Le comité directeur de l'ILF a jugé qu'il entre effectivement dans les missions de la fédération de lancer une telle initiative au niveau national (corpus de textes écrits, corpus oraux, corpus de référence ouvert à la diachronie longue).

La France, contrairement à d'autres pays ne s'est pas dotée d'un tel type de corpus. Or, les avancées scientifiques et technologiques, le développement des programmes et des infrastructures dans le domaine de ce qu'il est convenu d'appeler « Corpus » permettent de considérer que le contexte actuel se prête à l'accueil favorable d'une telle initiative.

Les deux journées s'articuleront autour de présentations de 45 minutes suivies de 15 minutes de discussion avec le public. À partir des conclusions de cette réflexion, un calendrier et des modalités de travail seront définis pour lancer concrètement l'initiative.

Nous serions heureux de vous accueillir nombreux à participer à ces journées qui se dérouleront les 28 et 29 mars prochains à l'amphithéâtre Emeraude de ParisTech à Paris (75013).

*Stéphanie GIRAULT et Franck NEVEU pour le Comité d'Organisation Scientifique*

-----  
Important :

**L'entrée est libre, mais l'inscription est obligatoire.** Vous devez impérativement remplir le formulaire d'inscription accessible sur le site de l'ILF (page d'accueil) et à cette adresse : <http://form.jotforme.com/form/30715229237351>

-----  
**PROGRAMME**

Jeudi 28 mars 2013

9H30-10H - Franck NEVEU (directeur de l'ILF, porteur du Consortium Corpus-Ecrits)  
Ouverture

10H-10H30 - Serge Heiden (ICAR - CNRS, ENS-Lyon) & Sophie Prevost (LATTICE - CNRS, ENS-Ulm)  
« Expérience d'instrumentation d'un corpus diachronique du français : couplage de la plateforme TXM avec le corpus GGHF »

10H30-11H30 - Peter Blumenthal (Département des langues romanes - Université de Cologne) & Denis Vigier (ICAR - Université Lumière Lyon II)

« Pour un corpus diachronique de référence du français »

11H30-12H30 - Agnès Tutin (LIDILEM, Université Stendhal-Grenoble 3)

« Un corpus de textes spécialisés : les écrits scientifiques »

12H30-14H - Pause déjeuner

14H00-15H - Frédéric Sabio (Laboratoire Parole et Langage (LPL) - Université de Provence)

« Corpus Oraux et annotation en 'unités d'énoncés' »

15H-16H - Irène Fegnolio (Institut des textes et manuscrits modernes (ITEM) - CNRS, ENS-Ulm)

« Quelle part archives et manuscrits peuvent-ils prendre à la constitution de corpus de référence ? »

16H-17H - Karèn Fort (Ecole nationale supérieure des mines de Nancy, LORIA)

« La charte Ethique et Big Data : pour des corpus sans zone d'ombre »

17H-18H - Benoît Sagot (Analyse Linguistique Profonde A Grande Echelle (ALPAGE) - INRIA Paris Rocquencourt et Université Paris 7 - Denis Diderot)

« Construction, annotation et exploitation: contributions réciproques entre le Traitement Automatique des Langues et un Corpus de référence »

Vendredi 29 mars 2013

9H30-10H30 - Pierre Larrivée (CRISCO, Université de Caen Basse-Normandie) & Florence Lefeuvre (Syled-Clesthia Université Sorbonne Nouvelle - Paris 3)

« Des corpus de référence pour une grammaire du français vernaculaire contemporain »

10H30-11H30 - Philippe Blache (Laboratoire Parole et Langage (LPL) - Université de Provence) & Jean-Marie Pierrel (porteur de l'Equipex ORTOLANG, ATILF) :

« ORTOLANG : un support technique pour le développement de corpus de référence »

11H30-12H30 - Bernard Colombat (HTL, Université Paris Diderot), Gilles Souvay (ATILF) & Arnaud Pelfrêne (ENS-Lyon)

« Quelles contributions le *CTLF* et le *Grand Corpus des grammaires françaises* peuvent-ils apporter à un corpus de référence du français? »

12H30-14H - Pause déjeuner

14H00-15H00 - Thierry Chanier (LRL, Université Blaise Pascal, Clermont-Ferrand)

"Expériences d'élaboration des corpus de référence du hollandais et de l'allemand. Projet de noyau de corpus nouvelles formes de communication en français".

15H-17H30 - TABLE RONDE

Participants : Olivier Baude (LLL-Université d'Orléans & DGLFLF), Philippe Blache (Laboratoire Parole et Langage (LPL) - Université de Provence), Bernard Laks (Modyco, Université Paris-Ouest-Nanterre-La Défense), Jean-Marie Pierrel (porteur de l'Equipex ORTOLANG, ATILF), Agnès Steuckardt (Praxiling, Université Paul-Valéry, Montpellier 3)