

Colloque international  
**Corpus de français parlés et français des corpus**

08-09 mai 2014  
Université de Neuchâtel, Av. du 1er-Mars 26 (salle C49)

*Comité d'organisation*

Marie-José Béguelin, Federica Diémoz (U. Neuchâtel)  
Mathieu Avanzi (U. Paris Diderot & U. Neuchâtel)

*Comité de soutien scientifique*

Alain Berrendonner (U. Fribourg) ; José Deulofeu (U. Provence) ; Andres Kristol (U. Neuchâtel) ;  
Corinne Rossari (U. Neuchâtel) ; Dominique Willems (U. Gand)

*Argumentaire scientifique*

Depuis un peu plus d'une décennie, on assiste dans le domaine de la francophonie à un intérêt croissant des chercheurs pour la constitution de grands corpus de français parlé, si bien que le retard par rapport aux autres langues comme l'anglais, l'espagnol ou le néerlandais, que regrettait Blanche-Benveniste [1999], est en train de diminuer. De nombreuses publications visant à présenter les différentes bases de données existantes [Bruxelles et al. 2009] ou à en faire le recensement [Cappeau & Sejjido 2005 ; Cappeau & Gadet 2007 ; Gadet 2013] ont vu le jour. Un guide des bonnes pratiques [Baude 2005] et des consortiums<sup>1</sup> ont été mis en place pour que les efforts des uns et des autres soient quelque peu canalisés et que les conventions de transcription et d'archivage se retrouvent d'un corpus à l'autre, ce qui permettrait de les rendre homogènes.

Dans ce contexte, le colloque projeté a pour but de réunir quelques-uns des acteurs ayant contribué de façon significative à la mise au point de corpus ou d'archives de français parlés. Au total, dix corpus seront présentés dans le cadre de ce colloque. Ces corpus contiennent des données enregistrées aux quatre coins de la francophonie (corpus PFC et I-PFC), des données de français « non-hexagonales » propres à un pays déterminé (la Suisse pour OFROM, la Belgique pour VALIBEL, le Québec pour CFPQ), des enregistrements réalisés à Paris et ses alentours (CFPP, MPF) et à Orléans (ESLO) et des enregistrements dont l'origine des locuteurs est moins contrôlée (CLAPI et ORFEO). Outre le critère géographique, ces corpus se distinguent par la variation des situations de parole (entretiens à dominante monologique pour les uns ; situations interactionnelles diverses enregistrées en contexte écologique pour les autres), mais aussi de par les objectifs originaux qui ont guidé leur constitution (étude de la phonologie, des interactions, de la syntaxe, de la prosodie, etc.), par les conventions de transcription adoptées (transcription orthographique ponctuée ou non, codage du para-verbal, etc.) et par les études quantitatives qu'ils permettent à l'heure actuelle.

Le colloque aura pour mission de faire le point sur les aspects qui caractérisent ces dix corpus de variétés de français. Mais nous souhaitons également qu'il permette d'établir un bilan sur ce que ces grands corpus nous apprennent sur les français parlés à l'orée du 21<sup>ème</sup> siècle, en France métropolitaine comme dans les autres pays de la francophonie.

---

<sup>1</sup> Cf. p. ex. les projets IRCOM (<http://ircom.corpus-ir.fr/site/accueil.php>) et ORTOLANG (<http://www.ortolang.fr>).

*Programme prévisionnel*

**Jeudi 08 mai 2014**

13.00-13.30 : Accueil des participants

13.30-14.00 : Présentation du colloque

14.00-15.00 : « Multicultural Paris French »

F. Gadet (U. Paris Ouest Nanterre)

15.00-16.00 : « La base CLAPI, des données vidéo pour comprendre les interactions »

V. Traverso (U. Lyon 2)

16.00-16.30 : Pause

16.30-17.30 : « Les ESLOs, du portrait sonore au paysage digital »

O. Baude (U. Orléans)

17.30-18.30 : « Le CFPP2000 : constitution, outils et analyses »

S. Branca & F. Lefevre (La Sorbonne Nouvelle, Paris 3)

**Vendredi 09 mai 2014**

08.30-09.30 : « PFC, codages et représentations : la question du schwa »

J. Durand (U. Toulouse-le-Mirail et I.U.F.) & I. Racine (U. Genève)

09.30-10.30 : « La liaison dans un corpus d'apprenants: le projet Interphonologie du Français Contemporain (IPFC) »

I. Racine (U. Genève) & S. Detey (U. Waseda)

10.30-11.00 : Pause

11.00-12.00 : « Recherches sur le français parlé en Suisse romande – la base de données OFROM »

M. Avanzi, M.-J. Béguelin, F. Diémoz (U. Neuchâtel)

12.00-13.00 : « VaLiBel - La gestion des corpus oraux à l'ère du numérique »

C. Bolly, L. Degand, M. Francard, A. C. Simon (U. Louvain-la-Neuve)

13.00-14.30 : Repas

14.30-15.30 : « Le corpus de français parlé au Québec (CFPQ) et la langue des conversations familiales. Exemple de mise à profit des données à partir d'un examen lexico-sémantique de la séquence quasi-figée *c'est ça* »

G. Dostie (U. Sherbrooke)

15.30-16.30 : « Le projet ORFEO : outils, annotations et résultats attendus »

J.-M. Debaisieux (La Sorbonne Nouvelle, Paris 3) & C. Benzitoun (U. Lorraine)

16.30-17.30 : Clôture du colloque