

Recherches linguistiques et corpus

Responsable Franck Neveu

Le thème *Recherches linguistiques et corpus* mis en place au sein du laboratoire STIH de la faculté des Lettres de Sorbonne Université a pour objet le développement d'une réflexion commune et croisée, d'ordre épistémologique et méthodologique, sur la notion de corpus telle qu'elle est exploitée aujourd'hui dans les **sciences du langage**, et, plus largement, dans les **sciences humaines et sociales**. On y traite des questions relatives aux notions de donnée, d'observable, d'empiricité, de théorie (lien type/occurrence), de variable contextuelle, d'annotation, de codage, de catégorisation, etc. On s'interroge sur la fonction des corpus dans l'activité de recherche. Ce thème transversal est organisé en **séminaires** ouverts notamment aux *chercheurs*, aux *enseignants-chercheurs*, aux *ingénieurs*, aux *doctorants* et aux *étudiants de master*. Les séances sont constituées de deux conférences, suivies de débats.

Mercredi 10 mars 2021, 14h-16h30

Faculté des Lettres, Sorbonne Université

Séance organisée par visioconférence

Lien zoom

<https://us02web.zoom.us/j/84787650246?pwd=NGRWaWtZcGhsaExFWHFicXZyQyt1QT09>

**Quels observables linguistiques pour une observation outillée ?
Expérimentations et réflexions sur le travail statistique sur corpus**

Gaël Lejeune

Faculté des Lettres, Sorbonne Université, laboratoire STIH

**Les animaux « liminaires » :
observations de terrain et précisions sémantiques.
Quand les sciences du langage se nourrissent de l'éthologie**

Astrid Guillaume

Faculté des Lettres, Sorbonne Université, laboratoire STIH

Séminaire ouvert

Contact : franck.neveu@sorbonne-universite.fr

Quels observables linguistiques pour une observation outillée ? Expérimentations et réflexions sur le travail statistique sur corpus

Gaël Lejeune

Maître de Conférences à Sorbonne Université, laboratoire STIH

Dans cette présentation j'illustrerai certains problèmes épistémologiques liés à la caractérisation automatique d'observables dans des textes. La question qui m'intéresse est de savoir dans quelle mesure on peut (ou on doit) avoir des observables qui soient motivés linguistiquement, identifiables automatiquement et efficaces dans une perspective descriptive ou applicative. J'illustrerai la question avec des exemples provenant de différents corpus, allant de l'imprimé du XVII^{ème} siècle (projet Antonomaz) aux tweets (projet Memes) en passant par les corpus de presse.

Les animaux « liminaires » : observations de terrain et précisions sémantiques. Quand les sciences du langage se nourrissent de l'éthologie

Astrid Guillaume

Maîtresse de Conférences HDR, Sorbonne Université, laboratoire STIH

Le terme « liminaire » est régulièrement utilisé associé aux animaux mais il reste très peu connu du grand public et des SHS. Qu'est-ce qu'un animal liminaire? Quel est le comportement d'un animal liminaire ? Comment définir précisément d'un point de vue lexical et donc sémantique la notion de « liminarité animalière » ?

Avec à l'appui un corpus de vidéos et de photos basé sur une observation de terrain, diurne et nocturne de cinq mois, d'animaux liminaires dans leur habitat, cette conférence aura plusieurs objectifs. D'abord montrer comment la précision du lexique peut se faire au contact d'une véritable observation de terrain, ensuite comment les sciences du langage peuvent s'approprier la méthodologie scientifique de l'éthologie, enfin comment la zoosémiotique se dresse comme une passerelle efficace reliant les sciences du langage et les sciences du vivant.