

Linguistique de corpus / informatique + MTU



Niveau d'étude
Bac +2



ECTS
3 crédits



Composante
UFR Langage,
lettres et arts
du spectacle,
information et
communication

- > **Langue(s) d'enseignement:** Français
- > **Forme d'enseignement :** Cours magistral
- > **Ouvert aux étudiants en échange:** Non

Présentation

Description

Ce cours a pour objectif d'étudier la variation des usages langagiers oraux et écrits à travers l'étude de corpus. Les principaux corpus du français et leurs applications seront présentés, ainsi que la problématique de la constitution de corpus et les protocoles de transcriptions et d'annotation. Les questions éthiques et juridiques sont également abordées.

Un autre objectif de ce cours est d'aborder les corpus oraux et écrits de manière complémentaire. À partir de cette complémentarité à la fois méthodologique et théorique, nous souhaitons mettre en avant le caractère empirique de la linguistique outillée et la multiplicité des approches et des applications concrètes.

Dans l'idée de rompre avec le côté technique et théorique de ce domaine, les problématiques théoriques concernant le recueil de données, le statut du chercheur, l'élaboration d'un protocole d'élicitation, les questions de transcriptions et d'annotations et les façons d'exploiter ces corpus sont abordées :

- via la lecture d'articles,
- via l'étude des corpus existants,

- et via des travaux pratiques : les étudiants sont ainsi amenés à construire leur propre (petit) corpus et/ou à travailler sur une exploitation ciblée d'un corpus existant pour amener l'étudiant à pratiquer par lui-même la linguistique de corpus lors de TD interactifs, dans une dynamique « de la théorie à la pratique ».

Les travaux pratiques proposés ont pour objectif de mettre en évidence quelques phénomènes de variations :

1. Variation des systèmes vocaliques ou Intonation en français.

Étude de quelques phénomènes lexicaux et syntaxiques dans des corpus oraux transcrits et des corpus écrits

Objectifs

Objectifs

Connaissances théoriques:

- le rôle des corpus en linguistique
- les principaux corpus du français et leur diversité
- la notion de protocole de recueil de données
- les objectifs des conventions de transcriptions
- les questions éthiques et juridiques
- les avantages et les limites de quelques outils et traitements automatiques de corpus
- les différentes postures d'exploration du corpus et de son analyse
- les différentes applications concrètes de la linguistique de corpus

Heures d'enseignement

Linguistique de corpus / informatique - CM	CM	12h
Linguistique de corpus / informatique - TD	TD	18h

Pré-requis recommandés

Les connaissances vues en cours de L1 en phonétique articulatoire et en L2 en phonétique acoustique sont indispensables.

La connaissance des notions de morphosyntaxe de base et une certaine aisance avec l'outil informatique est souhaitable.

Contrôle des connaissances

Contrôle continu

Période : Semestre 4

Compétences visées

Mise en œuvre de méthodes et d'outils du champ disciplinaire

- Concevoir et expliciter une problématique de linguistique et mettre en œuvre son traitement scientifique avec les outils et méthodes de référence, notamment les outils informatiques spécifiques de la discipline.
- Recueillir et exploiter des matériaux linguistiques (lexiques, retranscriptions, corpus...)

A savoir :

- Comprendre les différentes phases de la constitution de corpus, de la réflexion théorique préalable à l'exploitation des données.
- Savoir recueillir, transcrire, annoter des documents écrits et/ou oraux, comprendre une convention de transcription et savoir l'appliquer, comprendre les enjeux de la procédure d'annotation
- Apprendre à utiliser les ressources inhérentes à la linguistique outillée, corpus et outils informatiques de traitement de corpus
- Utiliser des outils automatiques, évaluer leur apport

C'est pour appuyer ces aspects informatiques que 6h de MTU sont couplés à ce cours.

Bibliographie

Detey S., Durand J., Laks B., Lyche Ch. (2010). Les variétés du français parlé dans l'espace francophone. Paris : Ophrys.

Baude O. (coord.), C. Blanche-Benveniste, M.-F. Calas, P. Cappeau, P. Cordereix, L. Goury, M. JACOBSON, I. de Lamberterie, C. Marchello-Nizia et L. Mondada, 2006, Corpus oraux - Guide des bonnes pratiques 2006. CNRS Editions & Presses Universitaire d'Orléans.

Habert B., Nazarenko A., Salem A (1997). *Les linguistiques de corpus*. Paris : Armand Colin.

Léon, Jacqueline. « Aux sources de la 'Corpus Linguistics'#: Firth et la London School », 2008 [🔗 https://halshs.archives-ouvertes.fr/halshs-00331646](https://halshs.archives-ouvertes.fr/halshs-00331646).

Mayaffre, Damon. « Rôle et place du corpus en linguistique. Réflexions introductives ». In *Actes du colloque JETOU'2005*, édité par Pascale Vergely, 5#17. Université de Toulouse-Le Mirail, 2005. [🔗 https://hal.archives-ouvertes.fr/hal-00553742](https://hal.archives-ouvertes.fr/hal-00553742).

Jacques, Marie-Paule. « Une linguistique outillée, pour quels objets#? » *Histoire Epistémologie Langage* 38, n° 2 (2016): 87#99. [🔗 https://doi.org/10.1051/hel/2016380207](https://doi.org/10.1051/hel/2016380207).

Zufferey, Sandrine. *Introduction à la linguistique de corpus*. ISTE Group, 2020.

Cori, Marcel, et Sophie David. « Les corpus fondent-ils une nouvelle linguistique#? » *Langages* n° 171, n° 3 (2008): 111#29.

Infos pratiques

Lieu(x) ville

› Grenoble

Campus

› Grenoble - Domaine universitaire