

- Faculté des lettres et sciences humaines
- www.unine.ch/lettres

Linguistique de corpus : introduction (2PR1466)

Filières concernées	Nombre d'heures	Validation	Crédits ECTS
Pilier B A - Humanités numériques	Cours: 2 ph	Voir ci-dessous	3
Pilier B A - linguistique et dialectologie	Cours: 2 ph	Voir ci-dessous	3
Pilier B A - sciences du langage	Cours: 2 ph	Voir ci-dessous	3

ph=période hebdomadaire, pg=période globale, j=jour, dj=demi-jour, h=heure, min=minute

Période d'enseignement:

- Semestre Automne

Equipe enseignante

Chloé Tahar

Contenu

Ce séminaire porte sur l'exploitation des corpus dans une démarche quantitative pour l'étude de la sémantique lexicale.

La linguistique de corpus étudie la langue à partir de son utilisation réelle dans des collections raisonnées de textes - c'est-à-dire des corpus. La sémantique lexicale, elle, est l'étude du sens des mots (qu'ils soient lexicaux, comme le mot "alphabet", ou grammaticaux, comme le mot "pourquoi").

Au fil des séances, nous prendrons appui sur une série d'études de cas pour explorer un (ou plusieurs) corpus à partir d'une question - et progressivement, d'une hypothèse - sur le sens des mots. Les mots dont nous interrogerons prioritairement le sens sont des mots modaux et des mots argumentatifs. Les moyens que nous emploierons pour explorer nos hypothèses sont des méthodes de mesure du vocabulaire (des plus simples aux plus complexes) dans les corpus textuels.

Forme de l'évaluation

Contrôle continu (30% de la note finale)

Dossier écrit (70% de la note finale) à rendre pour le 30 janvier au plus tard.

Modalités de rattrapage

En cas d'échec, le rattrapage consistera en un remaniement du dossier en s'appuyant sur les retours et commentaires de l'enseignante. La remise de ce dossier sera attendue pour le 11 mars 2024.

Documentation

Les lectures recommandées pour chaque séance sont sur le Moodle

Pré-requis

Aucun

Forme de l'enseignement

Séminaire. Il est nécessaire d'apporter avec soi un ordinateur portable pour effectuer les applications pratiques du séminaire sur le logiciel TXM.

Notez bien qu'il vous sera impossible d'effectuer les exercices sur une tablette.

Objectifs d'apprentissage

Au terme de la formation l'étudiant-e doit être capable de :

- Distinguer et organiser les différents sens et usages d'un mot
- Examiner la distribution d'un phénomène - c'est-à-dire le type de contexte dans lequel apparaît cet élément
- Mesurer si un phénomène est anormalement sous- ou sur-représenté dans une partie du corpus
- Identifier les associations significatives d'un phénomène - à savoir les formes lexicales avec lesquelles il se combine de façon statistiquement significative
- Analyser la répartition du stock lexical entre les différentes parties d'un corpus

- Faculté des lettres et sciences humaines
- www.unine.ch/lettres

Linguistique de corpus : introduction (2PR1466)