

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

Approches structurales : Comment utiliser les corpus (2LN1135)

Filières concernées	Nombre d'heures	Validation	Crédits ECTS
Pilier B A - Humanités numériques	Séminaire: 2 ph	Voir ci-dessous	6
Pilier B A - littérature française	Séminaire: 2 ph	Voir ci-dessous	6
Pilier B A - sciences du langage	Séminaire: 2 ph	Voir ci-dessous	6

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

Période d'enseignement:

- Semestre Automne

Equipe enseignante

Chloé Tahar
Bureau : Pierre-à-Mazel 7 (bureau 310)
Heures de disponibilité : sur rendez-vous
Adresse électronique pour le courriel : chloé.tahar@unine.ch

Contenu

La linguistique de corpus étudie la langue à partir d'exemples de son utilisation dans la vie réelle. Utilisant des méthodes quantitatives, cette discipline permet de révéler des régularités linguistiques à partir d'indications chiffrées sur la fréquence d'occurrence de certains phénomènes (mots, tournures syntaxiques, etc.) dans des collections raisonnées de textes. Ce cours introduit en premier lieu à la notion de linguistique de corpus en discutant de l'importance des faits de fréquence textuelle dans l'étude de la langue. La découverte des fonctions et la pratique du logiciel d'analyse statistique de données textuelles TXM permettront ensuite aux étudiant.e.s de mettre en évidence des régularités linguistiques dans un corpus. On construira des dispositifs de comparaison entre corpus afin d'analyser la variation linguistique dans le temps ou selon le système de communication (langue parlée, langue écrite).

Forme de l'évaluation

Il sera demandé aux étudiants de produire un dossier guidé qui se construira au fil des séances et d'en faire une présentation orale (environ 15 min.) lors des/de la dernière-s séance-s. Le dossier attendu comportera des analyses chiffrées de données de corpus sous la forme d'une série de tableaux et de graphiques, ainsi qu'une analyse et une interprétation des données. Le dossier devra être remis 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

Ce séminaire s'appuie en partie sur le manuel de Biber & Conrad (2012), *Corpus linguistics: investigating language structure and use*. Les lectures recommandées seront rendues disponibles sur la page Moodle du cours au fil des séances. Les diapositives des séances seront postées sur la page Moodle peu de temps après la séance, chaque lundi.

Pré-requis

Aucune connaissance spécifique n'est attendue en informatique ou statistique. Il s'agit d'un cours d'initiation à ces méthodes.

Forme de l'enseignement

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

- Analyser un corpus à l'aide d'un logiciel d'analyse de données textuelles
- Concevoir des graphiques pertinents pour traduire de manière efficace des données chiffrées

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

Approches structurelles : Comment utiliser les corpus (2LN1135)

- Interpréter les principaux calculs statistiques utilisés en textométrie

Compétences transférables

- Communiquer les résultats d'une analyse à l'oral et à l'écrit
- Gérer un projet scientifique de l'élaboration des hypothèses à l'analyse et l'interprétation des résultats