

		<b>Master</b>	<b>Technologies des langues</b>		
options	Informatique et Traduction	Traitement Automatique des Langues			
Mention	<b>Traduction et Interprétation</b>	Faculté	des <b>Langues</b>	Université de Strasbourg	



# Guide pédagogique 2023/2024

# Sommaire

Contacts.....	2
Page web du master.....	2
Calendrier universitaire.....	2
Informations importantes sur le plagiat.....	2
Description du master.....	3
Architecture générale du master.....	4
MASTER 1 SEMESTRE 1 (TRONC COMMUN).....	5
UE1 Transferts linguistiques 1.....	5
UE2 Anglais et linguistique.....	6
UE3 Informatique.....	6
UE4 Outils pour l'analyse de corpus.....	7
UE5 Professionnalisation et ouverture.....	8
MASTER 1 SEMESTRE 2 (TRONC COMMUN).....	9
UE1 Transferts linguistiques 2.....	9
UE2 Anglais et linguistique.....	9
UE3 Informatique.....	10
UE4 Initiation à la recherche.....	11
UE5 Professionnalisation et ouverture.....	11
MASTER 1 SEMESTRE 1 (OPTION Informatique et Traduction).....	13
UE2 Anglais et linguistique.....	13
MASTER 1 SEMESTRE 2 (OPTION Informatique et Traduction).....	14
UE2 Anglais et linguistique.....	14
MASTER 1 SEMESTRE 1 (OPTION Traitement Automatique des Langues).....	15
UE2 Anglais et linguistique.....	15
MASTER 1 SEMESTRE 2 (OPTION Traitement automatique des langues).....	18
UE2 Anglais et linguistique.....	18
MASTER 2 SEMESTRE 3 (TRONC COMMUN).....	21
UE1 Initiation à la recherche.....	21
UE2 Informatique.....	21
UE3 Communication scientifique.....	22
UE4 Traitement automatique des Langues et Traduction.....	23
UE5 Ouverture.....	25
MASTER 2 SEMESTRE 4 (TRONC COMMUN).....	25
UE1 Mémoire de recherche.....	26
UE2 Stage.....	26
Stage : Mode d'emploi 2023-2024.....	27

## Contacts

### **Doyen de la Faculté des Langues :**

Thomas MOHNIKE

Le Patio, Bâtiment IV, 3ème étage - Bureau 20

### **Responsable administratif et financier de la Faculté**

Étienne GUIDAT

Courriel : [eguidat@unistra.fr](mailto:eguidat@unistra.fr)

### **Responsables pédagogiques du Master**

Amalia TODIRASCU (contact pour le Master 2)

Le Portique côté rue, 5<sup>ème</sup> étage - Bureau 510

Courriel : [todiras@unistra.fr](mailto:todiras@unistra.fr)

Delphine BERNHARD (contact pour le Master 1)

Le Patio, Bâtiment V, bureau 51.20

Courriel : [dbernhard@unistra.fr](mailto:dbernhard@unistra.fr)

### **Secrétariat**

Cédric BRAGA

Le Patio, Bâtiment V, 1er étage, Le Patio – Bureau 51.30

Courriel : [c.braga@unistra.fr](mailto:c.braga@unistra.fr)

## Page web du master

<https://langues.unistra.fr/informatique/master-technologies-des-langues/#top>

## Calendrier universitaire

[https://www.unistra.fr/fileadmin/upload/unistra/etudes/Annee\\_universitaire\\_2023-2024/Calendrier\\_universitaire\\_2023-2024.pdf](https://www.unistra.fr/fileadmin/upload/unistra/etudes/Annee_universitaire_2023-2024/Calendrier_universitaire_2023-2024.pdf)

## Informations importantes sur le plagiat

« Le plagiat est l'acte de faire passer pour siens les textes ou les idées d'autrui. La fraude est un acte de tromperie fait pour gagner un avantage personnel, au détriment des autres. »

(Source: <http://www.integrite.umontreal.ca/definitions/fraude.html> (consulté le 5-07-2012) )

**Le plagiat peut conduire à un refus de délivrance de diplôme, d'inscription dans le niveau supérieur, par un refus d'autorisation de soutenance ou sanctionné par une note invalidant le semestre.**

Informations complémentaires :

<https://services-numeriques.unistra.fr/organisation-du-numerique/direction-du-numerique-dnum/prevention-du-plagiat-a-luniversite-de-strasbourg.html>

# Description du master

## Présentation générale

L'objectif du Master Technologies des Langues (anciennement Linguistique Informatique Traduction option Linguistique Informatique) est de permettre aux étudiantes et étudiants titulaires d'une licence de langues (Langues étrangères appliquées, Langues, littératures et civilisations étrangères et régionales) ou de lettres (Lettres Modernes, Sciences du langage) d'obtenir une double compétence, à la fois en langues ou linguistique (selon l'option: Informatique et Traduction ou Traitement automatique des langues) et en informatique. Cette double compétence leur permettra de candidater à des postes en lien avec leur première compétence en langue et demandant des compétences linguistiques et informatiques (constitution de ressources linguistiques numériques multilingues, annotation de données, analyse de données, évaluation et contrôle de qualité de systèmes de traitement automatique des langues, documentation et rédaction technique) ou des postes à dominante informatique nécessitant de solides compétences en langues (développement informatique dans le domaine de l'ingénierie des langues).

Ce parcours a pour objectif la formation de spécialistes dans le domaine des industries de la langue. Cette formation permet d'accéder à un doctorat dans un laboratoire de recherche spécialisé dans le domaine ou dans le cadre d'une thèse de recherche appliquée financée par une entreprise.

## Savoir-faire et compétences

### Informatique et analyse de données

- Maîtriser les techniques algorithmiques les plus courantes dans le développement logiciel, la programmation impérative et orientée objet ;
- Connaître les standards et technologies du web ;
- Maîtriser les traitements nécessaires à la collecte, au pré-traitement, à la représentation et à l'analyse de données textuelles, les expressions régulières, les principes de base des systèmes de gestion de bases de données.

### Traitement Automatique des Langues et linguistique de corpus

- Développer des applications dans le domaine du Traitement Automatique des Langues (fouille de textes, apprentissage automatique, extraction d'informations) ;
- Créer, maintenir, documenter et annoter des ressources linguistiques (dictionnaires, grammaires, corpus) pour les outils de Traitement Automatique des Langues, d'aide à la traduction ou à la rédaction technique ;
- Maîtriser les outils d'étiquetage et d'analyse syntaxique ;
- Connaître les méthodes de traitement statistique de corpus ;
- Connaître les concepts de base en terminographie, savoir construire des bases terminologiques ;

### Langues et sciences du langage

- Maîtriser deux langues de rédaction scientifique et technique : l'anglais et le français ;
- Maîtriser les techniques de communication orale en anglais et en français ;

### Gestion de projet et professionnalisation

- Être capable de collecter des informations et mettre en place des dispositifs de veille documentaire et informationnelle ;
- Être capable de mener à bien un projet de recherche personnel, de piloter des projets et connaître les outils de gestion du travail collaboratif en équipe.

### Spécificités de l'option informatique et traduction :

- Maîtriser les techniques avancées de communication orale et écrite en anglais ;
- Connaître le mode de fonctionnement des outils de recherche d'information et de traduction automatique.

### Spécificités de l'option Traitement Automatique des Langues :

- Maîtriser les notions fondamentales en linguistique ;
- Connaître les concepts de base en lexicologie, lexicographie ou d'autres domaines de la linguistique.

## Insertion professionnelle

Les diplômés pourront occuper des postes principalement dans le domaine de l'ingénierie linguistique (visant la création de ressources linguistiques et le développement des applications spécifiques de traitement automatique des langues) : infolinguiste, linguiste informaticien, ingénieur linguiste, consultant linguiste, analyste de données / *data analyst*, rédacteur technique, terminologue etc. Des débouchés sont aussi possibles dans le domaine de l'informatique et du développement Web : ingénieur d'études, ingénieur recherche et développement, développeur

## Poursuites d'études

La formation ayant à la fois une dimension recherche et une dimension professionnelle, il sera également possible pour les étudiants de poursuivre des études de doctorat en Traitement Automatique des Langues, dans un laboratoire de recherche spécialisé dans le domaine ou dans le cadre d'une thèse de recherche appliquée financée par une entreprise (contrat CIFRE).

## IMPORTANT : Passage du M1 au M2

**Attention** : pour accéder en M2, il faut obtenir une moyenne de **10/20 à chaque UE du M1**. La compensation au sein d'une UE est possible mais la compensation entre UE ne s'applique pas. Par ailleurs, il n'est pas possible d'être inscrit à la fois en M1 et en M2.

## Architecture générale du master

Voir <https://langues.unistra.fr/formation/masters/master-traduction-et-interpretation/technologies-des-langues-traitement-automatique-des-langues-informatique-et-traduction#data-rof-tab-cours>

# MASTER 1 SEMESTRE 1 (TRONC COMMUN)

## UE1 Transferts linguistiques 1

Cette UE comprend trois cours magistraux en amphithéâtre d'une heure chacun. Les étudiants doivent choisir 2 cours sur 3.

### 1) TI11GM22 - Histoire de la traduction et réflexions théoriques (Alain Volclair), 12H CM

#### Compétences :

Découvrir les différentes approches de la traduction au cours de l'histoire. Appréhender les différentes écoles et les différents courants pour asseoir ses choix traductionnels et éviter les écueils. Les "belles infidèles". Approche cibliste ou sourcière. Approche interprétative etc.

Légitimer l'apprenti traducteur et mesurer l'importance et le pouvoir de la médiation linguistique et culturelle au fil du temps.

#### Contenus :

Acquérir une culture de la traduction et replacer le traducteur dans son contexte historique.

#### Modalités de contrôle des connaissances :

**Session 1 :** examen écrit, durée 1h30

**Session 2 :** examen écrit, durée 1h30

### 2) LGC4GM11 - Méthodologie de la recherche et outils de traduction (Thierry Grass), 12H CM

#### Compétences :

Acquérir les méthodes du travail scientifique et de la recherche bibliographique ; connaître les outils de traduction (programmes de TAO, dictionnaires électroniques, sites spécialisés) existants.

#### Contenus :

Éléments techniques pour construire une recherche documentaire (délimiter un cadre ou projet de recherche, connaître les lieux ressources, les outils). Construire une bibliographie. L'environnement du traducteur en termes de matériel et de logiciels.

#### Modalités de contrôle des connaissances :

**Session 1 :** examen, durée 30 min

**Session 2 :** examen, durée 30 min

### 3) TI11GM22 - Terminologie (Gercelia Mendes), 12H CM

#### Compétences :

Langue de spécialité et langue courante ; terminologie nomenclature / étude descriptive / science ; caractéristiques de la langue de spécialité. Les caractères notionnels : intrinsèques & extrinsèques ; système notionnel. Le terme et la formation du terme ; définitions ; terminologie descriptive et prescriptive ; formation des termes : dérivations ; confixation ; composition ; lexicalisation. Terminographie ; dépouillement, renseignement, saisie ; fiches terminologiques unilingue et bilingue. Bibliographie ; bibliographie raisonnée ; référence bibliographique : livre, article (périodiques ou autres), Internet, multimédia, entrevue, etc. ;

#### Contenus :

Initiation à la terminologie/terminographie dans une perspective traductionnelle.

#### Modalités de contrôle des connaissances :

**Session 1 :** devoir à rendre au cours du semestre et examen écrit (1 heure)

**Session 2 :** examen, durée 1 heure

## **UE2 Anglais et linguistique**

Le cours d'anglais compréhension et expression orales commun aux étudiants des deux options.

### **1) LGB2GM12 - Anglais compréhension et expression orales / CEO (Mme Kautzmann), 18H TD**

Compétences à acquérir :

Écouter / comprendre des discours oraux

Réagir oralement (restitution, commentaire) à des discours oraux (conférences, débats, etc.)

Contenus :

Compréhension / restitution et commentaire oral de documents audiovisuels spécifiques

Modalités de contrôle des connaissances :

**Session 1 :** Exercices de communication orale en groupe + participation individuelle (Coef. 0,5)

**Session 2 :** Examen de compréhension orale de 1 h

Pour l'option Informatique et traduction : voir la description spécifique de l'UE2 (page 13)

Pour l'option Traitement Automatique des Langues : voir la description spécifique de l'UE2 (page 15)

## **UE3 Informatique**

### **1) LGC4GM31 - Algorithmique et programmation niveau 1 (Delphine Bernhard), 36H TD**

Compétences :

- Maîtriser les techniques de programmation en Python
- Connaître les concepts de base de l'algorithmique et des structures de données

Contenus :

Découverte du langage Python. Généralités autour de la notion d'algorithmique et des questions liées à l'informatisation des documents textuels. Écriture de programmes en utilisant les variables, les opérations d'affectation, les boucles, les instructions conditionnelles, les structures de données simples et complexes (listes et dictionnaires) ainsi que les expressions régulières. Mise en application de ces notions par des exercices impliquant de combiner ces divers éléments.

Modalités de contrôle des connaissances :

**Session 1 :** travail pratique noté + examen sur machine (1h) + épreuve orale (15 min)

**Session 2 :** épreuve orale (20 min)

### **2) LGC4GM32 - Technologies du Web (Kévin Michoud), 12H TD**

Compétences

Créer sa propre page Web avec les langages HTML et CSS.

Contenus :

Initiation à HTML (mise en forme de la page et du texte, titres, mise en valeur, listes, hyperliens, images, mise en forme de la page ; Flexbox, balises structurantes HTML5, sélecteurs et propriétés CSS2 et CSS3, types, ordre de priorité des types de sélecteurs, pseudo-classes et pseudo-éléments).

Modalités de contrôle des connaissances :

**Session 1 :** examen sur machine, durée 1h

**Session 2 :** examen sur machine, durée 1h

### **3) LGC4GM33 - XML (Amalia Todirascu), 12H TD**

Compétences à acquérir :

Maîtriser les techniques de représentation et de gestion des données en format normalisé (XML), connaître les normes XML utilisées pour la représentation des corpus.

Contenus :

Présentation générale du format XML. Architecture d'un document XML. Notions de bonne formation et de validité, outils de validation. Notion de DTD et de schéma XML. Présentation de normes XML reconnues (TEI, DocBook, XCES...).

Modalités de contrôle des connaissances :

**Session 1 :** examen sur machine 1h00

**Session 2 :** examen sur machine 1h00

#### **4) LGC4GM34 - Systèmes d'exploitation (Pablo Ruiz Fabo), 14H TD**

##### Compétences à acquérir :

- Connaître les différents types de fichiers
- Naviguer dans le système de fichiers, créer / supprimer / déplacer des dossiers et des fichiers en ligne de commande
- Connaître les principales commandes Linux utiles pour le traitement de fichiers textuels
- Savoir rechercher et modifier des chaînes de caractères : expressions régulières, description des contextes de modification et des remplacements
- Connaître les principes de la redirection et du chaînage de commandes
- Savoir écrire des scripts
- Connaître la console de Cygwin

Contenus : Linux, Windows, Commandes DOS, Commandes Linux systèmes de fichiers.

##### Modalités de contrôle des connaissances :

**Session 1 :** examen sur machine (1h)

**Session 2 :** examen sur machine (1h)

### ***UE4 Outils pour l'analyse de corpus***

#### **1) LT41GM22 – Introduction au TAL (Amalia Todirascu), 12H TD**

##### Compétences à acquérir :

- Connaître les concepts principaux du Traitement Automatique des Langues
- Maîtriser les outils d'étiquetage et d'analyse automatique de textes

##### Contenus :

Architecture d'un système en TAL : analyse et génération automatique. Ressources spécifiques pour l'analyse automatique : tokenisation, étiquetage, analyse syntaxique, analyse sémantique, analyse de discours. Approches en TAL : systèmes à base de règles, apprentissage statistique, apprentissage profond. Exemples d'applications : extraction d'information, extraction terminologique, fouille de textes.

##### Modalités de contrôle des connaissances :

**Session 1 :** dossier à rendre

**Attention ! Il n'y a pas de rattrapage pour ce cours**

#### **2) LGC4GM41 - Linguistique de corpus (Amalia Todirascu, Karim El-Haff), 24H TD**

##### Compétences à acquérir :

- Maîtriser les outils d'exploration de corpus,
- Connaître les principes de base de la constitution de corpus.

##### Contenus :

Cet enseignement aborde le traitement des documents numériques (définition, formats divers, encodage et manipulation), la définition et les critères de constitution d'un corpus, les bases textuelles et les corpus disponibles sur Internet. Dans ce cours seront présentés les outils d'exploration de corpus (concordanciers) disponibles en ligne (Frantext, FrWaC, SketchEngine), mais également indépendants (AntConc, TXM). Les mesures statistiques utilisées par les outils d'exploration de corpus seront abordés (information mutuelle, Loglikelihood). Une dernière partie du cours présente des outils d'exploitation des corpus étiquetés et annotés (TXM).

##### Modalités de contrôle des connaissances :

**Session 1 :** Contrôle continu - exercice noté ; examen final sur machine (1h)

**Session 2 :** contrôle sur machine 1h

#### **3) LT41GM23 - Outils d'annotation automatique (Amalia Todirascu), 12H TD**

##### Contenus :

Annotation automatique vs annotation manuelle. Accord inter-annotateurs. Annotations morpho-syntaxiques, syntaxiques, sémantiques: définition, enjeux, problèmes. Outils d'annotation morpho-syntaxique : étiquetage (TreeTagger, Unitex), lemmatisation (Flemm). Outils d'annotation syntaxique (analyse en dépendances par Bohnet, Talismane, ou Mind the Gap). Projets d'annotation : Multext, Universal POS, Universal Dependencies.

##### Modalités de contrôle des connaissances :

**Session 1** : dossier à rendre

**Attention ! Il n'y a pas de rattrapage pour ce cours**

## **UE5 Professionnalisation et ouverture**

### **1) LGC4GM51 - Bilan de compétences (Christophe Gérard), 8H TD**

Compétences à acquérir :

En phase avec le projet professionnel de chaque étudiant, l'objectif est de transmettre à l'étudiant une méthode concernant son bilan de compétences, mais aussi des critères concernant la rédaction du CV, la lettre de motivation et l'entretien d'embauche.

Contenus :

Cet enseignement permettra grâce à divers exercices pratiques de faire le point sur les connaissances et compétences acquises par les étudiants au cours de leur formation et de leurs expériences extra universitaires.

Modalités de contrôle des connaissances :

**Session 1** : un dossier comprenant un bilan de compétence, une analyse d'offre de stage / d'emploi, une interview réalisée avec un professionnel, un CV et une lettre de motivation, ainsi qu'une synthèse sur les éléments acquis en cours.

**Session 2** : présentation du dossier amendé qui a été déposé lors de la session 1, cette fois rédigé sur la base d'une nouvelle offre de stage d'emploi.

### **2) LGC4GM52 - Préparation du projet recherche/professionnel (Pablo Ruiz Fabo), 4H TD**

Compétences à acquérir :

- Faire une recherche documentaire sur le sujet choisi.
- Savoir identifier et éviter le plagiat.
- Gérer une bibliographie avec Zotero.

Contenus :

Le cours accompagnera la réalisation du dossier relatif au projet et traitera notamment des aspects suivants : Objectifs du projet de recherche / professionnel, Formes de communication scientifique, Constitution d'une bibliographie : techniques d'accès aux ressources électroniques, évaluation de l'information sur Internet, types de références, citation, plagiat, Respect de consignes de mise en forme (texte et bibliographie), Gestion de références bibliographiques avec Zotero, Formatage de bibliographies avec Zotero

Modalités de contrôle des connaissances :

**Session 1** : dossier à rendre.

**Session 2** : re-soumission du dossier

### **3) LGC4GX50 - Choix libre, 24H TD**

Selon le projet de l'étudiant·e, à chercher dans la liste des cours proposés à la Faculté des Langues et de Lettres.

# MASTER 1 SEMESTRE 2 (TRONC COMMUN)

## UE1 Transferts linguistiques 2

Cette UE comprend trois cours, les étudiants doivent choisir deux cours parmi les 3 proposés.

### 1) T111HM21 - Linguistique appliquée à la traduction (Christophe Gérard), 12H CM

Compétences à acquérir :

Le cours entend montrer qu'un savoir théorique peut devenir un atout de poids pour le traducteur. Contenus abordés : Traduction et linguistique : utilités et conditions d'un apport mutuel ; Les trois fossés entre les langues (structures, realia, connotations) : traduire l'équivalence lexicale ; variation linguistique et réduction au standard ; implication des genres textuels et fonctions langagières ; Qu'est-ce qu'un texte et comment y vient le sens ?

Objectif : appréhender les difficultés linguistiques qui rendent complexe tout acte de traduction, qu'il s'agisse des différences de lexique entre langues, du fait qu'une langue se constitue de différentes façons de la parler (et n'est pas un tout homogène), du fait que les genres textuels jouent un rôle central pour l'acte de traduction ou encore du fait que dans un texte la signification d'un mot choisi par un auteur dépend de celles des mots qui l'entourent.

Contenus :

Apport théorique en linguistique dans le but de fournir au futur traducteur une réflexion approfondie sur son propre travail et de prendre de la distance par rapport à la pratique.

Modalités de contrôle des connaissances :

**Session 1 :** examen écrit (1h30)

**Session 2 :** examen écrit (1h30)

### 2) T111HM22 - Institutions européennes (Patrizia Cantarelli-Volclair), 12H CM

Compétences à acquérir :

Principe de subsidiarité ; Principe de proportionnalité.

Ordre juridique communautaire ; Prise de décisions ; Finances.

Conseil européen ; Conseil ; Commission européenne ; Comité économique et social ; Comité des régions ; Parlement européen ; Cour de justice ; Cour des comptes.

Contenus :

Construction chronologique de l'Europe et étude spécifique de chacune des Institutions : fonctionnement – compétences – points faibles/forts.

Modalités de contrôle des connaissances :

**Session 1 :** devoir à rendre et un examen écrit (1 heure)

**Session 2 :** examen écrit (1 heure)

### 3) LGC4HM11 - Ressources électroniques (Gercelia Mendes ), 6H TD + travail en ligne

Compétences à acquérir :

Connaître les ressources électroniques utiles à la traduction ; connaître les techniques documentaires.

Contenus :

Dictionnaires électroniques. Ressources électroniques spécifiques (grammaires, bases lexico-sémantiques etc.). Corpus parallèles et comparables. Principes de la traduction automatique. Des exercices d'application seront proposés pendant le cours.

Modalités de contrôle des connaissances :

**Session 1 :** examen écrit, durée 1 heure

**Session 2 :** examen écrit, durée 1 heure

## UE2 Anglais et linguistique

Le cours de rédaction technique est commun aux deux options.

## **1) LGC1HM32 - Rédaction technique (Guyslaine Mylward), 18H TD**

Compétences à acquérir :

Améliorer les capacités de rédaction de documents techniques en anglais.

Contenus :

Le cours présentera les stratégies de réécriture en anglais, utilisés dans les domaines techniques. L'utilisation de glossaires et manuels d'annotation. Outils d'aide à la rédaction technique.

Modalités de contrôle des connaissances :

**Session 1** : dossier à rendre

**Session 2** : re-soumission du dossier

**Pour l'option Informatique et traduction : voir la description spécifique de l'UE2 (page 14)**

**Pour l'option Traitement Automatique des Langues : voir la description spécifique de l'UE2 (page 18)**

## **UE3 Informatique**

### **1) LGC4HM31 - Algorithmique et programmation niveau 2 (Delphine Bernhard), 24H TD**

Compétences à acquérir :

- Maîtriser les techniques de programmation avancées en Python.

- Développer des outils de traitement automatique des langues en Python.

Contenus :

Approfondissement des notions de programmation en langage Python : fonctions, tri et listes, dictionnaires et ensembles définis en compréhension, modules et packages, librairies Python pour traiter des pages web, programmation orientée objet.

Modalités de contrôle des connaissances :

**Session 1** : travail pratique noté + examen sur machine (1h) + épreuve orale (15 min)

**Session 2** : épreuve orale (20 min)

### **2) LGC4HM32 - Programmation Web (Gabriel Braun, Thomas Baguet), 18H TD**

Compétences à acquérir :

Maîtriser les techniques de programmation Web.

Contenus :

Programmation côté client de pages web dynamiques à l'aide de JavaScript. Programmation côté serveur avec PHP et Python.

Modalités de contrôle des connaissances :

**Session 1** : examen sur machine (1h).

**Session 2** : examen sur machine (1h)

### **2) LGC4HM33 - Bases de données (Pablo Ruiz Fabo), 18H TD**

Compétences à acquérir :

- Maîtriser les notions avancées sur les bases de données ;

- Maîtriser les techniques de création et de gestion d'une base de données en architecture client-serveur.

Contenus :

Modélisation de bases de données, apprentissage du langage SQL, création et manipulation de bases de données, Interface langage de script/MySQL, Applications pratiques.

Modalités de contrôle des connaissances :

**Session 1** : travail pratique noté (30 min) + examen sur machine (1h)

**Session 2** : examen sur machine (1h)

### **4) LGC4HM34 - Xpath et XSLT (Pablo Ruiz Fabo) 12H TD**

Compétences à acquérir :

Connaître le langage de programmation XSLT.

Contenus :

Introduction à XSLT. Structure d'une feuille de style. Modèle de données. Langage d'expressions défini par XPATH. Règles de transformation. Applications en corpus.

Modalités de contrôle des connaissances :

**Session 1** : travail pratique noté (30 min) + examen sur machine (1h)

**Session 2** : examen sur machine (1 h)

### **5) LGC4HM35 - Terminographie (Thierry Grass), 18H TD**

Compétences à acquérir : Maîtriser un logiciel de gestion de bases de données et les notions afférentes

Savoir constituer la terminologie d'un domaine dans une base de données.

Contenus :

Modèle conceptuel et logique des données. Les principes de la constitution de bases de données terminologiques, la fiche terminologique. Créer sa propre base terminologique à l'aide d'un logiciel de gestion de bases de données (LibreOffice Base).

Modalités de contrôle des connaissances :

**Sessions 1+2** : Constitution d'une base de données terminologique pour la traduction sur un domaine de spécialité avec LibreOffice Base.

### **6) LGC4HM36 – Analyse de données (Delphine Bernhard), 12H TD**

Compétences à acquérir : Collecter des corpus de textes à partir du web. Nettoyer, prétraiter, analyser et visualiser des données tabulées. Traiter des données textuelles pour obtenir des représentations vectorielles.

Contenus :

Collecte de données à partir du web : développement de programmes pour télécharger, analyser la structure hiérarchique et extraire le contenu textuel principal de pages HTML.

Traitement de données avec la bibliothèque pandas : nettoyage et prétraitement des données, analyse univariée, analyse bivariée, visualisation.

Vectorisation de textes : décomposition des textes en unités, normalisation et autres prétraitements, visualisations, pondération tf-idf.

Modalités de contrôle des connaissances :

**Session 1+2** : examen sur machine (1h30)

## ***UE4 Initiation à la recherche***

### **1) LGC4HM41 - Projet de recherche/professionnel (A.Todirascu, D.Bernhard, P. Ruiz Fabo)**

Compétences à acquérir :

Savoir rédiger une synthèse argumentée à partir de d'articles de recherche ou de documents professionnels (rapports, enquêtes, notes de synthèse, projets, etc.)

Contenus :

Réalisation d'un état de lieu sur un problème de recherche donné.

Modalités de contrôle des connaissances :

**Session 1** : rédaction d'une synthèse argumentée à partir de d'articles de recherche ou de documents professionnels (rapports, enquêtes, notes de synthèse, projets, etc.)

**Session 2** : re-soumission du document

## ***UE5 Professionnalisation et ouverture***

Cette unité d'enseignement est composée d'un choix libre (dans la liste des options de l'Université) ou d'un stage de découverte de 70 heures minimum, durée de 2 semaines (soit 10 jours de présence effective dans l'organisme d'accueil) à deux mois. Le stage donne lieu à la rédaction d'un rapport de stage de 10 à 15 pages (Times New Roman 12, interligne 1,5).

Compétences à acquérir pour le stage de découverte :

- S'insérer dans la vie professionnelle
- Connaître le fonctionnement de l'entreprise et/ou d'une équipe de recherche : horaires, organisation, travail en équipe

- Mobiliser et développer des compétences
  - Techniques
  - Personnelles

# MASTER 1 SEMESTRE 1 (OPTION Informatique et Traduction)

## UE2 Anglais et linguistique

Les deux cours sont obligatoires en UE2, en plus du module de compréhension orale

### 1) LGC4GM21 - Traduction spécialisée anglais-français (Christine Breyel-Steiner), 24H TD

#### Compétences à acquérir :

Savoir comprendre et analyser des textes techniques en anglais et les traduire le plus fidèlement possible vers le français.

#### Objectifs :

Acquérir une approche professionnelle de la traduction et de la méthodologie de la recherche terminologique et consolider ses qualités rédactionnelles.

#### Contenus :

Traduction de textes techniques et informatiques, représentatifs des réalités professionnelles du traducteur. Initiation à la recherche terminologique.

#### Modalités de contrôle des connaissances :

**Session 1** : une traduction durant le semestre (1h) et un examen écrit (traduction), 2h

**Session 2** : examen écrit (durée 1 heure)

### 3) LGC4GM20 - Compréhension et expression écrites anglais-français / CEE (Christine Breyel-Steiner), 18H TD

#### Compétences à acquérir :

Savoir traduire vers le français ou résumer en français des textes complexes rédigés en langue étrangère en tenant compte de notions telles que le type de texte, le public visé, les registres, les constructions idiomatiques, la synonymie etc.

#### Objectifs :

Apprendre à traduire en se détachant du texte source et en maîtrisant les techniques de rédaction.

#### Contenus :

Traduction de textes issus de domaines tels que l'environnement, le marketing, la publicité... ; analyse et correction de textes déjà traduits.

#### Modalités de contrôle des connaissances :

**Session 1** : examen écrit (durée 1H)

**Session 2** : examen écrit (durée 1H)

# MASTER 1 SEMESTRE 2 (OPTION Informatique et Traduction)

## UE2 Anglais et linguistique

L'option comporte trois modules :

- Anglais : compréhension et expression orales (18h TD)
- Anglais : rédaction technique (24h TD)
- 1 option obligatoire au choix parmi les matières proposées en Master Sciences du Langage en UE 4 (24h CM) :

**ATTENTION ! il n'y a pas d'examen de rattrapage pour ces matières !**

### a) LT41HM40 : Linguistique et phonétique clinique M. SOCK/Mme VAXELAIRE

Descriptif et bibliographie communiqués à la rentrée

### b) LT41HM41 : Linguistique romane ou diachronique Mme Capin : Le changement linguistique

Dans ce cours, nous étudierons les phénomènes de changement linguistique à travers différents types de changement dans les langues romanes, en diachronie et en synchronie contemporaine. Nous passerons en revue à travers cela les différents mécanismes de changement, les origines et causes du changement, ainsi que la méthodologie de recherche en linguistique diachronique

Références:

- Giacalone Ramat, A. & Hopper, P. (dir.), *The limits of grammaticalization*, Amsterdam, Benjamins.
- Givón, T. (1979), *Understanding Grammar*, New York, Academic Press.
- Haspelmath, M., et al. (dir.), (2001), *Language Typology and Language Universals*, Berlin, Mouton De Gruyter.
- Hickey, R. (dir.) (2010), *Motives for language change*, Cambridge University Press.
- Marchello-Nizia, C. (2006), *Grammaticalisation et changement linguistique*, De Boeck.
- Nørgård-Sorensen, J., Heltoft, L. & Schøsler, L., (2011), *Connecting Grammaticalisation*, Amsterdam, Benjamins.
- Stein, D., et al., (1995), *Subjectivity and Subjectivation*, Cambridge University Press.
- Traugott, E., & Heine, B. (dir.), *Approaches to grammaticalisation*, Amsterdam, John Benjamins, vol.1
- Traugott, E., & Dasher, R., (2002), *Regularity in Semantic Change*, Cambridge, Cambridge University Press

### c) L41HM43 : Lexicologie Mme Gerhard-Krait, Mme Vassiliadou

Les étudiants seront capables de :

- Distinguer les principes classificatoires en usage dans les typologies en usage
- Avoir un regard critique sur les classifications existantes
- Identifier les principes de formations du lexique RSE
- Identifier les principes de formations des néologismes
- Connaître les grandes sous-catégories en vigueur pour les catégories lexicales étudiées
- Dissocier différentes sous-catégories lexicales sur la base de critères morphologiques, syntaxiques et sémantiques
- Savoir exploiter d'autres bases de données lexicales pour en extraire et analyser des unités lexicales
- Savoir exploiter les données collectées dans des bases de données pour des analyses quantitatives et qualitatives
- Déterminer les usages et fonctions des unités lexicales dans les discours et suivant les genres discursifs

# MASTER 1 SEMESTRE 1 (OPTION Traitement Automatique des Langues)

## UE2 Anglais et linguistique

En plus du cours de compréhension orale en anglais, les étudiants doivent choisir 3 matières au choix en Master Sciences du Langage en UE 3 ou 4 (Phonétique, Morphologie, Syntaxe, Sémantique, Diachronie)

### 1) Linguistique, 36H TD

Pour le descriptif des modules, voir le guide pédagogique du Master Sciences du Langage

*Vous devrez choisir 3 matières parmi celles proposées.*

**ATTENTION ! il n'y a pas d'examen de rattrapage pour ces matières !**

#### a) LT41GM30 – Principaux courants en phonétique et phonologie (M. SOCK)

Ce séminaire permet d'apprécier différentes réactions aux écoles classiques en phonétique et en phonologie, et les diverses tentatives qui ont été faites pour revisiter ces approches.

Après avoir exposé quelques débats qui ont eu lieu depuis de nombreuses années en phonétique et en phonologie, et après avoir discuté plusieurs notions et concepts controversés dans le domaine, le séminaire tâchera de montrer une certaine convergence d'idées et de paradigmes, issus de divers travaux de recherche, aboutissant ainsi à de grands courants théoriques en phonétique et en phonologie.

#### Compétences acquises

savoir procéder à des analyses critiques de certains modèles (théoriques, conceptuels, physiques, statistiques, mathématiques...) de phonétique et de phonologie pour pouvoir confronter l'« ancien » au « nouveau », et illustrer le processus sous-jacent à l'émergence de ces théories phares dans le domaine de la production et de la perception de la parole.

#### b) LT41GM31 – Interface morphosyntaxe et sémantique (Mme TODIRASCU)

L'objectif de ce cours est de présenter la notion du figement en linguistique. Il s'agit du processus qui permet de transformer les séquences libres en expressions polylexicales figées et semi-figées. Le figement est un critère important dans la classification des expressions polylexicales variées (expressions idiomatiques, collocations, locutions etc.). L'étude de ce phénomène se fait en faisant appel aux tests linguistiques mais aussi aux outils statistiques. L'étude des expressions polylexicales et de leur degré de figement se fera dans une perspective contrastive.

#### c) LT41GM32 – Syntaxe : linéarisation de l'énoncé (Mme KUYUMCUYAN)

L'objectif du cours est d'étudier le rapport entre les fonctions grammaticales dans la proposition et l'organisation linéaire de l'énoncé, ainsi que les facteurs de perturbation de l'ordre réputé canonique SVO. Cette problématique conduit à interroger les organisations « déviantes » de l'énoncé simple en français : dislocation, thématization et autres procédés de (dé)focalisation. On abordera également l'organisation de la phrase averbale par le biais de la prédication.

#### Bibliographie

Blasco M. (1999), Les dislocations en français contemporain, Paris, Champion.

Delaveau A., Syntaxe – La phrase et la subordination, Armand Colin, 2001.

Groupe de Fribourg Grammaire de la période, P. Lang 2012

Le Goffic P., Grammaire de la phrase française, Hachette, 1994.

#### d) LT41GM33 – Sémantique (Mme Theissen)

Programme annoncé à la rentrée

#### e) LT41GM41 – Linguistique romane ou diachronique (D.Capin)

Pré-requis : connaître :

- la périodisation du français
- les catégories grammaticales

#### OBJECTIFS

Apprendre :

- les caractéristiques les plus importantes pour chaque type de changement étudié. Reconnaître les cadres théoriques associés à chaque type de changement ;
- apprendre à comparer les données dans des langues différentes et faire une synthèse ;
- apprendre à reconnaître les régularités dans le changement afin de mieux connaître la faculté du langage ;
- apprendre à trouver les universaux ;
- apprendre à décrire les diversités langagières.

Etre capable de retracer l'histoire d'un terme, d'une construction

Approfondir et élargir les connaissances linguistiques dans les domaines de la morphologie, la syntaxe et la sémantique en diachronie et en synchronie

Comprendre : 1) le changements linguistique ; 2) un énoncé dans une langue différente ; 3) comprendre et faire le compte-rendu d'un article scientifique dédié à la variation et au changement linguistique.

#### **f) LT41GM40 –Linguistique et phonétique cliniques (Mme VAXELAIRE)**

Ce séminaire aborde les apports de la phonétique à l'étude des pathologies de la voix et de la parole. La phonétique s'occupe de l'étude des Ce séminaire commence par rappeler et consolider des notions de phonétique clinique déjà évoquées en L3. Il traite des apports de la phonétique à l'étude des pathologies de la voix et de la parole. La phonétique s'occupe de l'étude des sons de la parole. La clinique se préoccupe, elle, des déviations dues à des pathologies, par rapport à la normalité phonétique et phonologique.

Des pathologies sont examinées dans le cadre conceptuel de la perturbation et des réajustements en production et en perception de la parole.

##### **Compétences acquises**

- savoir décrire et étudier, sur les plans théorique et expérimental, diverses pathologies, en relation avec des notions phonétiques et phonologiques.
- Les étudiants sont accueillis au sein du laboratoire de l'Institut de Phonétique de Strasbourg (IPS) pour y conduire des expérimentations.

Une bibliographie et des documents seront présentés lors des premières séances du séminaire et tout au long du semestre.

#### **g) LT41GM34 –Linguistique romane ou diachronique (T.Revol)**

Contenu du cours :

Études philologiques à partir de textes médiévaux en italien ou en espagnol (éventuellement d'autres langues romanes). Les textes sont fournis en cours.

Méthodes :

- recherches des caractéristiques linguistiques des textes : graphies, phonétique, morphologie, syntaxe, stylistique
- perspective historique : études diachroniques, du latin à la langue du texte et à la situation actuelle
- perspective contrastive : comparaison avec le français et les autres langues romanes

Bibliographie sommaire :

Pierre Bec, Manuel pratique de philologie romane (t. 1), Paris, Picard, 1970.

Jack Schmidely, Manuel Alvar Ezquerra, Carmen Hernández González, De una a cuatro lenguas. Del español al portugués, al italiano y al francés, Madrid, Arcos / Libros S. L., 2001.

Compétences à acquérir dans l'UE2 :

Maîtrise de notions théoriques en linguistique dans les domaines de la linguistique synchronique ou diachronique, de la phonétique ou de la sémantique (en fonction des options choisies).

Contenus :

Linguistique synchrone ou diachronique, phonétique, sémantique selon l'option choisie.

Modalités de contrôle des connaissances : NB ces matières sont évaluées en contrôle continu selon le guide du master Sciences du langage.

**Session unique :** selon le ROF

#### **h) LT41GM43– Lexicologie (Mme TODIRASCU, M Meyer, Mme LAMMERT)**

Le lexique RSE Sur la base d'un corpus commun de discours RSE (entreprises de mode), il s'agira d'analyser

les pratiques langagières exploitées dans ces discours tout en fournissant une méthode d'analyse du lexique. Les thèmes abordés au semestre 1 seront les suivants:

- les noms de marques et de produits : tests linguistiques, annotations
- notions de sémantique distributionnelle
- les noms collectifs et les noms collectifs humains.

Les étudiants seront capables de :

- Distinguer les principes classificatoires en usage dans les typologies lexicales
- Avoir un regard critique sur les classifications existantes
- Identifier les principes de formations des formes lexicales étudiées
- Connaître les grandes sous-catégories en vigueur pour les formes lexicales étudiées
- Dissocier différentes sous-catégories de formes lexicales sur la base de critères morphologiques, syntaxiques et sémantiques
- Connaître le maniement de certaines bases de données
- Savoir exploiter d'autres bases de données lexicales pour en extraire et analyser les unités lexicales étudiées
- maîtriser des outils d'extraction d'expressions polylexicales

Modalités de contrôle des connaissances : NB ces matières sont évaluées en contrôle continu selon le guide du master Sciences du langage.

**Session unique** : selon le ROF

**ATTENTION ! il n'y a pas d'examen de rattrapage pour ces matières !**

# MASTER 1 SEMESTRE 2 (OPTION Traitement automatique des langues)

## UE2 Anglais et linguistique

En plus du cours de rédaction technique en anglais, les étudiants doivent choisir 3 matières au choix en Master Sciences du Langage en UE 3 ou 4 (Phonétique, Morphologie, Syntaxe, Sémantique, Diachronie)

**ATTENTION ! il n'y a pas d'examen de rattrapage pour ces matières !**

### 1) Sciences du langage, 36H TD

Pour le descriptif des modules, voir le guide pédagogique du Master Sciences du Langage . 3 matières à choisir parmi:

#### UE3

#### **LT41HM31 : Morphologie : la nominalisation Mme GERHARD**

La formation des noms dits construits, qu'ils soient nouvellement insérés dans le lexique ou qu'ils soient installés depuis plus ou moins longtemps, répond à des règles, que les locuteurs ont intériorisées, et qui leur permettent globalement de construire ou d'interpréter des formes lexicales à partir des éléments qui les composent.

Ce cours ciblera exclusivement la construction des noms dits déverbaux (sortie, construction, téléchargement, etc.) avec pour objectif de discuter de la possibilité que les éléments qui entrent dans la structure formelle et sémantique d'un mot construit puisse prédire son fonctionnement sémantique et syntaxique et qu'elle puisse s'appréhender de manière systématique. La problématique sera abordée au prisme de différents points de vue théoriques.

#### Bibliographie

- BERRENDONNER A., 1995, Redoublement actantiel et nominalisations, *Scolia* 5, 215-244.
- CHOMSKY N., 1970, Remarks on Nominalizations, in R. Jacobs et P. Rosenbaum (eds), *Readings in English Transformational Grammar*, Waltham (Mass.), Ginn-Blaisdell.
- CORBIN D. (1987), *Morphologie dérivationnelle et structuration du lexique*, Tübingen, Niemeyer.
- DUBOIS J., 1969, *la Phrase et ses transformations*, Paris, Larousse.
- FRADIN B., 2003, *La morphologie morphématique combinatoire. Présentation et critique*, Nouvelles approches en morphologie, sous la direction de Fradin Bernard. Presses Universitaires de France, 27-78, (<https://www.cairn.info/nouvelles-approches-en-morphologie>).
- HONESTE M.-L., 2007, « Approche cognitive de la notion de "forme nominalisée du verbe" en français », *Faits de Langues*, 59-68, .
- HUYGHE R., 2015, Les nominalisations "d'achèvement graduel" en français. *Le Français Moderne*, (<https://hal-univ-diderot.archives-ouvertes.fr/hal-01244487>).
- HUYGHE R., JUGNET A., 2010, Nominalisations statives et transfert aspectuel: quel héritage sémantique?, *Congrès Mondial de Linguistique Française 2010*, 2010, États-Unis. 1713-1726. hal- 00975635, (<https://hal-univ-diderot.archives-ouvertes.fr/hal-00975635>).
- LEFEUVRE F., 2014, Les nominalisations prédicatives en discours. *Le Français Moderne*, <halshs-01519597>.
- MAGRI-MOURGUES V., 2014, La nominalisation. Fait de langue ou fait de style ?, in Laure Himy-Piéri, Jean-Francois Castille, Laurence Bougaul (éds), *Le style, découpeur de réel*, (éds), Presses universitaires de Rennes, 125-137, (<http://www.openedition.org/6540>).
- Lexique* 20, 2012, *Nominalisations : nouveaux aspects*, Presses du Septentrion.

#### **LT41HM32 : Syntaxe : enchâssement/emboîtement Mme KUYUMCUYAN**

L'objectif du cours est de revenir sur la notion de phrase enchâssée. Après un rappel des diverses définitions du phénomène, et une présentation comparée des diverses classifications proposées dans les grammaires les plus courantes, on procèdera à un recensement critique de certaines structures, clivées et pseudo-clivées notamment, pour montrer que la subordination ne se limite pas à l'enchâssement, et peut servir d'autres fins dans la langue.

#### Bibliographie

Combettes B. 2002 Aspects de la grammaticalisation de la phrase complexe en ancien et en moyen français, *Verbum* 24, 109-128.

Delaveau A. 2001, *Syntaxe – La phrase et la subordination*, Armand Colin.

Bonnard H. 2001, *Les trois logiques de la grammaire française*, Bruxelles, Duculot.

Groupe de Fribourg 2012, *Grammaire de la période*, P. Lang

Le Goffic P. 1994, *Grammaire de la phrase française*, Hachette.

Muller C. 1996, *La subordination en français*, Paris, A. Colin.

Roubaud M.-N. 2000, *Les constructions pseudo-clivées en français contemporain*, Paris, Champion.

### **LT41HM33 : Sémantique : L'adverbe Mme THEISSEN**

La catégorie adverbiale est, par son hétérogénéité, une classe complexe où se côtoient les problèmes morphologiques, sémantiques et syntaxiques. Le cours aura pour objectif de présenter, d'une part, les difficultés soulevées par cette classe qualifiée parfois de « fourre-tout » ou « classe reliquat » et, d'autre part, d'examiner les réponses qui ont été apportées.

Bibliographie

Molinier C. & Levrier F., 2000, *Grammaire des adverbes : Description des formes en –ment*, Droz, Genève.

Nøkke H., éd., 1990, *Classification des adverbes*, Langue française 88.

Nøjgaard M., 1992, *Les adverbes français. Essai de description fonctionnelle*, *Historisk-filosofiske Meddelelser* 66 1 (1992), 2 (1993), 3 (1995), Copenhagen, Munksgaard.

### **LT41HM34 : Linguistique romane ou diachronique Mme CAPIN**

Philologie linguistique galloromane

Le cours s'attachera à montrer les spécificités des pratiques éditoriales en usage en linguistique historique et les techniques d'analyse élaborées par la philologie linguistique (scriptologie, glossographie, modèles textuels, utilisation d'outils informatiques). Il prendra appui sur les textes galloromans médiévaux (français, occitan, francoprovençal), en déterminant l'espace, les périodes et les genres textuels concernés.

Le cours comportera des travaux pratiques sur des textes notamment documentaires, qui permettent d'appréhender de manière particulièrement immédiate autant l'élaboration éditoriale d'un texte médiéval que les différentes opérations d'analyse.

UE4

### **LT41HM40 Linguistique et phonétique clinique M. SOCK/Mme VAXELAIRE**

### **LT41HM43 Lexicologie : le lexique des discours RSE LT41LM43 – Lexicologie**

Intervenantes : Mme Gerhard, Mme Vassiliadou

Sur la base d'un corpus de discours RSE (entreprises de la mode), il s'agira d'étudier divers aspects de la langue pratiquée dans ce genre discursif tout en initiant à une méthodologie lexicologique.

Modalités de contrôle de connaissances : conforme MECC du master SDL

### **LT41HM41 : Linguistique romane ou diachronique Mme Capin**

Le changement linguistique. Processus de changement en linguistique (romane, germanique et slave) : perspectives synchroniques et diachroniques

Descriptif du cours

Le cours présente les différents types de changements (analogie, réanalyse, grammaticalisation vs dégrammaticalisation, lexicalisation, pragmatification, subjectification, transcatégorisation) en appliquant une grille de critères qui permettent de les différencier. Les exemples pour illustrer chaque type de changement sont pris dans les corpus de langues romanes, germaniques et slaves. Sont recensés les principaux changements dans le domaine nominal et verbal en diachronie et en synchronie.

Références:

Bibliographie

COMBETTES B., (2003), « Introduction : grammaticalisation et changement linguistique », *Verbum*, XXV/ 3,

25-40.

- COMBETTES B. (dir.) (2003), Évolution et variation en français préclassique. Études de syntaxe. Champion.
- COMBETTES B, MARCHELLO-NIZIA C. et PREVOST S. (coord.) (2003), Grammaticalisation en français, Verbum XXV N° 4
- GIACALONE RAMAT A. et HOPPER P. (eds), The limits of grammaticalization, Amsterdam, Benjamins.
- HARRIS M. & RAMAT P., (eds) (1987), Historical Development of Auxiliaries, Berlin, NY, Amsterdam, Mouton de Gruyter.
- HASPELMATH M., et al. (eds), (2001), Language Typology and Language Universals, Berlin, Mouton De Gruyter.
- HICKEY R. (éd) (2010), Motives for language change, Cambridge, Cambridge University Press.
- HEINE B. et KUTEVA T. (2002), Word Lexicon of grammaticalisation, Cambridge University Press.
- HEINE B., CLAUDI U et HÜNNEMEYER F. (1991), Grammaticalization : a conceptual framework, Chicago, Chicago University Press.
- HOPPER P. et TRAUOGOTT E., (2003), Grammaticalisation, Cambridge University Press.
- STEIN D. et al.,(1995), Subjectivity and Subjectivation, Cambridge University Press.
- TRAUGOTT E., HEINE B. (eds), Approaches to grammaticalisation, Amsterdam, Jonh Benjamins, vol.1
- TRAUGOTT E., DASHER R., (2002), Regularity in Semantic Change, Cambridge, Cambridge University Press.

Modalités de contrôle des connaissances :

**Session unique** : selon les enseignements choisis, à vérifier dans le guide du master Sciences du langage.

**ATTENTION ! il n'y a pas d'examen de rattrapage pour ces matières !**

# MASTER 2 SEMESTRE 3 (TRONC COMMUN)

## UE1 Initiation à la recherche

### 1) LT41GM10 - Séminaire de recherche, 6H TD

#### Compétences à acquérir :

Connaître le fonctionnement d'une équipe de recherche ainsi que les sujets traités.

#### Contenus :

Option TAL : participation aux séminaires de recherche de l'équipe LiLPa (UR 1339). Dates disponibles sur les pages suivantes : <http://lilpa.unistra.fr/seminaires/>

Option Informatique et traduction :

Participation aux séminaires de recherche de l'équipe LiLPa (UR 1339). Dates disponibles sur les pages suivantes : <http://lilpa.unistra.fr/seminaires/>

Participation aux séminaires des mardis de l'ITIRI (dates à venir, les mardis 18h-20h)

Session 1 : satisfecit. Chaque participation devra être justifiée par la signature de l'attestation de présence. Cette signature devra être demandée en fin de séance à l'organisateur ou à l'organisatrice de l'événement.

### 2) LGC4KM11 - Méthodologie de la recherche (Delphine Bernhard), 6H TD

#### Compétences à acquérir :

Savoir répondre à une offre de stage et connaître les éléments fondamentaux pour la réussite du stage.

Savoir présenter des travaux de recherche.

#### Contenus :

Préparation au stage : rechercher un stage, répondre à une offre de stage, savoir se présenter à l'entretien, rédiger un cahier de charges, connaître le déroulement d'un stage, rédiger un rapport de stage.

Présentation d'un travail de recherche à l'oral et à l'écrit. Méthodologie d'écriture du mémoire.

#### Modalités de contrôle des connaissances :

**Session 1** : exposé oral du projet de recherche (15 minutes)

**Session 2** : refaire l'exposé.

## UE2 Informatique

### 1) LGC4KM21 - Algorithmique et programmation niveau 3 (Delphine Bernhard), 24H TD

#### Compétences à acquérir :

Connaître les notions de base de la programmation orientée objet avec Java

#### Contenus :

Programmation orientée objet. Langage Java. Classes, objets. Héritage, polymorphisme.

#### Modalités de contrôle des connaissances :

**Session 1** : travail pratique noté, épreuve orale (15min)

**Session 2** : épreuve orale (20min)

### 2) LGC4KM22 - Génie logiciel et gestion de projets (Kévin Michoud), 12H TD

#### Contenus :

Modélisation d'un projet informatique à l'aide de diagrammes UML.

- Gestion de projet :
  - Décomposition d'un projet en tâches et sous-tâches.
  - Planification du déroulement des tâches (diagramme de Gantt).
  - Méthodes agiles
- Utilisation d'outils de gestion de versions

#### Modalités de contrôle des connaissances :

**Session 1** : exercices notés

**Session 2** : exercice noté

### **3) LGC4KM23 - Fouille de textes et apprentissage automatique (Delphine Bernhard), 30H TD**

Compétences à acquérir :

- Préparer les données pour l'apprentissage : guide d'annotation, annotation manuelle des données, transformation du format des données.
- Appliquer des algorithmes d'apprentissage automatique.
- Sélectionner le meilleur algorithme d'apprentissage en fonction de la tâche.
- Évaluer les résultats de l'apprentissage de manière quantitative et qualitative.

Contenus :

Techniques d'apprentissage automatique, supervisé et non supervisé. Prétraitement des données textuelles pour l'apprentissage. Principaux algorithmes d'apprentissage. Évaluation des résultats. Initiation à Weka et aux bibliothèques Python pour l'apprentissage automatique.

Modalités de contrôle des connaissances :

**Session 1** : travail pratique noté, épreuve orale (15min)

**Session 2** : épreuve orale (20min)

### **4) LGC4KM24 - Méthodes statistiques pour l'analyse de données textuelles (Pablo Ruiz Fabo), 12H TD**

Compétences à acquérir :

Maîtriser les techniques d'analyse statistique de corpus ; Connaître les bases du langage R

Contenus :

Analyse quantitative de corpus textuels à l'aide de méthodes statistiques. Initiation au langage R.

Modalités de contrôle des connaissances :

**Session 1** : exercices notés

**Session 2** : exercices notés

### **5) LGC4KM25 - Projet tutoré (Pablo Ruiz Fabo), 6H TP**

Compétences à acquérir :

Mobiliser des connaissances acquises dans les autres matières.

Contenus :

Le projet tutoré est proposé sur une durée de 7 semaines. L'objectif est de réaliser une mini-application dans le domaine du traitement automatique des langues ou de la fouille de textes. Les étudiants pourront ainsi mobiliser des compétences acquises en programmation orientée objet, en génie logiciel et en apprentissage automatique.

Modalités de contrôle des connaissances :

**Session 1** : dossier à rendre

**Session 2** : re-soumission du dossier

## ***UE3 Communication scientifique***

Cette UE contient deux matières obligatoires pour tous les étudiants.

### **1) LGB2KM41 - Rédaction d'articles en anglais (Ellen Carter), 12H TD**

Compétences à acquérir :

Apprendre le style et le registre de l'anglais académique et scientifique, apprendre les techniques de structuration de l'article scientifique (abstract, introduction, keywords, findings and conclusion) et les techniques de documentation (citing and in-text referencing, using foreign language sources, footnotes and endnotes, composing a list of works cited).

Objectifs :

L'objectif de ce module est de simuler la rédaction d'un article scientifique dans le domaine du TAL ou de la traduction, ou un document de synthèse accompagnant un projet professionnel précis.

Contenus :

Introduction à l'écriture d'articles scientifiques en anglais. Introduction aux styles de citation et référencement

spécifiques (MLA, APA, Chicago, AMS, etc.).

Modalités de contrôle des connaissances :

**Session 1 + 2 :** rédaction en anglais d'une question de recherche. L'article doit comprendre entre 1200 et 1500 mots et doit correspondre aux exigences de l'écriture scientifique et académique anglaise (choix, utilisation et documentation des sources, structure, style et langue). L'article doit porter sur un domaine de recherche choisi par l'étudiant et convenu d'avance avec l'enseignant. La qualité de la langue ainsi que le raisonnement scientifique devront correspondre aux normes de publication internationale.

## **2) LGB2KM42 - Expression scientifique orale en anglais (Sandrine Baudry), 12H TD**

Compétences à acquérir :

Savoir effectuer une présentation orale sur un sujet défini. Rédiger et mettre en forme une présentation argumentée.

Contenus :

Présentation d'une communication avec support PAO. Apprentissage des techniques de rédaction d'un exposé académique.

Modalités de contrôle des connaissances :

**Session 1 :** Présentation orale scientifique en anglais (15 min) + débat (5 min) + participation au cours du semestre

**Session 2 :** Présentation orale scientifique en anglais (20 min)

## ***UE4 Traitement automatique des Langues et Traduction***

### **1) LT41KM20 - Annotation de haut niveau (Amalia Todirascu) 12H TD**

Contenus :

Annotation sémantique. Exemples d'annotation sémantique : les entités nommées. CasSys. Annotation discursive. Plateformes d'annotation (Glozz, Inception, SACR pour les chaînes de référence <http://boberle.com/projects/sacr/>, TXM).

Modalités de contrôle des connaissances :

**Session 1 :** Dossier à rendre

**Attention, il n'y a pas de rattrapage pour ce cours**

### **2) LGC4KM41 - Traduction automatique (Pablo Ruiz Fabo), 18H TD**

Compétences à acquérir :

- Maîtriser les outils d'alignement de corpus
- Maîtriser les outils d'extraction de ressources linguistiques à partir de corpus alignés
- Maîtriser les outils d'exploration de corpus parallèles et comparables
- Développer un modèle de traduction automatique neuronale

Contenus :

- Histoire de la traduction automatique et paradigmes : À base de règles, statistique et neuronale.
- Évaluation : mesures manuelles et automatiques, intrinsèques et extrinsèques
- Corpus alignés et corpus comparables.
  - Alignement de phrases et de segments
  - Extraction d'un corpus aligné à partir de sites web bi- ou multilingues
- Extraction de ressources linguistiques lexicales ou grammaticales à partir de corpus comparables et alignés.
  - Extraction de lexiques bilingues
  - Extraction de paraphrases
- Traduction statistique (SMT)
  - Systèmes à base de segments (PBMT)
  - Systèmes factorisés
  - Pratique avec l'outil Moses
- Deep learning et réseaux de neurones pour la traduction automatique
  - Introduction au TAL neuronal
  - Introduction à la traduction neuronale

- Modèles encodeur-décodeur : sequence to sequence (*seq2seq*) avec attention
  - Exemples avec plusieurs bibliothèques : Keras, TensorFlow, OpenNMT
  - Le modèle *Transformer*

Modalités de contrôle des connaissances :

**Session 1** : dossier à rendre

**Session 2** : re-soumission du dossier

### 3) LGC4KM42 - Recherche d'information (Pablo Ruiz Fabo) 18H TD

Compétences à acquérir :

- Connaître le mode de fonctionnement des moteurs de recherche
- Développer un moteur de recherche pour données textuelles en utilisant des bibliothèques existantes
- Connaître les principes de base du référencement naturel (SEO)

Contenus :

- Historique des moteurs de recherche.
- Mode de fonctionnement (crawling, pré-traitement et indexation des documents, traitement des requêtes utilisateur)
- Paradigmes de recherche :
  - Recherche booléenne
  - Modèles vectoriels : pondération des termes, calculs de similarité entre requête et document
  - Modèles probabilistes (notions)
  - Evaluation en recherche d'information
- Création d'un moteur de recherche avec la librairie Solr
  - Configuration de l'index
    - Options linguistiques pour le français (lemmatisation)
  - Indexation
  - Affichage des résultats de recherche
- Enrichissement de la recherche avec des métadonnées obtenues par le TAL (entités nommées, motifs lexicaux, ...)
- Exploitation d'annotations obtenues par le TAL pour la recherche par facettes dans Solr
- Recherche web
  - Référencement naturel (SEO) et critères de positionnement.
  - Algorithme PageRank
  - Chatbots et agents conversationnels :
    - À base de règles, à base de corpus, agents conversationnels à base de frames
    - Pratique d'entraînement d'un chatbot

Modalités de contrôle des connaissances :

**Session 1** : dossier à rendre

**Session 2** : re-soumission du dossier

### 4) LGC4KM43 - Dictionnaires électroniques (Beatriz Sanchez), 6H CM + 12H TD

Compétences à acquérir :

Savoir constituer un dossier complet en terminologie pour un donneur d'ordre, en utilisant les corpus et des outils d'extraction terminologique.

Contenus :

Ce séminaire, de 12 heures est structuré en deux parties, présentiel et à distance, de 6 heures chacune.

Séances en présentiel (6 heures) :

Contenus théoriques

- Notions théoriques sur la lexicographie
- Ressources lexicographiques en ligne
- Conception mono-fonctionnelle des dictionnaires électroniques
- Compilation de corpus spécialisés comparables
- Extraction terminologique des termes simples et complexes
- Identification des collocations, structures phraséologiques.
- La prosodie sémantique
- Identification de structures sémantiques avec CQL

- Conception d'un dictionnaire électronique: la macrostructure
- Description des unités lexicales: la microstructure

Séances à distance (6 heures) :

Les étudiants travailleront sur un projet personnel guidé pour créer un dictionnaire électronique en mettant en pratique les concepts et méthodes abordés pendant les séances en présentiel.

Modalités de contrôle des connaissances :

**Session 1** : constitution d'un dossier.

**Session 2** : re-soumission du dossier.

## **UE5 Ouverture**

Liste communiquée à la rentrée. Il faut suivre un cours de 24 h ou 2 cours de 12h parmi la liste des options SDL (Faculté de Lettres) ou proposées à la Faculté de langues :

Options proposés en Sciences du langage

### **1) LT41KM33 – Sémantique référentielle Mme Todirascu (12H)**

Ce cours a comme objectif de présenter les chaînes de référence, leurs propriétés et leur fonctionnement dans plusieurs genres discursifs. On aborde les problèmes liés à l'annotation, manuelle et automatique des chaînes de référence.

### **2) LT41KM43 – Lexicologie Mme TODIRASCU, M Meyer, Mme LAMMERT (24H)**

Le lexique RSE

Sur la base d'un corpus commun de discours RSE (entreprises de mode), il s'agira d'analyser les pratiques langagières exploitées dans ces discours tout en fournissant une méthode d'analyse du lexique.

Les thèmes abordés au semestre 1 seront les suivants:

- les noms de marque et de produits (tests linguistiques, annotation manuelle et automatique)
- la sémantique distributionnelle
- les noms collectifs et les noms collectifs humains

### **3) LT41KM30 – Parole et cognition M. SOCK (12h)**

L'objectif de ce séminaire est de montrer que la production de la parole est un processus cognitif et moteur complexe, régi par une orchestration temporelle des articulateurs, organisés dans le temps et dans l'espace, pour *contribuer* à l'*émergence* des catégories sensori-motrices *viables* : les sons/phonèmes d'une langue spécifique. En mettant l'accent sur la dimension motrice du phénomène parole, l'on se rapproche, avec quelques ajustements épistémologiques, de la démarche de STETSON, qui disait dans son ouvrage *Motor Phonetics* (1928 ; 1951 ; rééd. 1988) que la parole était plutôt un ensemble de gestes rendus audibles qu'un ensemble de sons produits par des mouvements.

### **4) LT41KM31 – Morphologie Mme GERHARD-KRAIT (12H)**

### **5) LT41KM32 – Syntaxe Mme KUYUMCUYAN/Mme THEISSEN (12H)**

Pour les options en Langues, consulter le site Web de la faculté des Langues.

## **MASTER 2 SEMESTRE 4 (TRONC COMMUN)**

Le quatrième semestre a lieu en stage et comprend deux modules.

## ***UE1 Mémoire de recherche***

### **LGC4LM11 – Mémoire de recherche**

Le mémoire de recherche constitue une synthèse argumentée sur le sujet choisi en première année et les travaux effectués durant les deux années de master (80 pages : police Times new roman 12 / Arial 11, interligne simple 1,15). Davantage de détails à ce sujet sont donnés dans le cours LGC4KM11 - Méthodologie de la recherche.

Modalités de contrôle des connaissances :

Mémoire de recherche et soutenance (présentation de 20 minutes devant un jury avec diaporama)

## ***UE2 Stage***

### **LGC4LM21 – Stage**

Le stage peut se faire dans une entreprise spécialisée dans le domaine du TAL, dans l'équipe d'accueil Lilpa ou dans un laboratoire de recherche universitaire, selon les modalités stipulées par la loi (convention de stage).

Compétences à acquérir :

- S'insérer dans la vie professionnelle
- Connaître le fonctionnement de l'entreprise et/ou d'une équipe de recherche : horaires, organisation, travail en équipe
- Construire et réaliser un projet du début à la fin
- Mobiliser et développer des compétences
  - Techniques
  - Personnelles

Modalités de contrôle des connaissances :

**Session 1** : Rapport de stage d'environ 30 pages (Davantage de détails à ce sujet sont donnés dans le cours LGC4KM11 - Méthodologie de la recherche)

**Session 2** : Refaire un stage ou re-soumettre le rapport de stage

## Stage : Mode d'emploi 2023-2024

Il est obligatoire d'avoir défini toutes les modalités pédagogiques du stage avant de faire une demande de convention de stage en ligne (voir la liste des étapes à suivre).

Prenez contact avec vos enseignants avant le début du stage, que le stage soit obligatoire ou facultatif.

Il est obligatoire d'être inscrit administrativement pour 2023-2024.

### Contact pour les stages des étudiants de la Faculté des langues

Monsieur Stéphane Trompette, pôle scolarité langues, à [strompet@unistra.fr](mailto:strompet@unistra.fr) / 03 68 85 66 89

### A quoi sert une convention de stage ?

La convention de stage tripartite définit les conditions du stage et, pour l'essentiel, les garanties dont bénéficie l'étudiant. Elle couvre l'étudiant en cas d'accident.

### Processus de demande de convention

La demande de convention de stage en ligne doit être faite au minimum :

- 1 mois avant le début du stage pour un stage en France
- 2 mois avant le début du stage pour un stage à l'étranger.

**Attention ! Si vous avez déjà commencé le stage, la convention de stage ne pourra pas en tenir compte.** La convention de stage ne sera **datée qu'à partir du moment où la demande de convention de stage aura été faite** à votre scolarité via l'ENT. Votre scolarité **ne délivre pas de convention de stage lorsque le stage est terminé.**

### Comment créer une demande de convention de stage ?

Vous devez faire la demande en ligne via votre Environnement Numérique de Travail (ENT) sur le site <https://ent.unistra.fr/>

Après identification, sélectionnez l'onglet « Mon dossier » puis l'onglet « Demande de convention de stage ».

### Éléments à collecter avant la saisie de sa demande en ligne

À votre sujet :

- Le diplôme pour lequel le stage est réalisé
- Votre numéro de sécurité sociale
- Votre type d'affiliation à la sécurité sociale (en qualité d'ayant droit d'assuré social, en qualité d'étudiant, par une assurance volontaire ou en qualité d'étudiant étranger)
- Votre caisse d'assurance maladie (Caisse Primaire d'Assurance Maladie (LMDE, MGEL,..) ou autre Caisse)

Au sujet de votre organisme d'accueil :

- Raison sociale (nom de l'organisme d'accueil)
- Adresse complète de votre organisme
- Numéro SIRET à 14 chiffres pour les stages en France (info : <http://www.manageo.fr/>)

- Code NAF/APE à 4 chiffres + 1 lettre (secteur d'activité de l'entreprise) pour les stages en France
- Effectif (nombre de salariés)
- Type d'établissement
- Le nom du service qui vous accueille

**Au sujet de votre tuteur professionnel dans l'organisme d'accueil :**

- Nom, Prénom, Fonction
- Service de rattachement
- Coordonnées (téléphonique, mail ou télécopie)

**Au sujet de votre stage :**

- Type de stage (obligatoire ou facultatif)
- Thématique du stage
- Sujet et fonctions du stage
- Dates de stage Attention ! Le stage doit être terminé le 30 septembre 2020 au plus tard
- Période d'interruption - Présence éventuelle le samedi, le dimanche, la nuit et les jours fériés
- Temps de présence (temps partiel ou plein)
- Durée du stage en heures et nombre de jours de travail hebdomadaire(6 mois maximum dans la même entreprise)
- Choix de la convention de stage : Convention de stage en français, Convention de stage en franco-anglais, Convention de stage en franco-allemand, Convention de stage en franco-espagnol, Convention de stage formation continue
- Montant horaire net de la gratification :

**Au sujet de l'enseignant référent à l'université :**

- Nom, Prénom

**Au sujet du signataire de la convention de stage dans l'organisme d'accueil :**

- Nom, Prénom, Fonction (PDG, RH, Gérant, Président, Inspecteur d'académie pour l'enseignement primaire public...)
- Coordonnées (téléphonique, mail ou télécopie)

Fin de saisie et envoi de votre demande de convention de stage : lorsque vous êtes à l'Etape 8 - Récapitulatif de la convention, un message vous informe que « votre convention a bien été créée ». Après vérification des données et de la législation, votre scolarité valide votre demande de convention de stage et vous envoie à votre adresse mail UNISTRA la convention de stage en format pdf ainsi que la Charte des stages dans les meilleurs délais.

Vous vérifiez la convention de stage (**pour tout changement ou erreur dans la convention de stage saisie, ne pas créer une nouvelle convention de stage mais nous contacter par mail ou téléphone**), vous l'imprimez en **3 exemplaires**, vous les signez et les faites signer par **l'enseignant référent de votre diplôme et par le représentant de l'organisme d'accueil** (les signatures et cachets sont obligatoires).

Vous déposez les 3 originaux ou les envoyez à votre scolarité accompagnées des pièces suivantes :

- Votre assurance responsabilité civile couvrant la période du stage
- Une grande enveloppe affranchie pour 100g (1, 55 euros) libellée votre adresse pour le retour des conventions de stages signées par l'Université de Strasbourg.

Votre scolarité fait signer les conventions par le Président de l'Université de Strasbourg, vous envoie 2 exemplaires signés avec la fiche d'évaluation et l'attestation de stage.

Pour tout avenant à votre convention de stage initiale (interruption, prolongation...), vous devez faire la demande via votre ENT ET nous contacter.

## Étapes à suivre pour la mise en place d'une convention de stage

**IMPORTANT :** Toutes ces étapes doivent être réalisées **dans l'ordre** et en respectant les **délais minimum** avant le début du stage. Dans le cas contraire, le début du stage pourra être reporté ou la signature de la convention pourra être refusée.

N°	Description	Délai minimum	Fait le (date)
1	Une fois votre candidature pour un stage acceptée, remplir la fiche de projet de stage et l'envoyer à l'enseignant-e référent-e (encadrant-e du mémoire). L'enseignant-e référent-e validera ou non votre projet de stage en fonction de l'adéquation entre les missions du stage, les objectifs de la formation et vos objectifs professionnels. Si vous avez des doutes sur l'adéquation d'une offre de stage, nous vous conseillons vivement de contacter votre encadrant-e avant même de candidater à l'offre.	6 semaines (stage en France) 10 semaines (stage à l'étranger)	
2	Si le projet de stage est validé, demander la convention de stage via l'ENT (voir le guide pédagogique pour davantage de détails)	1 mois (stage en France)	
3	Une fois la convention validée par le Pôle Sclolarité Langues, elle est envoyée au format PDF sur votre adresse Unistra. Imprimer la convention en 3 exemplaires et les signer	2 mois (stage à l'étranger)	
4	Faire signer les trois exemplaires de la convention par les personnes suivantes dans l'organisme d'accueil : <ul style="list-style-type: none"> <li>représentant-e de l'organisme d'accueil</li> <li>tuteur-riche dans ce même organisme</li> </ul>		
5	Faire signer les trois exemplaires de la convention par son enseignant-e référent-e au sein de la Faculté des Langues		
6	Déposer les trois exemplaires signés avec l'attestation d'assurance responsabilité civile au Pôle Sclolarité Langues. Le Pôle les transmet pour signature à la doyenne.		
7	Réception des exemplaires signés Un mail sur votre adresse Unistra vous prévient que vous pouvez retirer votre convention signée. Possibilité de retrait en personne, ou envoi par voie postale (voir conditions dans le guide pédagogique) ou demande d'envoi de copie scannée de la convention par mail à strompet@unistra.fr.	Tous les exemplaires doivent être signés <b>AVANT</b> le début du stage	