

# Master Sciences Humaines et Sociales

## Mention : Traitement automatique des langues

### Parcours : Recherche et développement

Présentation	
<b>Nature</b> Formation diplômante Diplôme national  <b>Niveau de recrutement</b> Bac + 3, Bac + 4 <a href="#">Conditions d'admission</a>  <b>Durée de la formation</b> 2/4 semestres  <b>Lieu(x) de formation</b> Campus de Nanterre, Autres sites  <b>Stages</b> Non	<b>Accessible en</b> <ul style="list-style-type: none"><li>• <a href="#">Formation initiale</a></li><li>• <a href="#">Formation continue</a></li></ul> <b>Modes d'enseignement</b> En présentiel

Le Master de Traitement Automatique des Langues prépare des étudiants issus des sciences humaines aussi bien que des sciences exactes à être opérationnels dans le secteur du traitement automatique de l'information et l'ingénierie linguistique. Les compétences acquises sont à la fois **linguistiques, statistiques et informatiques**.

Le master TAL-R&D est commun à 3 établissements : l'Université Paris Nanterre, l'Université Sorbonne Nouvelle Paris 3, l'INALCO.

La recherche et le développement dans le domaine du TAL et des industries de la langue font appel à des méthodes et des savoirs multiples. Il s'agit de :

- **Disposer** des méthodes d'analyse linguistique : maîtriser les manipulations débouchant sur des descriptions détaillées de faits de langue, connaître les fondements des différents domaines des sciences du langage (phonétique et phonologie, morphologie, syntaxe et sémantique) ;
- **Connaître** les bases de la recherche et de l'extraction d'information, de la constitution et de la gestion de corpus (écrits ou oraux) et de ressources, y compris multilingues : les corpus sont des mines d'information pour une description réaliste d'emplois d'une langue, les techniques de la recherche et de l'extraction d'information permettent de rapatrier les documents ou les parties de documents jugés pertinents pour un besoin de recherche particulier ;
- **Exprimer** les règles et les régularités à l'œuvre dans les corpus, par le biais des grammaires formelles et des traitements quantitatifs pour savoir passer d'une description linguistique d'un texte à une représentation plus formelle permettant sa prise en charge par des logiciels.

## Spécificités

▣ **Structuration générale de la Mention Traitement automatique des langues :**

**En M1** : Un parcours unique Traitement automatique des langues vous offre différents choix en vue de la spécialisation de M2

**En M2** : Vous pouvez, au sein de la mention, vous spécialiser dans un des 2 parcours suivants :

- 1/ « Recherches et développement (R&D) »
- 2/ « [Ingénierie linguistique](#) - DEFI »

## Admission

### Publics

**Niveau de recrutement** : Bac + 3, Bac + 4

### Conditions d'admission

- **Vous êtes étudiant(e) en formation initiale :**

Vous trouverez les conditions d'admissions, modalités et critères de recrutement par formations, dans l'encadré "Candidatures 2019-2020", en bas à droite. Sélectionnez M1 ou M2 suivant votre niveau puis l'UFR et l'intitulé de votre formation. Retrouvez les dates d'inscription sur e-Candidat : <https://ecandidat.parisnanterre.fr/#!accueilView> (menu à gauche "Offre de formation", sélectionnez la formation).

- **Vous êtes un(e) adulte en reprise d'études (formation continue) :**

L'équipe du SFC est à votre disposition pour vous aider à :

1/ **identifier les différents dispositifs de validation** qui vous permettent d'entrer dans cette formation ou d'obtenir ce diplôme (VAE/VES) (cf. supra, rubrique Admission).  
2/ **connaître les multiples dispositifs de financement** pour concrétiser votre projet (DIF/CPF, CIF, Contrat de professionnalisation, Période de professionnalisation, Plan de formation, Congé de formation professionnelle, Congé pour VAE, chèque VAE, etc.).

3/ **monter un dossier de financement et obtenir des devis.** Pour en savoir plus sur ces procédures, rendez-vous sur le site du [SFC](#), rubrique : Contactez-nous pour remplir le [formulaire de contact](#)

- **Les dispositifs de validation d'acquis à connaître :**

Que vous soyez étudiant(e), salarié(e) ou demandeur d'emploi en formation continue, plusieurs dispositifs de validation vous permettent :

**1/ d'entrer dans cette formation sans avoir le diplôme de niveau inférieur :**

- Validation des Acquis Académiques (VAA)
- Validation des études, expériences professionnelles ou acquis personnels (anciennement VAPP)

**2/ d'obtenir ce diplôme sans pour autant reprendre des études**

- Validation des Acquis de l'Expérience (VAE)
- Validation des Etudes Supérieures (VES)

[Pour en savoir plus sur ces dispositifs...](#)

## Programme

**Vous trouverez ci-dessous la version abrégée de la formation. Le livret pédagogique est accessible dans l'encadré de droite "Supports pédagogiques formation tout au long de la vie"**

### ▾ 1ère ANNEE - Semestre 1

#### **UE TAL et ingénierie linguistique 1** (15 ects)

- Gestion informatique du multilinguisme
- Informatique et phonétique
- Programmation et projet encadré 1
- Modélisation linguistique pour l'analyse automatique de texte
- Grammaire formelle

#### **UE Document numérique et informatique 1** (6 ects)

- Base de données
- Programmation et algorithmique 1

### **UE Linguistique 1** (9 ects)

- Syntaxe formelle
- Machine creativity and text generation
- un cours au choix

### ▣ **1ère ANNEE - Semestre 2**

#### **UE TAL et ingénierie linguistique 2** (12 ects)

- Statistique et analyse multidimensionnelle
- Programmation et projet encadré 2
- Outil de traitement de corpus
- Corpus parallèles et comparables

#### **UE Document numérique et informatique 2** (9 ects)

- Recherche et extraction d'information
- Document structuré
- Programmation et algorithmique 2

#### **UE Linguistique 2** (6 ects)

- Introduction à la fouille de textes
- un cours au choix

#### **UE Langue** (3 ects)

### ▣ **2ème ANNEE - Semestre 3**

#### **UE Ingénierie en TAL (18 ects)**

*6 enseignements à prendre parmi ceux qui suivent (ou d'autres à choisir en accord avec le directeur de recherche)*

##### **Site Université Paris Nanterre**

- Modélisation des langues : mots et grammaire
- TAL et ingénierie des connaissances
- Document structuré et écriture numérique
- Base de données et Web dynamique
- Programmation et programmation orientée objet
- Méthodologie de la Recherche. Epistémologie du TAL
- Corpus annotés et développement de ressources linguistiques
- Linguistique outillée et traitements statistiques
- Langages du Web sémantique

##### **Site Paris 3**

- Analyse du discours et lexicométrie
- Fouille de textes

##### **Site INALCO**

- Acquisition, modélisation et représentation des connaissances
- Multimodalité du document numérique
- Sémantique des textes multilingues
- Genres, textes et usages
- Lexicologie, terminologie, dictionnairique

##### **Site Paris 7**

- Analyse de données pour le TAL
- Analyse syntaxique automatique
- Analyse sémantique automatique
- Analyse et génération de discours

#### **UE Linguistique (9 ects)**

- 2 ou 3 enseignements de linguistique à prendre en accord avec le directeur de recherche

#### **UE Langue ( 3 ects)**

### ▣ **2ème ANNEE - Semestre 4**

#### **UE Stage (9 ects)**

- Stage en laboratoire ou en entreprise

### **UE Mémoire (15 ects)**

- Mémoire de recherche

### **UE Méthodologie de la recherche (6 ects)**

- Méthodologie de la recherche

Débouchés

### **Année post-bac de sortie**

Bac + 5

## **Débouchés professionnels**

### **S e c t e u r s d ' a c t i v i t é :**

Le secteur du traitement automatique des langues comprend aussi bien les grands acteurs du web et du logiciel, qu'un grand nombre de PME qui développent des outils personnalisés pour les entreprises et les particuliers.

### **M é t i e r s :**

chercheur (établissement public ou secteur R&D) // enseignant-chercheur // ingénieur linguiste (établissement privé ou public ou projets internationaux) // informaticien spécialisé dans le traitement des langues // gestionnaire de ressources textuelles (courrier électronique, documentation) // veilleur (économique, stratégique, technologique) par analyse de documents textuels // développeur de ressources linguistiques (ontologies, mémoires de traduction, dictionnaires, treebanks) // développeur d'outils d'aide aux handicapés

Mis à jour le 25 avril 2019

## **► UFR de rattachement**

[Philosophie Information-Communication Langage Littérature Arts du Spectacle](#)

Paris Nanterre

LLIA

Ricoeur,

ue de la République

NANTERRE

0 97 56 13

0 97 47 11

[et](#)

## **► Supports pédagogiques formation tout au long de la vie**

[pédagogique](#)

[RNCP](#)

[des formations éligibles au CPF](#)

[de la formation](#)

[charger la fiche formation](#)

## ► Renseignements pratiques

Secrétariat

MANCHI

(L, bureau R06)

[parisnanterre.fr](http://parisnanterre.fr)

Formation (SCUIO-IP)

[e.parisnanterre.fr](http://e.parisnanterre.fr)

Bureau E.14

[parisnanterre.fr](http://parisnanterre.fr)

Formation continue(SFC)

[formation-continue@liste.parisnanterre.fr](mailto:formation-continue@liste.parisnanterre.fr)

Tel : 01 40 97 78 66

Bâtiment M

Site : [formation-continue.parisnanterre.fr](http://formation-continue.parisnanterre.fr)

## ► CANDIDATURES 2019-2020

[Candidatures en M1](#)

[Candidatures en M2](#)

## ► ATTENTION

Le présent document correspond à notre ancienne offre de formation pour la période 2019-2020.

La nouvelle offre de formation pour l'année 2020/2021 est consultable dès à présent [sur notre site dédié](#)

**Offre de  
formation  
2020-2021**



▶ Documents à télécharger

[- LICENCE](#)

[- MASTER](#)

[Graduate Courses - LICENCES](#)

[Graduate Courses - MASTERS](#)

[terre.fr/master-sciences-humaines-et-sociales-br-mention-traitement-automatique-des-langues](http://terre.fr/master-sciences-humaines-et-sociales-br-mention-traitement-automatique-des-langues)