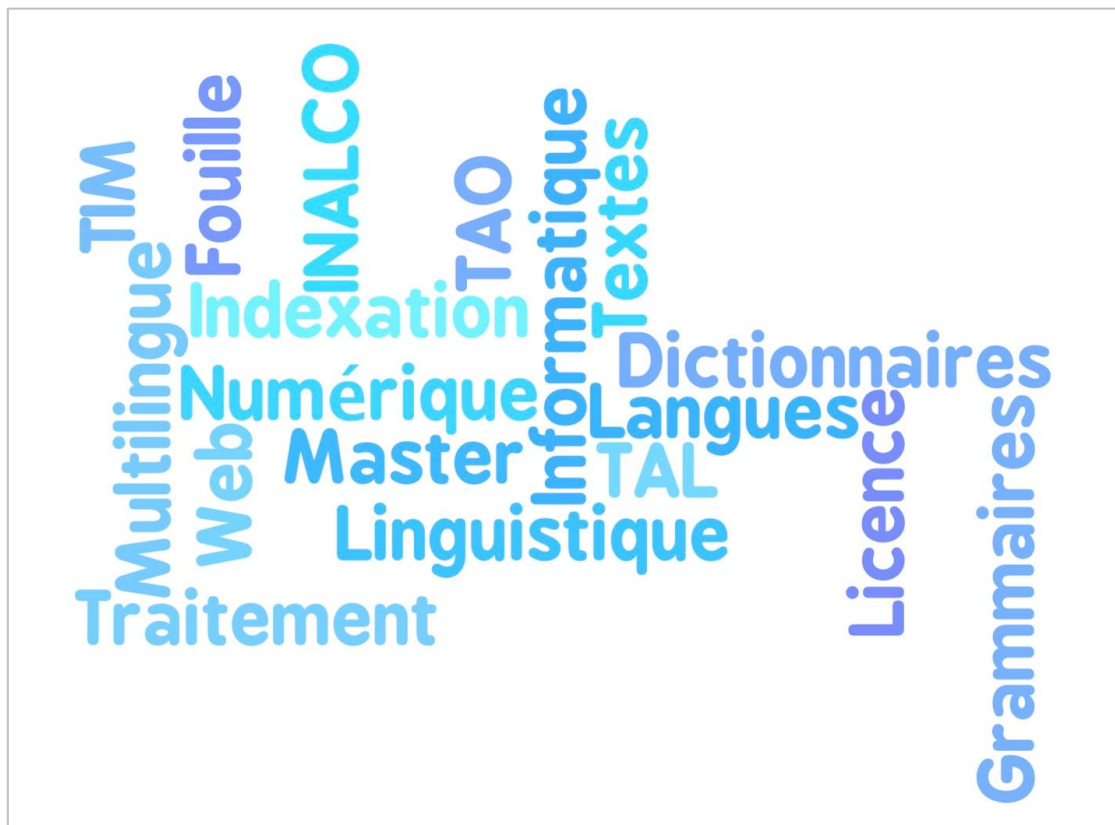


# Master Traitement Automatique des Langues

Parcours Ingénierie Multilingue

Parcours Technologie de la Traduction et  
Traitement des Données Multilingues

Parcours Recherche & Développement



## Sommaire

Présentation générale la filière TIM.....	2
Contacts .....	2
Objectif du Master .....	3
Public visé .....	3
Conditions d'admission :.....	3
Débouchés professionnels.....	3
Les enseignements.....	5
Master 1 traitement automatique des langues.....	6
M2 - parcours Ingénierie multilingue .....	7
M2 - parcours Technologie de la Traduction et Traitement des Données Multilingues .....	8
M2 - parcours Recherche et développement .....	9



## Présentation générale la filière TIM

L'offre de formation du département TIM (Textes, Informatique, Multilinguisme) a pour objectif de former les étudiants en langue aux métiers de la société de l'information demandant de fortes compétences linguistiques.

À cette fin, cette formation leur apporte les connaissances théoriques et méthodes pratiques pour le Traitement Automatique des Langues (TAL), domaine fondamentalement pluridisciplinaire (langue(s), linguistique, informatique théorique et pratique, etc.).

Le traitement automatique des langues (TAL) est de manière générale le domaine recouvrant toutes les applications informatiques ayant trait aux langues et aux textes et qui suppose la modélisation de phénomènes langagiers afin de doter les systèmes informatiques de capacités à analyser ou générer des textes dans des langues variées. L'ingénierie linguistique est l'ensemble des produits et services qui ont le TAL pour objet. La spécificité de cette formation est de positionner le TAL dans un contexte multilingue. Sont enseignées de manière approfondie les compétences pour le traitement de volumineux corpus dans des langues diverses, parfois rares, à tous les niveaux (encodage, analyses morphologiques, syntaxiques, sémantiques etc.).

### Contacts

#### Filière Textes, Informatique, Multilinguisme

Directeur de la filière : M. Mathieu VALETTE – [mvalette@inalco.fr](mailto:mvalette@inalco.fr)

Responsable de formation : M. Bassir HAMID – [bassir.hamid@inalco.fr](mailto:bassir.hamid@inalco.fr)

Gestionnaire pédagogique : Mme Lucie TAILLARD – [lucie.taillard@inalco.fr](mailto:lucie.taillard@inalco.fr)

Bureau 3.25A

65, rue des Grands Moulins

75013 Paris

01 81 70 11 36

## Objectif du Master

L'objectif du Master est de former des spécialistes capables d'appréhender un domaine par essence pluridisciplinaire (langue, linguistique, informatique), tant dans un cadre recherche qu'industriel. A l'image de l'Inalco, le Master a une forte composante multilingue seule à même de répondre, entre autres, aux nouveaux besoins créés par la convergence de phénomènes tels que la mondialisation, l'élargissement de l'accès au web, la définition de normes informatiques pour toutes les aires linguistiques, la croissance de l'échange de documents électroniques.

## Public visé

Le Master appuie sa démarche de formation sur les acquis en langues des étudiants. Il s'adresse

- à des étudiants alliant des connaissances significatives en informatique et souhaitant valoriser des compétences en langue dans la perspective de travailler dans le domaine de l'ingénierie linguistique,
- à des étudiants en langues ou en linguistique n'ayant pas obligatoirement de compétences techniques mais désireux d'acquérir une spécialisation favorisant leur intégration dans le monde du travail, en particulier dans les métiers de la traduction, de l'édition et de la gestion de contenus multilingues.

## Conditions d'admission :

Les étudiants sont admis sur dossier.

Les étudiants doivent justifier d'un niveau suffisant dans une langue enseignée à l'Inalco et faire preuve de compétences en linguistique ou en informatique. Les étudiants étrangers doivent en outre justifier d'une connaissance suffisante du français.

## Débouchés professionnels

Les métiers que suscitent les industries de la langue sont multiples. Des possibilités s'offrent dans les entreprises spécialisées ou ayant des secteurs spécialisés dans le développement d'outils TAL (Xerox, IBM, Microsoft, SINEQUA SA, France Telecom, SYSTRAN, TEMIS, LexiQuest, ...) pour la conception et la maintenance de logiciels. Ces

mêmes entreprises et bien d'autres (petites et moyennes, organismes nationaux et internationaux), pour les services qu'elles proposent ou pour leurs besoins propres, peuvent également faire appel à des spécialistes dans les domaines de la production, de l'organisation et de la gestion de l'information.

Les métiers sont essentiellement des métiers de niveau ingénieur: traducteur et responsable de projet de traduction, rédacteur de documentations techniques monolingues et/ou multilingues, rédacteur en lexicographie et terminographie pour une utilisation traditionnelle ou électronique ou dans des applications TAL, intervenant dans la constitution, l'organisation, la gestion de bases de données monolingues et/ou multilingues utilisables dans de nombreuses applications, expert pour l'amélioration de la couverture linguistique d'applications existantes, participant tant du point de vue linguistique qu'informatique à la conception et au développement d'outils d'ingénierie linguistique et de sites Web : TAO, e-learning, recherche et filtrage d'informations, documentation automatique, outils bureautiques évolués, outils de veille technologique, analyse des réseaux sociaux, fouille de textes (fouille d'opinion, analyse de sentiment, détection des émotions, extraction d'information (terminologies, noms propres, noms de marque, nouveaux métiers de la traduction : postédition.

## Les enseignements

Le Master TAL est organisé en 3 parcours, avec un tronc commun en M1.

- « Recherche et Développement »
- « Ingénierie Multilingue »
- « Technologie de la Traduction et Traitement des Données Multilingues »

Le parcours « Recherche et Développement » est cohabilité avec l'Université Sorbonne Nouvelle – Paris 3 et avec l'Université Paris Ouest Nanterre La Défense.

Les tableaux ci-après décrivent l'organisation des enseignements et la répartition des ECTS.



## Master 1 traitement automatique des langues

Semestre 1		30 ECTS
<b>LANGUES VIVANTES ÉTRANGÈRES - 1 EC</b>		<b>4</b>
Anglais ou une langue de l'inalco (présentiel ou mooc)	Inalco	
<b>LINGUISTIQUE</b>		<b>7</b>
Pratiques textuelles et traduction	Inalco	
Lexique et morphologie	Inalco	
Phonétique et synthèse de la parole	Paris3	
<b>MODÈLES ET FORMALISMES</b>		<b>6</b>
Langages réguliers	Inalco	
Principes des bases de données	Inalco	
<b>INFORMATIQUE</b>		<b>7</b>
Corpus arborés et parsing	Paris Nanterre	
Programmation objet 1	Inalco	
Langages de script	Inalco	
<b>INGÉNIERIE</b>		<b>6</b>
Gestion informatique du multilinguisme	Inalco	
Programmation et Projet encadré 1	Paris3 / Inalco	

Semestre 2		30
<b>LANGUES VIVANTES ÉTRANGÈRES - 1EC</b>		<b>4</b>
Anglais ou une langue de l'inalco (présentiel ou mooc)	Inalco	
<b>LINGUISTIQUE</b>		<b>5</b>
Traduction automatique et assistée	Inalco	
Sémantique lexicale et textuelle	Inalco	
<b>MODÈLES ET FORMALISMES</b>		<b>5</b>
Documents structurés	Paris2	
<b>INFORMATIQUE</b>		<b>5</b>
Programmation objet 2	Inalco	
Outils de traitement de corpus	Inalco	
<b>INGÉNIERIE</b>		<b>5</b>
Statistiques textuelles	Paris3	
Corpus parallèles et comparables	Inalco	
Extraction d'informations	Inalco	
<b>PROJET</b>		<b>6</b>
Programmation et Projet encadré 2	Paris3 / Inalco	

## M2 - parcours Ingénierie multilingue

<b>Semestre 1</b>	<b>30</b>
<b>LINGUISTIQUE</b>	<b>6 ECTS</b>
Sémantique des textes multilingues 1 Lexicologie, terminologie et dictionnaire 1 Genres, textes et usages	
<b>INGÉNIERIE</b>	<b>9 ECTS</b>
Traitement statistique de corpus Fouille de textes 1 Réseau de neurones pour la reconnaissance de l'oral et applications linguistiques	
<b>INFORMATIQUE</b>	<b>6 ECTS</b>
Programmation objet 1 Langages de script	
<b>MODÈLES ET FORMALISMES</b>	<b>9 ECTS</b>
Documents structurés Calculabilité Acquisition, modélisation et représentation des connaissances	

<b>Semestre 2</b>	<b>30</b>
<b>LINGUISTIQUE</b>	<b>3 ECTS</b>
Sémantique des textes multilingues 2 Lexicologie, terminologie et dictionnaire 2	
<b>INGÉNIERIE</b>	<b>6 ECTS</b>
Fouille de textes 2 Technique web Web, BD, sites multilingues et localisation Méthodes en apprentissage automatique Ingénierie des connaissances Outils de traitement de corpus	
<b>INFORMATIQUE</b>	<b>3 ECTS</b>
Programmation objet 2 Programmation itérative et récursive	
<b>STAGE ET MÉMOIRE</b>	<b>18 ECTS</b>



## M2 - parcours Technologie de la Traduction et Traitement des Données Multilingues

<b>Semestre 1</b>	<b>30</b>
<b>LINGUISTIQUE</b>	<b>6 ECTS</b>
Sémantique des textes multilingues 1 Lexicologie, terminologie et dictionnaire 1 Genres, textes et usages	
<b>INGÉNIERIE</b>	<b>9 ECTS</b>
Traitement statistique de corpus Fouille de textes 1 Outils de TAO 1 Écriture et multilinguisme	
<b>TRADUCTION TECHNIQUE</b>	<b>9 ECTS</b>
Traduction technique 1 Traductologie 1 Conduite de projet de traduction 1	
<b>INFORMATIQUE</b>	<b>3 ECTS</b>
Base de données pour le web Langages de script	
<b>MODÈLES ET FORMALISMES</b>	<b>3 ECTS</b>
Documents structurés Acquisition, modélisation et représentation des connaissances	

<b>Semestre 2</b>	<b>30</b>
<b>LINGUISTIQUE</b>	<b>3 ECTS</b>
Sémantique des textes multilingues 2 Lexicologie, terminologie et dictionnaire 2	
<b>INGÉNIERIE</b>	<b>3 ECTS</b>
Fouille de textes 2 Outils de TAO 2 Web, BD, sites multilingues et localisation Ingénierie des connaissances	
<b>TRADUCTION TECHNIQUE</b>	<b>6 ECTS</b>
Traduction technique 2 Traductologie 2 Conduite de projet de traduction 2	
<b>STAGE ET MÉMOIRE</b>	<b>18 ECTS</b>

## M2 - parcours Recherche et développement

<b>Semestre 1</b>		<b>30 ECTS</b>
<b>LINGUISTIQUE</b>		<b>9 ECTS</b>
Modélisation des langues (Paris Nanterre)		3
<b>TALA512x (2 EC au choix)</b>		
Fonctionnement et dysfonctionnement prosodique		3
Phonétique comparée des langues : introduction à la recherche		3
<b>INGÉNIERIE en TAL (5 EC au choix, 1 EC obligatoire)</b>		<b>18 ECTS</b>
TAL et ingénierie des connaissances		3
Document structuré et écriture numérique		3
Base de données et Web dynamique		3
Programmation et programmation orientée objet		3
<b>Méthodologie de la Recherche. Epistémologie du TAL (obligatoire)</b>	Paris X	3
Corpus annotés et développement de ressources linguistiques		3
Linguistique outillée et traitements statistiques		3
Langages du Web sémantique		3
Fouille de textes		3
Expérimentation et modalisation dans les humanités numériques	Paris 3	3
Syntaxe neuronale		3
Sémantique des textes multilingues		3
Lexicologie, terminologie, dictionnaire		3
Genres, textes et usages		3
Traitements statistiques de corpus	Inalco	3
Fouille de textes		3
Réseau de neurones pour la reconnaissance de l'oral et applications linguistiques		3
<b>LANGUE</b>		<b>3 ECTS</b>
Langue vivante	Inalco	3
<b>Semestre 2</b>		<b>30 ECTS</b>
STAGE (en laboratoire ou en entreprise)		9
MÉMOIRE de Recherche		15
MÉTHODOLOGIE DE LA RECHERCHE		6

