

# Master

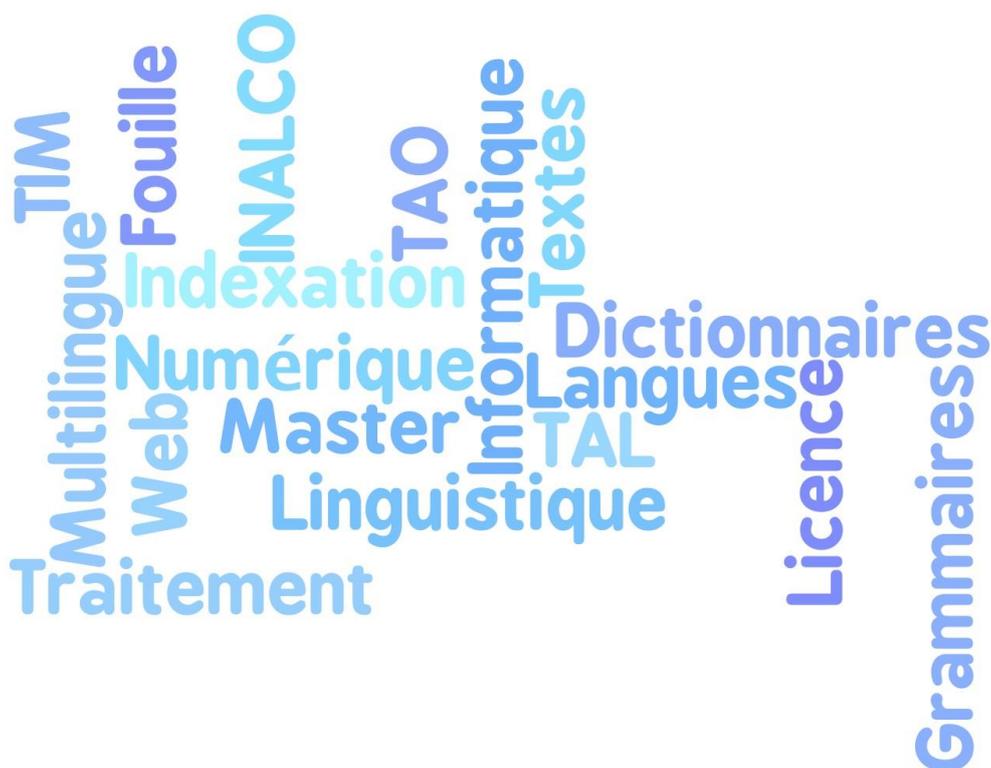
## Traitement Automatique des Langues

Parcours Ingénierie Multilingue

Parcours Technologies de la Traduction et

Traitement des Données Multilingues

Parcours Recherche & Développement



Département Textes, Informatique, Multilinguisme



Institut national  
des langues  
et civilisations orientales

Document non contractuel,  
sous réserve de modification

## Sommaire

Présentation générale du département .....	2
Contacts.....	2
Objectifs de la formation du Master .....	2
Public visé .....	2
Conditions d'admission : .....	2
Débouchés professionnels .....	2
Les enseignements .....	3
Master Traitement Automatique des Langues – M1 .....	4
Master Traitement Automatique des Langues – M2 - parcours Ingénierie multilingue .....	6
Master Traitement Automatique des Langues – M2 - parcours Traductique et gestion de l'information .....	8
Master Traitement Automatique des Langues – M2 - parcours Recherche et développement .....	10



A word cloud containing the following terms: Corpus, Algorithmique, Logique, Automates, C++, Indexation, Programmation, and Syntaxe. The words are arranged in a roughly cross-like shape, with 'Corpus' and 'Algorithmique' at the top, 'Automates' on the left, 'C++' and 'Indexation' at the bottom, and 'Programmation' and 'Syntaxe' on the right.



A word cloud containing the following terms: XML, Lexiques, Textes, Documents, Morphologie, Sémantique, and HTML. The words are arranged in a roughly cross-like shape, with 'XML' and 'Lexiques' at the top, 'Textes' and 'Documents' on the left, 'Morphologie' and 'Sémantique' on the right, and 'HTML' at the bottom.

## Présentation générale du département

L'offre de formation du département TIM (LLCER, parcours Traitement Numérique Multilingue, Master Traitement Automatique des Langues) 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 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

### Département Textes, Informatique, Multilinguisme

Renseignements pédagogiques : Jean-Michel Daube  
[jmdaube@inalco.fr](mailto:jmdaube@inalco.fr) / 01 81 70 10 36

Secrétariat : Mme Amandine Polliart  
bureau 32



65, rue des Grands Moulins  
75013 Paris



01 81 70 11 36

### Équipe de recherche ERTIM



2, rue de Lille  
75007 Paris

[www.er-tim.fr](http://www.er-tim.fr)

[ertim@inalco.fr](mailto:ertim@inalco.fr)

## Objectifs de la formation 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, son établissement d'accueil, 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 :

Sont admis de plein droit

- les étudiants titulaires de la Licence LLCER, parcours Traitement Numérique Multilingue).

Sont admis sur dossier :

- les étudiants titulaires d'une licence de TAL ou de linguistique informatique délivrée par un autre établissement et
- les étudiants provenant d'autres filières ou d'autres formations et justifiant d'un niveau suffisant dans une langue enseignée à l'INALCO et faisant 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 »
- « Traductive et Gestion de l'information »

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 Traitement Automatique des Langues – M1

### Semestre 1

<b>TAL4A01</b>		<b>COMPÉTENCES AVANCÉES EN LANGUES</b>		<b>6 ECTS</b>	
	LTP4A02J	Langues de l'inalco anglais	26		
<b>TAL4A02</b>		<b>LINGUISTIQUE</b>		<b>6 ECTS</b>	
	LTE4A02E	Pratiques textuelles et traduction	26		
	LTP4A02A	Lexique et morphologie	26		
	LTE4A02C	Informatique et phonétique	26		
<b>TAL4A03</b>		<b>MODÈLES ET FORMALISMES</b>		<b>6 ECTS</b>	
	LTP4A02K	Langages réguliers	26		
	LTP4A02F	Syntaxes formelles	26		
	LTP4A02E	Principes des bases de données	39		
<b>TAL4A04</b>		<b>INFORMATIQUE</b>		<b>6 ECTS</b>	
	LTE4A02D	Bases de données pour linguistes	26		
	LTP4A02G	Programmation objet 1	39		
	LTP4A02H	Langages de script	26		
<b>TAL4A05</b>		<b>INGÉNIERIE</b>		<b>6 ECTS</b>	
	LTE4A02B	Gestion informatique du multilinguisme	48		
	LTP4A02I	Projet encadré			
<b>Total semestre 1</b>			<b>334</b>		

**Semestre 2**

<b>TAL4B01</b>	<b>COMPÉTENCES AVANCÉES EN LANGUES</b>			<b>3 ECTS</b>
	LTP4B02J	Langues de l'inalco anglais	26	
<b>TAL4B02</b>	<b>LINGUISTIQUE</b>			<b>3 ECTS</b>
	LTE4B02F	Techniques de traduction	19,5	
	LTE4B02A	Sémantique lexicale et textuelle	39	
<b>TAL4B03</b>	<b>MODÈLES ET FORMALISMES</b>			<b>6 ECTS</b>
	LTE4B02E	Documents structurés XML	26	
	LTP4B02K	Langages hors contexte	26	
<b>TAL4B04</b>	<b>INFORMATIQUE</b>			<b>6 ECTS</b>
	LTP4B02G	Programmation objet 2	39	
	LTP4B02F	Outils de traitement de corpus	19,5	
<b>TAL4B05</b>	<b>INGÉNIERIE</b>			<b>6 ECTS</b>
	LTE4B02B	Statistiques et analyse de corpus	26	
	LTE4B02C	Corpus parallèles et comparables	19,5	
	LTE4B02D	Extraction d'informations	26	
<b>TAL4B06</b>	<b>PROJET</b>			<b>6 ECTS</b>
	LTP4B02I	Projet encadré 2	39	
<b>Total semestre 2</b>			<b>305,5</b>	<b>30</b>

## Master Traitement Automatique des Langues – M2 - parcours Ingénierie multilingue

### Semestre 1

<b>TIM5A02</b>	<b>LINGUISTIQUE</b>			<b>6 ECTS</b>
	LTE5A02A	<b>Sémantique des textes multilingues 1</b>	<b>26</b>	
	LTP5A02A	<b>Lexicologie, terminologie et dictionnaire 1</b>	<b>26</b>	
	LTP5A02B	<b>Genres, textes et usages</b>	<b>26</b>	
<b>TIM5A03</b>	<b>MODÈLES ET FORMALISMES</b>			<b>9 ECTS</b>
	LTE5A02D	<b>Documents structurés</b>	<b>39</b>	
	LTP5A02C	<b>Calculabilité</b>	<b>26</b>	
	LTP5A02D	<b>Acquisition, modélisation et représentation des connaissances</b>	<b>39</b>	
<b>TIM5A04</b>	<b>INFORMATIQUE</b>			<b>6 ECTS</b>
	LTP5A02E	<b>Programmation objet 1</b>	<b>39</b>	
	LTP5A02F	<b>Langages de script</b>	<b>39</b>	
<b>TIM5A05</b>	<b>INGÉNIERIE</b>			<b>9 ECTS</b>
	LTP5A02O	<b>Ingénierie des connaissances</b>	<b>26</b>	
	LTP5A02M	<b>Outils de traitement de corpus</b>	<b>19,5</b>	
	LTP5A02N	<b>Traitement statistique de corpus</b>	<b>39</b>	
	LTP5A02P	<b>Fouille de textes 1</b>	<b>26</b>	
	LTR5A02D	<b>Écriture et multilinguisme</b>	<b>26</b>	
<b>Total semestre 1</b>			<b>396,5</b>	

## Semestre 2

<b>TIM5B02</b>	<b>LINGUISTIQUE</b>			<b>3 ECTS</b>
		LTE5B02A	<b>Sémantique des textes multilingues 2</b>	<b>26</b>
		LTP5B02A	<b>Lexicologie, terminologie et dictionnaire 2</b>	<b>26</b>

<b>TIM5B03</b>	<b>INFORMATIQUE</b>			<b>3 ECTS</b>
		LTP5B02E	<b>Programmation objet 2</b>	<b>39</b>
		LTP5B02D	<b>Programmation itérative et récursive</b>	<b>26</b>

<b>TIM5B04</b>	<b>INGÉNIERIE</b>			<b>6 ECTS</b>
		LTP5B02G	<b>Technique web</b>	<b>26</b>
		LTP5B02F	<b>Bases de données pour le web</b>	<b>39</b>
		LTP5B02P	<b>Fouille de textes 2</b>	<b>26</b>
		LTP5B02L	<b>Méthodes en apprentissage automatique</b>	<b>26</b>

<b>TAL5B05</b>	<b>STAGE ET MÉMOIRE</b>			<b>18 ECTS</b>

<b>Total semestre 2</b>				<b>234</b>
-------------------------	--	--	--	------------

## Master Traitement Automatique des Langues – M2 -parcours Traductique et gestion de l'information

### Semestre 1

<b>TTG5A02</b>	<b>LINGUISTIQUE</b>				<b>6 ECTS</b>
		LTE5A02A	<b>Sémantique des textes multilingues 1</b>	<b>26</b>	
		LTP5A02A	<b>Lexicologie, terminologie et dictionnaire 1</b>	<b>26</b>	
		LTP5A02B	<b>Genres, textes et usages</b>	<b>26</b>	
<b>TTG5A03</b>	<b>MODÈLES ET FORMALISMES</b>				<b>3 ECTS</b>
		LTE5A02D	<b>Documents structurés</b>	<b>39</b>	
		LTP5A02D	<b>Acquisition, modélisation et représentation des connaissances</b>	<b>39</b>	
<b>TTG5A04</b>	<b>INFORMATIQUE</b>				<b>3 ECTS</b>
		LTP5A02G	<b>Base de données pour le web</b>	<b>39</b>	
		LTP5A02F	<b>Langages de script</b>	<b>26</b>	
<b>TTG5A05</b>	<b>INGÉNIERIE</b>				<b>9 ECTS</b>
		LTP5A02O	<b>Ingénierie des connaissances</b>	<b>26</b>	
		LTP5A02H	<b>Outils de TAO 1</b>	<b>39</b>	
		LTP5A02N	<b>Traitement statistique de corpus</b>	<b>39</b>	
		LTP5A02P	<b>Fouille de textes 1</b>	<b>26</b>	
		LTR5A02D	<b>Écriture et multilinguisme</b>	<b>26</b>	
<b>TTG5A06</b>	<b>TRADUCTION TECHNIQUE</b>				<b>9 ECTS</b>
		LTP5A02J	<b>Traduction technique 1</b>	<b>39</b>	
		LTP5A02K	<b>Traductologie</b>	<b>26</b>	
		LTP5A02L	<b>Conduite de projet de traduction 1</b>	<b>26</b>	
<b>Total semestre 1</b>				<b>468</b>	

## Semestre 2

<b>TTG5B02</b>	<b>LINGUISTIQUE</b>				<b>3 ECTS</b>
		LTE5B02A	<b>Sémantique des textes multilingues 2</b>	<b>26</b>	
		LTP5B02A	<b>Lexicologie, terminologie et dictionnaire 2</b>	<b>26</b>	
<b>TTG5B03</b>	<b>INGÉNIERIE</b>				<b>3 ECTS</b>
		LTP5B02G	<b>Technique web</b>	<b>26</b>	
		LTP5B02H	<b>Outils de TAO 2</b>	<b>39</b>	
		LTP5B02P	<b>Fouille de textes 2</b>	<b>26</b>	
		LTP5B02C	<b>Indexation et gestion de l'information</b>	<b>39</b>	
<b>TTG5B04</b>	<b>TRADUCTION TECHNIQUE</b>				<b>6 ECTS</b>
		LTP5B02I	<b>Traduction technique 2</b>	<b>39</b>	
		LTP5B02J	<b>Traductologie 2</b>	<b>26</b>	
		LTP5B02K	<b>Conduite de projet de traduction 2</b>	<b>26</b>	
<b>TAL5B05</b>	<b>STAGE ET MÉMOIRE</b>				<b>18 ECTS</b>
<b>Total semestre 2</b>				<b>273</b>	

Master Traitement Automatique des Langues – M2 - parcours Recherche et développement

**Semestre 1**

<b>TRD5A01</b>	<b>INGÉNIERIE EN TAL</b>		<b>9 ECTS</b>
	LTP5A02B	<b>Genres, textes et usages</b>	<b>26</b>
	LTE5A02A	<b>Sémantique des textes multilingues 1</b>	<b>26</b>
	LTP5A02A	<b>Lexicologie, terminologie et dictionnaire 1</b>	<b>26</b>
	LTP5A02D	<b>Acquisition, modélisation et représentation des connaissances</b>	<b>39</b>
	LTP5A02O	<b>Ingénierie des connaissances</b>	<b>26</b>
<b>TRD5A02</b>	<b>LINGUISTIQUE</b>	2 ou 3 enseignements	<b>9 ECTS</b>

<b>TRD5A03</b>	<b>INGÉNIERIE DES TEXTES ET DES DOCUMENTS</b>		<b>9 ECTS</b>
	LTP5A02Q	<b>Corpus annotés et développement de ressources linguistiques</b>	
	LTP5A02R	<b>Linguistique outillée et traitements statistiques</b>	
	LTP5A02S	<b>Langage du web sémantique</b>	
<b>TRD5A04</b>	<b>LANGUES VIVANTES</b>		<b>3 ECTS</b>

**Semestre 2**

<b>TRD5B01</b>	<b>STAGE</b>		<b>9 ECTS</b>
<b>TRD5B02</b>	<b>MÉTHODOLOGIE DE LA RECHERCHE</b>		<b>6 ECTS</b>
<b>TRD5B05</b>	<b>MÉMOIRE DE RECHERCHE</b>		<b>15 ECTS</b>

A word cloud of French terms related to linguistics and digital language processing. The words are arranged in a roughly triangular shape, pointing to the right. The terms include:

- Numérique
- INALCO
- Multilingue
- Web
- Informatique
- Textes
- TAL
- TIM
- Langues
- Grammaires
- Dictionnaires
- TAO
- Indexation
- Master
- Linguistique
- Traitement
- Licence
- Fouille