

---

**Ano Letivo** 2022-23

---

**Unidade Curricular** LINGUÍSTICA COMPUTACIONAL

---

**Cursos** CIÊNCIAS DA LINGUAGEM (2.º ciclo) (\*)

(\*) Curso onde a unidade curricular é opcional

---

**Unidade Orgânica** Faculdade de Ciências Humanas e Sociais

---

**Código da Unidade Curricular** 15091005

---

**Área Científica** CIÊNCIAS DA LINGUAGEM

---

**Sigla**

---

**Código CNAEF (3 dígitos)** 223

---

**Contributo para os Objetivos de Desenvolvimento Sustentável - ODS (Indicar até 3 objetivos)** 4,9,16

---

**Línguas de Aprendizagem** Português

---

**Modalidade de ensino**

Presencial

---

**Docente Responsável**

Jorge Manuel Evangelista Baptista

---

DOCENTE	TIPO DE AULA	TURMAS	TOTAL HORAS DE CONTACTO (*)
Jorge Manuel Evangelista Baptista	OT; S	S1; OT1	39S; 50T

\* Para turmas lecionadas conjuntamente, apenas é contabilizada a carga horária de uma delas.

---

ANO	PERÍODO DE FUNCIONAMENTO*	HORAS DE CONTACTO	HORAS TOTAIS DE TRABALHO	ECTS
1º	S2	39S; 50T	280	10

\* A-Anual;S-Semestral;Q-Quadrimestral;T-Trimestral

---

**Precedências**

Sem precedências

---

**Conhecimentos Prévios recomendados**

Conhecimentos básicos de informática na ótica do utilizador.

---

**Objetivos de aprendizagem (conhecimentos, aptidões e competências)**

Promover uma abordagem científica e tecnologicamente assistida ao estudo da linguagem humana, com especial referência ao léxico e a gramática do português.

No final desta UC o estudante terá uma visão abrangente dos métodos, recursos e ferramentas que as TIC permitem adotar na descrição linguística e as novas questões de investigação que suscitam.

### Conteúdos programáticos

Introdução à Linguística Computacional: objectivos e métodos. Segmentação do texto: texto, parágrafos, sentenças, palavras, tokens e separadores. Morfossintaxe: parte do discurso/categorias morfossintáticas, lema, palavras inflectidas, morfologia de dois níveis. Sintaxe: análise de dependência. Semântica: entidades nomeadas e ontologias; resolução de anáfora; anotação de papéis semânticos. Análise textual/de discurso Tradução automática. Processamento de linguagem natural e linguística (computacional). Aprendizagem mecânica (ou Inteligência Artificial) em Linguística (computacional).

---

### Metodologias de ensino (avaliação incluída)

Metodologia de ensino centrada no aluno. Avaliação por frequência: (i) leitura de textos seleccionados e resposta a questionários on-line, segundo calendário a indicar (50%); e (ii) elaboração em trabalho autónomo de pequenos projetos, segundo calendário a indicar, e resposta em aula a questionários on-line através da Tutoria Eletroónica (50%). Critério de assiduidade: mínimo 75% das horas de contacto para admissão a exame normal. Em todos os elementos de avaliação (incluindo exames) será atribuída uma percentagem de 20% à correcção linguística e discursiva e à estruturação textual.

---

### Bibliografia principal

- Clark, A., Fox, C., & Lappin, S. (Eds.). (2012). The handbook of computational linguistics and natural language processing. John Wiley & Sons.
- Jurafsky, D. and Martin, J.H. [2021-12-29]. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Prentice Hall. (3ª ed. preprint) <https://web.stanford.edu/~jurafsky/slp3/>
- Manning, C.D. and Schütze, H. (1999). Foundations of Statistical Natural Language Processing. The MIT Press.
- Mitkov, R. (Ed.). (2022). The Oxford handbook of computational linguistics. Oxford University Press. (3ª ed.)

---

**Academic Year** 2022-23

---

**Course unit** COMPUTACIONAL LINGUISTICS

---

**Courses** LANGUAGE SCIENCES (\*)  
Common Branch

(\*) Optional course unit for this course

---

**Faculty / School** FACULTY OF HUMAN AND SOCIAL SCIENCES

---

**Main Scientific Area**

---

**Acronym**

---

**CNAEF code (3 digits)** 223

---

**Contribution to Sustainable  
Development Goals - SGD  
(Designate up to 3 objectives)** 4,9,16

---

**Language of instruction** Portuguese

---

**Teaching/Learning modality** Face-to-face

**Coordinating teacher** Jorge Manuel Evangelista Baptista

Teaching staff	Type	Classes	Hours (*)
Jorge Manuel Evangelista Baptista	OT; S	S1; OT1	39S; 5OT

\* For classes taught jointly, it is only accounted the workload of one.

Contact hours	T	TP	PL	TC	S	E	OT	O	Total
	0	0	0	0	39	0	5	0	280

T - Theoretical; TP - Theoretical and practical ; PL - Practical and laboratorial; TC - Field Work; S - Seminar; E - Training; OT - Tutorial; O - Other

#### Pre-requisites

no pre-requisites

#### Prior knowledge and skills

Basic computer skills in user's perspective

#### The students intended learning outcomes (knowledge, skills and competences)

To promote a scientifically and technologically assisted approach to the study of human language, with special reference to the lexicon and grammar of Portuguese.

At the end of this course, the student will have a comprehensive view of the methods, resources and tools that ICTs allow to be adopted in language description and the new research questions that they raise.

#### Syllabus

Introduction to Computational Linguistics: aims and methods. Text segmentation: text, paragraphs, sentences, words, tokens, and separators. Morphosyntax: part-of-speech, lemma, inflected words, two-level morphology. Syntax: dependency parsing. Semantics: named entities and ontologies; anaphora resolution; semantic role labelling. Textual parsing. Machine translation. Natural language processing and (computational) linguistics. Machine learning (or Artificial Intelligence) in (computational) Linguistics.

### Teaching methodologies (including evaluation)

Student-centered teaching methodology. Assessment by frequency: (i) reading selected texts and answering online questionnaires, according to a given schedule (50%); and (ii) elaboration of small projects on their own, according to a given schedule, and answering in class to online questionnaires through the Tutoria Eletronica (50%). Attendance criteria: minimum 75% of the contact hours for admission to the regular exam (Exame Normal). In all the elements of assessment (including exams) a percentage of 20% will be given to linguistic and discursive correctness and textual structure.

---

### Main Bibliography

Clark, A., Fox, C., & Lappin, S. (Eds.). (2012). The handbook of computational linguistics and natural language processing. John Wiley & Sons.

Jurafsky, D. and Martin, J.H. [2021-12-29]. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Prentice Hall. (3ª ed. preprint) <https://web.stanford.edu/~jurafsky/slp3/>

Manning, C.D. and Schütze, H. (1999). Foundations of Statistical Natural Language Processing. The MIT Press.

Mitkov, R. (Ed.). (2022). The Oxford handbook of computational linguistics. Oxford University Press. (3ª ed.)