

24-25

MÁSTER UNIVERSITARIO EN  
HUMANIDADES DIGITALES: MÉTODOS Y  
BUENAS PRÁCTICAS

# GUÍA DE ESTUDIO PÚBLICA



## DESARROLLO Y EXPLOTACIÓN DE CORPUS ORALES PARA LA INVESTIGACIÓN EN LINGÜÍSTICA

CÓDIGO 2704017-

UNED

**24-25**

**DESARROLLO Y EXPLOTACIÓN DE  
CORPUS ORALES PARA LA  
INVESTIGACIÓN EN LINGÜÍSTICA  
CÓDIGO 2704017-**

# ÍNDICE

PRESENTACIÓN Y CONTEXTUALIZACIÓN  
REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA  
ASIGNATURA  
EQUIPO DOCENTE  
HORARIO DE ATENCIÓN AL ESTUDIANTE  
COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE  
RESULTADOS DE APRENDIZAJE  
CONTENIDOS  
METODOLOGÍA  
SISTEMA DE EVALUACIÓN  
BIBLIOGRAFÍA BÁSICA  
BIBLIOGRAFÍA COMPLEMENTARIA  
RECURSOS DE APOYO Y WEBGRAFÍA  
IGUALDAD DE GÉNERO

Nombre de la asignatura	DESARROLLO Y EXPLOTACIÓN DE CORPUS ORALES PARA LA INVESTIGACIÓN EN LINGÜÍSTICA
Código	2704017-
Curso académico	2024/2025
Título en que se imparte	MÁSTER UNIVERSITARIO EN HUMANIDADES DIGITALES: MÉTODOS Y BUENAS PRÁCTICAS
Tipo	CONTENIDOS
Nº ETCS	4
Horas	100
Periodo	SEMESTRE 2
Idiomas en que se imparte	CASTELLANO

## PRESENTACIÓN Y CONTEXTUALIZACIÓN

La denominada lingüística de corpus es una disciplina de investigación que implica la utilización de grandes cantidades de datos, ya sean escritos u orales, y herramientas informáticas para la descripción de la lengua y sus usos comunicativos.

El objetivo de esta asignatura es introducir a los/as estudiantes en los conceptos básicos y el trabajo práctico que implica el desarrollo y la utilización para la investigación en lingüística de corpus orales anotados. En ella se presta especial atención a aquellos aspectos de la lingüística de corpus que tienen que ver con el uso y preparación de herramientas informáticas para el desarrollo y explotación de este tipo de corpus. La asignatura, de 4 ECTS, se imparte en el segundo semestre del máster y tiene un carácter optativo.

## REQUISITOS Y/O RECOMENDACIONES PARA CURSAR ESTA ASIGNATURA

Aunque no existen requisitos obligatorios para cursar esta asignatura, es recomendable el conocimiento de alguna lengua extranjera moderna, preferentemente inglés o francés, para poder acceder a un mayor número de fuentes de recursos (datos, artículos, libros...) que no siempre están traducidos al castellano, o se traducen muy posteriormente a su publicación. El tener soltura con herramientas informáticas y estadísticas ayudará también al éxito en esta materia.

## EQUIPO DOCENTE

Nombre y Apellidos	JUAN MARIA GARRIDO ALMIÑANA (Coordinador de asignatura)
Correo Electrónico	jmgarrido@flog.uned.es
Teléfono	91398-6820
Facultad	FACULTAD DE FILOLOGÍA
Departamento	LENGUA ESPAÑOLA Y LINGÜÍSTICA GENERAL
Nombre y Apellidos	PAOLO ROSEANO .
Correo Electrónico	paolo.roseano@flog.uned.es
Teléfono	91398-6341
Facultad	FACULTAD DE FILOLOGÍA
Departamento	LENGUA ESPAÑOLA Y LINGÜÍSTICA GENERAL

## HORARIO DE ATENCIÓN AL ESTUDIANTE

Las/os estudiantes pueden contactar con las/os profesoras/es para resolver dudas sobre la asignatura en primer lugar a través del foro de la asignatura en el campus virtual correspondiente, a través del correo electrónico o por teléfono en el horario que se indica. Si se desea una entrevista personal, debe concertarse previamente. En todo tipo de comunicación con el profesorado se deberá indicar la asignatura a la que se refiere y utilizar el correo de la UNED.

### •Juan María Garrido Almiñana

- Horario de atención:
- Martes de 16 a 20 horas
- Miércoles de 10 a 14 horas
- Dirección postal:
- Facultad de Filología, Departamento de Lengua Española y Lingüística General, UNED, Paseo Senda del Rey, 7. 7ª planta. Despacho 7.15., 28040 Madrid
- Teléfono:
- 91.398.6820
- Correo electrónico:
- jmgarrido@flog.uned.es

### •Paolo Roseano

- Horario de atención:
- Miércoles de 10 a 14 horas
- Dirección postal:
- Facultad de Filología, Departamento de Lengua Española y Lingüística General, UNED, Paseo Senda del Rey, 7. 7ª planta, Despacho 7.06A, 28040 Madrid
- Teléfono:
- 91.398.6841
- Correo electrónico:
- paolo.roseano@flog.uned.es

## COMPETENCIAS QUE ADQUIERE EL ESTUDIANTE

### B Á S I C A S Y G E N E R A L E S

- CG1 - Administrar el trabajo en equipos multidisciplinares dedicados al ámbito de las Humanidades Digitales de forma eficiente, abordando los posibles conflictos de manera constructiva.
- CG2 - Conocer e identificar las nuevas técnicas y herramientas digitales para su empleo en la práctica profesional e investigadora en el ámbito de las Humanidades Digitales.

- CG3 - Describir y aplicar las tecnologías para la gestión y organización de la información y la documentación en el ámbito de las Humanidades Digitales.
- CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
- CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
- CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
- CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
- CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

### **ESPECÍFICAS**

- CE8 - Conocer y saber aplicar diferentes técnicas y tipos de representación de datos digitales y del resultado de su análisis, en el ámbito de las Humanidades Digitales.
- CE7 - Aplicar las tecnologías digitales en el tratamiento y la preservación de datos de diferente tipología en el ámbito de las Humanidades Digitales.

## **RESULTADOS DE APRENDIZAJE**

- Conocer qué se entiende por Lingüística de corpus como metodología de investigación, con especial atención a su uso para el estudio de la lengua oral.
- Mostrar nociones básicas sobre las principales tareas que implica el desarrollo de un corpus de lengua oral.
- Mostrar nociones básicas sobre las principales tareas que implica la explotación de un corpus de lengua oral para la investigación en el ámbito de la lingüística.
- Utilizar de forma autónoma algunas de las herramientas computacionales que pueden usarse para el desarrollo de corpus orales.
- Familiarizarse con el uso de las principales herramientas computacionales existentes para la explotación automatizada de corpus orales.

## CONTENIDOS

### BLOQUE I. Introducción

- Tema 1. La lengua oral
- Tema 2. La Lingüística de corpus

### BLOQUE II. La explotación de corpus orales

- Tema 3. Corpus orales para la investigación en Lingüística
- Tema 4. Herramientas computacionales para la explotación de corpus en Lingüística

### BLOQUE III. El desarrollo de corpus orales para la investigación lingüística

- Tema 5. El desarrollo de corpus orales
- Tema 6. Herramientas computacionales para el desarrollo de corpus orales

## METODOLOGÍA

La materia está planteada para su realización a través de la metodología general de la UNED, en la que se combinan distintos recursos y los medios impresos con los audiovisuales y virtuales. La metodología estará basada en los siguientes elementos:

1. Materiales de estudio: guía de estudio y web; textos obligatorios; materiales audiovisuales; bibliografía, etc.
2. Participación y utilización de las distintas herramientas del Entorno Virtual de Aprendizaje.
3. Tutorías en línea y telefónica: participación en los foros; comunicación e interacción con el profesorado.
4. Evaluación continua y sumativa: actividades prácticas de evaluación continua; pruebas presenciales; ejercicios de autoevaluación.
5. Trabajo individual o en grupo: lectura analítica de cada tema; elaboración de esquemas; realización de las actividades de aprendizaje propuestas.

## SISTEMA DE EVALUACIÓN

### TIPO DE PRUEBA PRESENCIAL

Tipo de examen

No hay prueba presencial

### CARACTERÍSTICAS DE LA PRUEBA PRESENCIAL Y/O LOS TRABAJOS

Requiere Presencialidad

No

Descripción

Deberá realizarse un trabajo individual sobre alguno de los contenidos tratados en la asignatura. El equipo docente propondrá posibles temas, pero podrán proponerse otros atendiendo a intereses personales. En todo caso, el equipo docente deberá dar su aprobación final al tema finalmente escogido.

#### Criterios de evaluación

Deberá demostrarse que se han asimilado correctamente los conocimientos trabajados en la asignatura sobre desarrollo y explotación de corpus orales para la investigación en lingüística.

#### **Deberá cuidarse la redacción, ortografía y puntuación.**

Ponderación de la prueba presencial y/o los trabajos en la nota final La calificación obtenida en el trabajo representa un 60% de la calificación final de la asignatura.

Fecha aproximada de entrega 15/01/2024

#### Comentarios y observaciones

No está permitido el uso de herramientas de Inteligencia Artificial Generativa para la elaboración del trabajo individual.

#### **PRUEBAS DE EVALUACIÓN CONTINUA (PEC)**

¿Hay PEC? Si, PEC no presencial

#### Descripción

Habrán dos pruebas de evaluación continua a lo largo del curso. Serán actividades prácticas relacionadas con los contenidos tratados en el temario de la asignatura.

#### Criterios de evaluación

Ponderación de la PEC en la nota final La calificación obtenida en las pruebas de evaluación continua aportan un 30% (15% cada una) de la calificación final de la asignatura.

Fecha aproximada de entrega

#### Comentarios y observaciones

No está permitido el uso de herramientas de Inteligencia Artificial Generativa para la elaboración de las PEC.

#### **OTRAS ACTIVIDADES EVALUABLES**

¿Hay otra/s actividad/es evaluable/s? Si, no presencial

#### Descripción

Se valorará la participación activa en la asignatura.

#### Criterios de evaluación

Se tendrá en cuenta la participación en los foros, realización de actividades prácticas no evaluadas y otras actividades que puedan plantearse a lo largo de la asignatura..

Ponderación en la nota final La valoración de la participación aporta el 10% de la calificación final.

Fecha aproximada de entrega

#### Comentarios y observaciones

### ¿CÓMO SE OBTIENE LA NOTA FINAL?

La calificación final será el resultado de la suma ponderada de las calificaciones de:

El trabajo (TRA) (60% de la calificación final)

Las pruebas de evaluación continua (PEC1 y PEC2) (30% de la calificación final, 15% cada una)

La participación (PAR) (10% de la calificación final)

**Cada una de las pruebas tendrá una calificación máxima de 10 puntos.**

**La calificación final será el resultado de aplicar la siguiente fórmula:**

**$(TRA * 60 / 100) + (PEC1 * 15 / 100) + (PEC2 * 15 / 100) + (PAR * 10 / 100)$**

## BIBLIOGRAFÍA BÁSICA

ISBN(13): 9780191750250

Título: THE OXFORD HANDBOOK OF CORPUS PHONOLOGY

Autor/es: Jacques Durand; Gjert Kristoffersen; Ulrike Gut

Editorial: : OXFORD UNIVERSITY PRESS

ISBN(13): 9788418625329

Título: USO DE CORPUS EN CLASE DE ELE. LA LENGUA REAL COMO MODELO. 1 edición

Autor/es: Elvira García, Wendy

Editorial: DIFUSIÓN

## BIBLIOGRAFÍA COMPLEMENTARIA

- Adolphs, S. y Carter, R. (2013). *Spoken corpus linguistics: From monomodal to multimodal*. Routledge. <https://doi.org/10.4324/9780203526149>
- Adolphs, S. y Knight, D. (2010). Building a spoken corpus: What are the basics? En A. O'Keeffe y M. McCarthy (Ed.), *The Routledge handbook of corpus linguistics*, pp. 38–52. Routledge. <https://doi.org/10.4324/9780203856949>
- Alcántara, M. (2008). Los retos en el análisis de los corpus de última generación. En R. Monroy y A. Sánchez (Ed.), *25 años de Lingüística Aplicada en España: hitos y retos / 25 years of Applied Linguistics in Spain: Milestones and challenges*. Servicio de Publicaciones de la Universidad de Murcia, pp. 701–706. <https://www.um.es/lacell/aesla/contenido/pdf/6/alcantar.pdf>
- Garrido, J. M., Renau, I. (2022). Lingüística de corpus y discurso en lengua española. En López, C., Carranza, I. E., Van Dijk, T. A. (eds.)- *Estudios del discurso. The Routledge Handbook of Spanish Language Discourse Studies*, Routledge, pp. 65-84.
- Llisterri, J., Machuca, M. J., de la Mota, C., Riera, M. y Ríos, A. (2005). Corpus orales para el desarrollo de las tecnologías del habla en español. *Oralia. Análisis del discurso oral*, 8, pp. 289–325. <https://joaquimllisterri.cat/publicacions/>



- Parodi, G., Cantos-Gómez, P., Howe, Ch. (eds.) (2022). *Lingüística de corpus en español / The Routledge Handbook of Spanish Corpus Linguistics*, Routledge.
- Llisterri\_Machuca\_Mota\_Riera\_Rios\_05\_Corpus\_Orales\_Tecnologias\_Habla\_Espanol.pdf
- Schiel, F., Draxler, C., Baumann, A., Ellbogen, T. y Steffen, A. (2012). *The production of speech corpora*. Bavarian Archive for Speech Signals.  
<https://doi.org/10.5282/ubm/epub.13693>

## RECURSOS DE APOYO Y WEBGRAFÍA

- European Language Resources Association (ELRA)*: <http://www.elra.info/en/>
- Linguistic Data Consortium (LDC)*: <https://www ldc.upenn.edu/>

## IGUALDAD DE GÉNERO

En coherencia con el valor asumido de la igualdad de género, todas las denominaciones que en esta Guía hacen referencia a órganos de gobierno unipersonales, de representación, o miembros de la comunidad universitaria y se efectúan en género masculino, cuando no se hayan sustituido por términos genéricos, se entenderán hechas indistintamente en género femenino o masculino, según el sexo del titular que los desempeñe.