



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

## Facultad de Humanidades y Ciencias de la Educación

### Actividades Educación Permanente 2025

#### Curso: Desarrollo de corpus y metadatos en LSU

**Área temática:** Lingüística de corpus

#### **Antecedentes, fundamentación y breve justificación de por qué es un curso o actividad de Educación Permanente**

Este curso de educación permanente que se propone tiene como antecedentes directos varias

actividades de investigación, extensión y enseñanza sobre estudios y producción de corpus en lengua de señas uruguaya desde el año 2018 hasta la actualidad. La línea de investigación y producción de corpus de LSU ha sido llevada adelante por un grupo de trabajo interinstitucional e interdisciplinar de la Udelar, integrado por el departamento de Estudios Sordos y Lenguaje de la TUILSU (FHCE), el Centro de Investigación Básica en Psicología (FPsico) y el Departamento de Ingeniería Eléctrica (FIng). Como parte de los antecedentes de actividades y proyectos del grupo sobre la temática de corpus en lengua de señas podemos mencionar al Proyecto Comunicación Bimodal en Lengua de Señas del programa Semilleros de Actividades Interdisciplinarias (2018), el Fortalecimiento de Colectivos Interdisciplinarios (2021) y el actual Núcleo Interdisciplinario en Multimodalidad (NIM) (2025) del Espacio Interdisciplinario. El objetivo de producción de un corpus en LSU con descripciones de metadatos se continúa actualmente con la financiación del programa I+D de CSIC y el NIM. El desarrollo y producción de estas actividades se han volcado en enseñanza curricular de la TUILSU, en la formación de intérpretes y traductores en LSU, y en actividades de extensión a través de Espacios de Formación Integral (CSEAM) y actividades en el medio con organizaciones e instituciones civiles vinculadas a la Comunidad Sorda y a la Lengua de Señas Uruguaya (Asociación de Sordos del Uruguay, Centro de Investigación de la Persona Sorda, Instituto de la Comunidad Sorda). La conformación de un corpus en LSU a través de una metodología científica y la descripción y sistematización de metadatos tienen un valor importante en el contexto científico en el Uruguay sobre la Lengua de Señas y los estudios sordos. Por el momento, los actuales estudios lingüísticos y los esfuerzos por realizar estudios psicolingüísticos en Lengua de Señas Uruguaya (LSU) se

enfrentan a cuerpos de mensajes dispersos, de cada especialista, sin metadatos comunes. En un contexto internacional de desarrollo y uso de grandes corpus para los estudios de esta naturaleza, la situación actual antes descrita para los estudios de LSU es una desventaja metodológica y reduce la validez de los estudios o, en algunos casos, los imposibilita. Esta situación afecta también las áreas de enseñanza de lengua como objeto formativo en todos los niveles educativos y a la accesibilidad clara y organizada de un corpus de lengua de uso general. Al generar un corpus público y abierto se pretende solventar esas limitaciones, usar esos metadatos para sustentar estudios de frecuencia u otros necesarios para estudios psicolingüísticos de lengua señada y contar con una estrategia para el crecimiento y actualización de ese corpus. En esa línea de trabajo, desde la Línea de Lenguaje del Centro de Investigación Básica en Psicología (CIBPsi), hemos hecho estudios comparativos de Asociación de Palabras, Análisis descriptivo del léxico temporal de la LSU y otros en curso (decisión léxica, procesamiento numérico, metáfora temporal y uso de información no lingüística en el procesamiento de señas). Actualmente, y como producto del trabajo realizado, hemos desarrollado y estandarizado los componentes centrales del corpus, particularmente en lo pertinente al desarrollo de criterios de etiquetación gramatical (fonológica, morfológica, sintáctica, semántica y pragmática) y de un sistema de glosado.

La enseñanza y capacitación continua en estudios de corpus se torna parte inherente a todo corpus de lengua sobre todo en aquellos que registran la producción lingüística de una comunidad sorda, lingüísticamente minoritaria e históricamente marginada. La capacitación en estudios de corpus en LSU favorece al abordaje de la lengua de señas en variadas perspectivas y áreas de acción, investigación y enseñanza en torno a la LSU y a la comunidad sorda en general, tanto para la formación, capacitación y perfeccionamiento de intérpretes oyentes y/o traductores sordos, docentes de enseñanza de lengua, investigadores y profesionales relacionados. A nivel general, si bien la formación de intérpretes y traductores LSU-español tiene un estándar alto, con el resultado de profesionales bien formados para esa tarea, la descripción gramatical de la LSU y de su uso son en sí mismo insumos para mejorar esa formación, para que el área académica alcance nuevas y mejores capacidades (p.e., investigadoras) y para un mejor conocimiento de los usos de la LSU. Estos tres elementos redundan en que los estudiantes y egresados de la TUILSU y otras instituciones educativas alcancen un mejor perfil de formación.

**Fecha de dictado: Del 13 de agosto a 30 de noviembre**

| Cargo | Nombre | Departamento/Sección |
|-------|--------|----------------------|
|-------|--------|----------------------|

|                                      |                   |                          |   |
|--------------------------------------|-------------------|--------------------------|---|
| <b>Responsable del curso</b>         | Profesora Adjunta | Ma. Eugenia Rodino Bueno | Área de Estudios Sordos/ TUILSU, FHCE, UDELAR                               |
| <b>Encargado del curso</b>           | Profesor adjunto  | Roberto Aguirre          | Centro de Investigación Básica en Psicología. FPsico, UDELAR                |
| <b>Otros participantes del curso</b> | Ayudante          | Juliana Madera           | Centro de Investigación Básica en Psicología. Fpsico, y TUILSU-FHCE, UDELAR |
|                                      | Ayudante          | Bruno Pérez              | Centro de Investigación Básica en Psicología. FPsico, UDELAR                |
|                                      | Ayudante          | Hernán Quintero          | Centro de Investigación Básica en Psicología. Fpsico y TUILSU-FHCE, UDELAR  |
|                                      | Ayudante          | Gabriel Morales          | Centro de Investigación Básica en Psicología. FPsico, UDELAR                |
|                                      | Ayudante          | Maximiliano Amaral       | Centro de Investigación Básica en Psicología. FPsico, UDELAR                |
|                                      | Adjunto           | Ariel Stassi             | Centro Universitario Regional Litoral Norte, UDELAR                         |
|                                      | Adjunto           | Luis Chiruzzo            | Instituto de Computación. FIng, UDELAR                                      |

|                           |                 |
|---------------------------|-----------------|
| <b>Carga horaria aula</b> | <b>30 horas</b> |
| <b>Plataforma EVA</b>     | <b>SI</b>       |
| <b>Trabajo final</b>      | <b>SI</b>       |

|                    |  |
|--------------------|--|
| <b>Cupos</b>       | 20   |
| <b>Condiciones</b> | Se requiere intérprete de lengua de señas en caso de cursar estudiantes sordos |
| <b>Servicios</b>   | Todos  |

### **Forma de evaluación (describa):**

A los estudiantes se les asignará un conjunto de videos con mensajes en LSU. Se evaluará el cumplimiento de la descripción de los metadatos de dicho conjunto de videos. El estudiante deberá etiquetar de manera adecuada –en los tres niveles: clase gramatical, codificación TRELSU, traducción al español– al menos el 60 % de los videos del corpus asignados. Se exigirá asistencia de al menos el 80 % de las sesiones. Los estudiantes contarán con las descripciones teóricas, metodológicas e informáticas necesarias para realizar su tarea. Se usará una rúbrica para especificar y detallar los componentes y sus puntajes concretos.

### **Conocimientos previos requeridos/recomendables (si corresponde):**

Conocimiento de la Lengua de Señas Uruguaya.

### **Objetivos:**

Capacitar a los estudiantes en el desarrollo de corpus de la LSU, en la descripción gramatical (metadatos) del mismo, en la reflexión sobre los rasgos lingüísticos de dicha lengua y su modalidad.

Fomentar en los futuros profesionales de la traducción el interés por una cultura científica en la descripción de la LSU y de su uso.

Aportar a los futuros y actuales profesionales de la traducción una herramienta de trabajo y una capacitación que mejore su capacidad descriptiva de la LSU y su uso.

### **Contenidos:**

Unidad 1: Generalidades teóricas y metodológicas de los estudios de corpus. Características y metodología de corpus. Comparación de corpus en lenguas de señas y en lenguas orales.

Unidad 2: Corpus en LSU. Metodología de recolección de datos. Diagramación de estímulos, estandarización de las condiciones de filmación, características de los participantes sordos, etc.

Unidad 3: Definición de Metadatos. Herramientas descriptivas del corpus en lengua de señas. Alcances y fidelidad de los datos. Niveles de análisis lingüísticos. Uso del sistema de glosado como tecnología de registro de las LS. Estudios estadísticos de metadatos, Herramientas y plataformas de estimación de pose, captura de movimiento, manejo y análisis de datos multimodales.

### **Bibliografía básica:**

1. Cormier, K., & Fenlon, J. (2014). *BSL Corpus Annotation Guidelines*, v. 1  
<http://www.bslcorpusproject.org/wp-content/uploads/BSLCorpusAnnotationGuideli>

nes\_23October2014.pdf: Deafness, Cognition and Language Research Centre, University College London.

2. Crasborn, O., Hulbosch, M., & Sloetjes, H. (2012). Linking Corpus NGT annotations to a lexical database using open source tools ELAN and LEXUS. En O. Crasborn, E. Efthimiou, E. Fotinea, T. Hanke, J. Kristoffersen, & J. Mesch (Eds.), *Proceedings of the fifth workshop on the representation and processing of sign languages: Interactions between corpus and lexicon* (pp. 19–22). Paris: European Language Resources Association.
3. Fojo, A., García, P., Pérez, B., Madera, J. .... Aguirre, R. (En prensa). Guía para la recolección de datos de corpus y el glosado de la Lengua de Señas Uruguaya. En R. Aguirre (Ed.). *Estudios interdisciplinarios de la Lengua de Señas Uruguaya*. CSIC.
4. Fojo, A., & Massone, M. I. (2012). *Estructuras lingüísticas de la Lengua de Señas Uruguaya*. TUILSU: Udelar.
5. Institut d'Estudis Catalans (2018). *Corpus de la Llengua de Signes Catalana. Portal de la Llengua de Signes Catalana*. <https://blogs.iec.cat/lsc/corpus/>
6. Johnston, T., & De Beuzeville, L. (2016). *Auslan corpus annotation guidelines*. Auslan Corpus.
7. Johnston, T. (2014). *Auslan Corpus Annotation Guidelines*, [http://media.auslan.org.au/attachments/Johnston\\_Auslan\\_CorpusAnnotationGuidelines\\_14June\\_2014.pdf](http://media.auslan.org.au/attachments/Johnston_Auslan_CorpusAnnotationGuidelines_14June_2014.pdf). Sydney, Australia: Macquarie University.
8. Johnston, T., and Schembri, A. (2013). Corpus analysis of sign languages. En: C. A. Chapelle (Ed.), *Encyclopedia of applied linguistics* (pp. 1312–1319). Oxford: Wiley Blackwell.
9. Peluso, L. (2014). Nueva versión del modelo de descripción fonológica TRELSU: Matriz segmental-articulatoria, configuración y movimiento. *Lengua de Señas e Interpretación*, (5), 63–95.
10. Stassi, A. E., Randall, G., Delbracio, M. Tancredí, M., Aguirre, R. .... Koleszar, V. (En prensa). Base de datos de Lengua de Señas Uruguaya para reconocimiento automático. En R. Aguirre (Ed.). *Estudios interdisciplinarios de la Lengua de Señas Uruguaya*. CSIC.