

Curso de transcripción con ELAN



Instituto da Lingua Galega

**CURSO DE
TRANSCRIPCIÓN
CON ELAN**

23 de febreiro e 2 de marzo (modalidade en liña)
24 de febreiro e 3 de marzo (modalidade presencial)

Esta actividade é parte do proxecto de I+D+i PID2019-110352GB-I00 financiado por MCIN/AEI/ 10.13039/501100011033/.

USC INSTITUTO DA LINGUA GALEGA

Reyes Rodríguez Rodríguez

Proxecto COCLI-MIL
Instituto da Lingua Galega

Febreiro 2022



1. Introducción

- 1.1. Que é ELAN
- 1.2. Para que serve
- 1.3. Posibilidades de explotación

2. Crear un arquivo de ELAN

- 2.1. Incorporar audio e/ou vídeo

3. Elementos relevantes

- 3.1. Visores: anotacións, onda, liñas
- 3.2. Controis: reprodución, volume e velocidade
- 3.3. Atallos
- 3.4. Modos: anotación, segmentación e transcrición

4. Crear un padrón

- 4.1. Vocabularios controlados
- 4.2. Tipos lingüísticos
- 4.3. Liñas de transcrición
 - 4.3.1. Liñas nais
 - 4.3.2. Liñas fillas

5. Proceso

- 5.1. Segmentación
- 5.2. Transcrición
 - 5.2.1. Criterios de transcrición
 - 5.2.2. Comezar a transcribir

6. Posibilidades

- 6.1. Buscas
- 6.2. Estatística
- 6.3. Importacións
- 6.4. Exportacións

1.1. Que é ELAN?

- Programa/Ferramenta de anotación e transcripción
- Permite sincronizar audio + vídeo
- Transcripción manual e semiautomática
- Permite anotar diferentes participantes e múltiples niveis e liñas
- <https://tla.mpi.nl/tools/tla-tools/elan/download/>

1.2. Para que serve ELAN?

- Sincronizar o audio (+vídeo) coa transcripción = correlación temporal texto-voz
- Visualizar transcripción + audio + vídeo ó mesmo tempo
- Analizar conversas con varios falantes
- Crear varias liñas para cada falante coa información que nos interese (ex.: transcripción ortográfica, fonética, análise do discurso...)
- Crear un corpus > buscas
- Exportar o material

1.3. Posibilidades de explotación

- Análise desde infinitas perspectivas: discurso, prosodia, fonética, variación, gramática, pragmática...
- Traballar a oralidade
- Analizar lingua de signos
- Material para o ensino da lingua
- Extracción de datos

2. Crear un arquivo ELAN

- Resumen do proceso:

1. Seleccionar audio (+ vídeo)
2. Organizar os participantes
3. Decidir e deseñar as liñas e a súa estrutura > CV, tipos e liñas
4. Construír e decidir os criterios de transcripción (silencios, solapamentos...)
5. Outras funcións de interese (modo, velocidade, aparencia...)
6. Segmentar
7. Transcribir
8. Aproveitamento: exportar, estatísticas, buscas...

Comezamos...

MERS-CORILGA-COCIDOGALEGO-01-2018

Canle Youtube: O Galego Mola

Título vídeo: COCIDO GALEGO tradicional

Data: 12/02/2018

Muller, tramo de idade: 35-50 anos

https://www.youtube.com/watch?v=zQI5xzU7Be8&list=PL0t_NwZqmZgQdftQqsCTiZdU7Fx1QOi9g

✓ Ficheiros na nosa Web (.wav e o .mp4)

1º) Incorporar audio (+ vídeo)

- Empezar unha nova transcripción : abrimos o programa ELAN > [Arquivo](#) > [Novo](#) > [Engadir](#) (1º audio; 2ª vídeo = [Editar](#) > [Ficheiros vinculados](#) > [Engadir...](#)> [Aplicar](#))
 - Formatos que soporta (ver en [Formato do Arquivo](#))
 - Consello: activar a copia de seguridade automática: [Arquivo](#) > [Copia de seguridade](#) > [1 minuto](#)
Ollo! Mellor ir gardando cada pouco manualmente
 - Incluír o audio xa editado previamente (recortar, engadir pitidos...)
- Recoméndase que o nome do audio sexa idéntico ao nome do arquivo ELAN.

OICO-URB-CREY-OURENSE-02-2020

[Tipo - Hábitat - Córpora de procedencia - Concello de gravación ou - Nº de gravación - Ano de gravación]

- O audio sempre o ten que estar localizable (pc, lapis...).

3. Elementos relevantes I

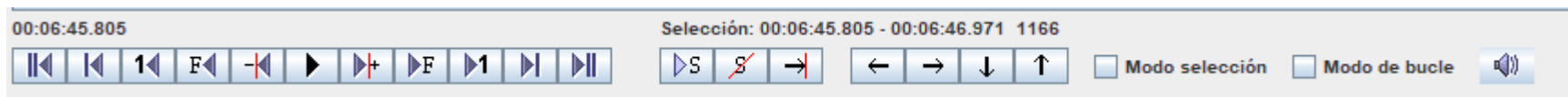
3.1. Visores

- Sincronización total
- **Visor do vídeo:** pode moverse, cambiar tamaño
- **Visor de densidade de anotacións:** especie de liña de tempo que permite navegar por todo o ficheiro, permite ver cantas anotacións se concentran nun momento.
- **Visor de formas de onda do arquivo .wav** (ou da pista do audio do vídeo). Estreitar/agrandar, barra vermella vaise movendo cando se escoita o son, cando seleccionas un intervalo de tempo > a forma da onda resáltase en azul.
- **Panel dos nomes de liña** vemos o nome das liñas e se nos poñemos enriba > despregable coas características de cada unha. Tamén botón dereito e podemos escoller varias opcións (ver por xerarquía, esconder algunha liña...)

3.2. Controis: reprodución, volume e velocidade

Opcións do reprodutor multimedia

- Controis:



Ver as táboas do manual páx. 119 – 120 coas explicacións de cada botón:

<https://www.mpi.nl/corpus/manuals/manual-elan.pdf>


- Volume
- Velocidade

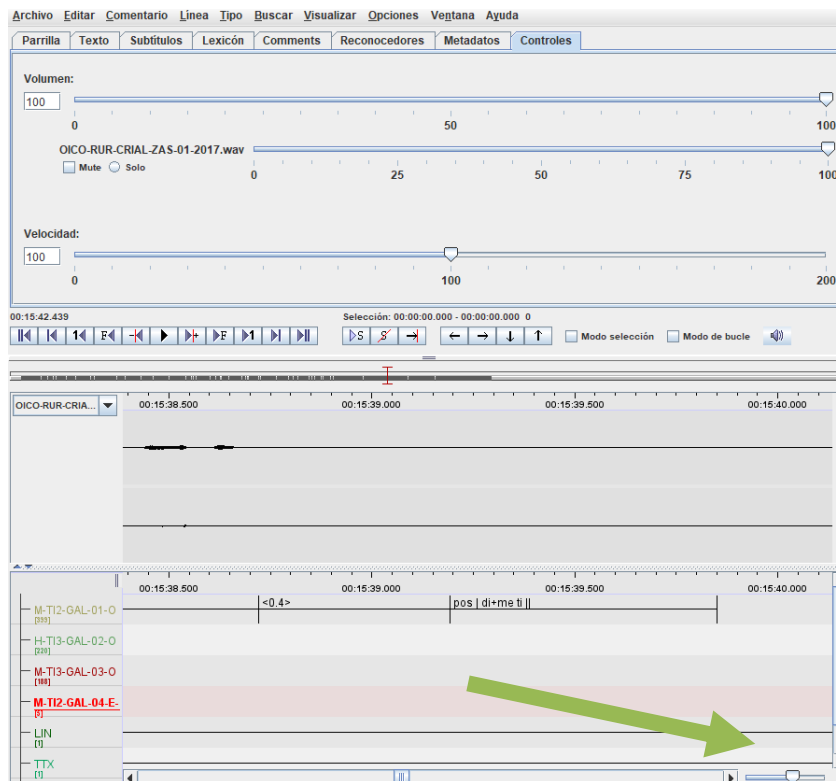
The image shows a software interface with a tabbed menu at the top: **Parrilla**, **Texto**, **Subtítulos**, **Lexicón**, **Comments**, **Reconocedores**, **Metadatos**, and **Controles**. The **Controles** tab is active. Below the tabs, there are two main control sections:

Volumen:
A horizontal slider is set to 100. The scale ranges from 0 to 100 with major ticks every 25 units. Below the slider, the file name **OICO-RUR-CRIAL-ZAS-01-2017.wav** is displayed. To the left of the slider are two radio buttons: **Mute** (checked) and **Solo** (unchecked).

Velocidad:
A horizontal slider is set to 100. The scale ranges from 0 to 200 with major ticks every 25 units.

Otras opciones:

- Aumento/diminución do tamaño da xanela
- Activar as liñas, movelas...
- Cambio do zoom temporal: Ctrl + roda do rato ou aumento/diminución da onda do son 







3.3. Atallos

Ver as táboas do manual páx. 119 – 120 cos atallos de cada botón:

<https://www.mpi.nl/corpus/manuals/manual-elan.pdf>

Táboa 1.3. controis de anotación

| ícona | Significado | atallo |
|---|---|--------------|
|  | Vaia á anotación anterior sobre o nivel de anotación activa | ALT+ESQUERDA |
|  | Vaia á seguinte anotación sobre o nivel de anotación activa | ALT+DEREITA |
|  | Vaia á anotación anterior. | ALT+ARRIBA |
|  | Vaia á anotación a continuación. | ALT+ABAIXO |



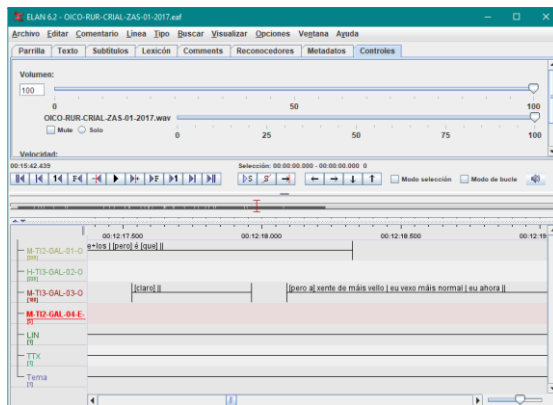
Preparación previa:

- ✓ Que necesito neste ELAN?
- ✓ Que información quero traballar?
- ✓ Necesito crear vocabulario(s) controlado(s)?
- ✓ Que tipos de liña teño que crear
- ✓ Que Liñas? Cantas? Xerarquía/dependencia?
- ✓ Precisión da segmentación?
- ✓ Criterios de transcripción?

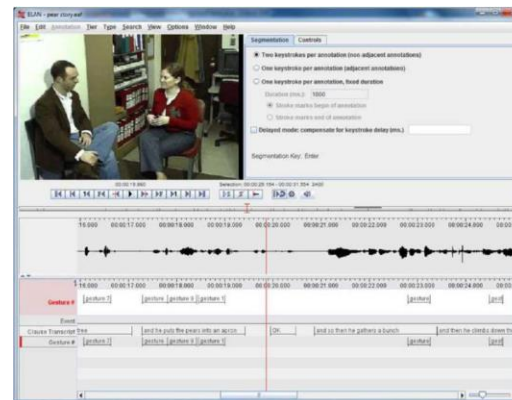
3.4. Modos de trabajo

Opciones > Modo de ...

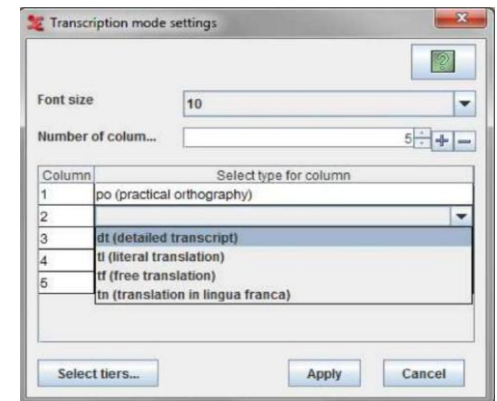
- Anotación



- Segmentación



- Transcripción

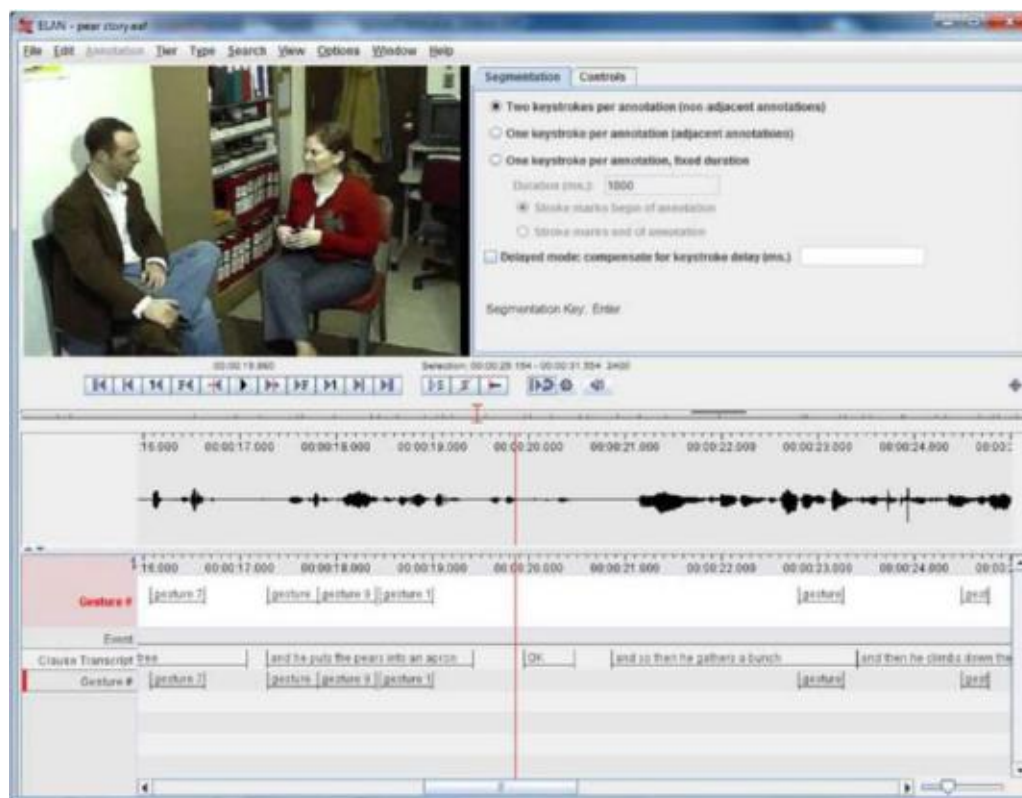


Modo de Anotación

- Xenérico (abre por defecto)
- Todas as opción de visualización, edición e busca
- Cómodo (completo)
- Editar, gardar, importar, exportar...
- Cargar ficheiro(s)
- Crear liñas e tipos lingüísticos
- Crear comentarios
 - Borrado
 - Solapamentos
 - Mover comentario
- Crear silencios

Modo de Segmentación

- Únicamente segmentar, non transcribir
- Creación rápida da segmentación
- Coa tecla **Intro** márcase o inicio e o fin do segmento



Modo de Transcripción

- Obx:
 - Aumentar a velocidade do traballo de transcripción: escoitar – escribir – escoitar – escribir...
 - Transcribir grandes cantidades de datos lingüísticos
- Consiste:
 - Cubrir os segmentos (a transcripción) nunha columna vertical
 - Diferentes cores por falantes
 - Permite facer revisión da segmentación
 - Clic sobre a liña > volve reproducir o segmento
 - *Olló! Presupón que a segmentación inicial xa está feita previamente (no modo anotación ou segmentación).
- Aparece un diálogo de configuración (1ª vez) / Axustes > seleccionar liñas: seleccionar/mostrar/ocultar as liñas, o tamaño da letra...

4. Crear un padrón

- Vocabularios controlados
- Tipos lingüísticos
- Liñas de transcripción

4.1. Vocabularios controlados (CV) I

- Consiste nunha serie de valores predefinidos que un usuario pode escoller para unha anotación
 - MORFOLOXÍA: substantivo, adxectivo, verbo, adverbio, preposición, conxunción, determinante, pronome...
 - SINTAXE: suxeito, predicado, complemento directo, complemento indirecto...
 - XESTOS: sorrí, asente, dubida...
- Obxectivo: facer a tarefa do anotador máis sinxela, máis rápida e axuda a non cometer erros
- Asociado ao tipo de liña
- **Editar > Editar vocabulario controlado > cadro de diálogo**

4.1. Vocabularios controlados (CV) II

... **Cadro de diálogo:**

Nome: **XESTOS** > **Engadir** > en **Valor de entrada:** introducimos as diferentes opcións (sorrí, dubida... > **Engadir** cada un > **Pechar**)

Editar vocabulario controlado

Vocabulario controlado

VC actuales

Nombre de VC

Descripción

Entradas

Valor de entrada

Descripción de entrada

Categoría de datos ISO

4.2. Tipo lingüístico I

- Cada liña está asociada a un Tipo lingüístico.
- Un tipo de liña fai referencia aos datos lingüísticos das liñas fillas. Á liña nai non se lle asigna ningún tipo = [*None*]
- Nomes para tipos: *ortografía, enunciados, fonética...*
- Cada tipo especifica unha serie de restriccións que están agrupadas en 4 estereotipos:

| | |
|-----------------------------|---|
| Symbolic subdivision | Crea segmentos dependentes da anotación principal, nos que introduce cada súa palabra ortográfica (entendida como a separada por espazos) de maneira automática. |
| Included in | Crea comentarios dependentes da liña principal, tomando como inicio e fin aqueles da liña da que dependa. Permite movemento bastante libre dos comentarios creados (en tempo inferior de duración, NON superior). |
| Time Subdivision | Crea comentarios dependentes da liña principal, sendo inamovible o valor temporal. |
| Symbolic association | Crea comentarios dependentes e ligados. Permite transcribir ambas as dúas liñas/comentarios en columnas paralelas (no modo de transcripción). |

- Importar tipos de liña
- CV (se nos interesa)

4.2. Tipo lingüístico II

- 4 estereotipos

| | |
|-------------------------|--|
| Symbolic subdivision | Crea segmentos dependientes da anotación principal, nos que introduce cadansúa palabra ortográfica (entendida como a separada por espazos) de maneira automática. |
| Included in | Crea comentarios dependientes da liña principal, tomando como inicio e fin aqueles da liña da que dependa. Permite movemento bastante libre dos comentarios creados (en tempo inferior de duración, NON superior). |
| Time Subdivision | Crea comentarios dependientes da liña principal, sendo inamovible o valor temporal. |
| Symbolic association | Crea comentarios dependientes e ligados. Permite transcribir ambas as dúas liñas/comentarios en columnas paralelas (no modo de transcripción). |

4.3. Liña I

- Liña: conxunto de anotacións que comparten as mesmas características (Ex.: ORT. 'información transcrita dos falantes')
- 2 tipos de liñas: que marcan a xerarquía
 - Liñas nai = independentes: contén anotacións vinculadas a un intervalo de tempo > ORT
 - Liñas fillas = dependentes: contén anotacións que están vinculadas a anotacións doutro nivel. Normalmente non van ligadas ao eixe temporal, é dicir, herdán a liña temporal da liña nai á que están asociadas > ESTÁNDAR, FONÉTICA...

*Olo! Se eliminas unha liña nai > eliminas as fillas

4.3. Liña II

- Liña > Engadir nova liña > cadro de diálogo
Ex. nome: M-TI2-GAL-01-ORT, H-TI3-GAL-02-ORT, Muller 1, Tema...
- Importar e exportar liñas: Liña > Importar tiras > Navegar > buscamos ficheiro > Seleccionar > Importar
- Para comprobar que a xerarquía está ben feita: clic dereito encima da liña > clasificar liñas > clasificar por xerarquía
- Copiar o xogo de liñas dun participante para outro: Liña > copiar tira > marco X en Copiar tiras dependentes tamén > seguinte > seguinte > Terminar > OK = Aparece un novo xogo de liñas ao que lle teño que cambiar as características do participante anterior = Liña > Cambiar ppdes liña > Seleccionar o que queremos modificar (_cp).
- Borrar liña: Liña > Borrar liña

4.3. Liña IV

- Engadir liña > Cadro de diálogo
- Información que cubrir: participante, anotador, tipo, CV...

Cambiar Atributos de Línea

Líneas Actuales

| Nombre de... | Línea Madre | Tipo Lingüí... | Participante | Anotador | Lengua po... | Content La... |
|---------------|-------------|----------------|--------------|----------|--------------|---------------|
| M-TI2-GAL-... | - | ORTOGRÁ... | ELA | Reyes | - | glg |
| H-TI3-GAL-... | - | ORTOGRÁ... | EL | Reyes | - | glg |
| M-TI3-GAL-... | - | ORTOGRÁ... | ELA | Reyes | - | glg |
| M-TI2-GAL-... | - | ORTOGRÁ... | ELA | Reyes | - | glg |
| LIN | - | ORTOGRÁ... | LIN | Reyes | - | |
| TTX | - | ORTOGRÁ... | TTX | Reyes | - | |

Añadir **Cambiar** **Suprimir** **Importar**

Seleccionar Línea: **M-TI2-GAL-01-ORT**

Nombre de Línea: M-TI2-GAL-01-ORT

Participante: ELA

Anotador: Reyes

Línea Madre: none

Tipo Lingüístico: ORTOGRÁFICO

Lengua por Defecto: Ninguno

Content Language: Galician (glg) - glg - <http://cdb.iso.org/lg/CDB-00138516-001>

Más opciones...

Cambiar **Cerrar**

4.3. Liña III: modelo (plantilla)

- Crear un modelo: abro o ELAN que quero gardar como plantilla > Menú: **Arquivo > Gardar como plantilla.**
- Como usar un modelo: abro programa ELAN > **Arquivo > Novo > meto primeiro o .wav (vídeo) e agora selecciono no cadro de diálogo (no centro) “plantilla” > busco o ficheiro > selecciono > Aceptar.**



5. Proceso

5.1. Segmentación






5.2. Transcripción

5.2.1. Criterios de transcripción

5.2.2. Comezar a transcribir

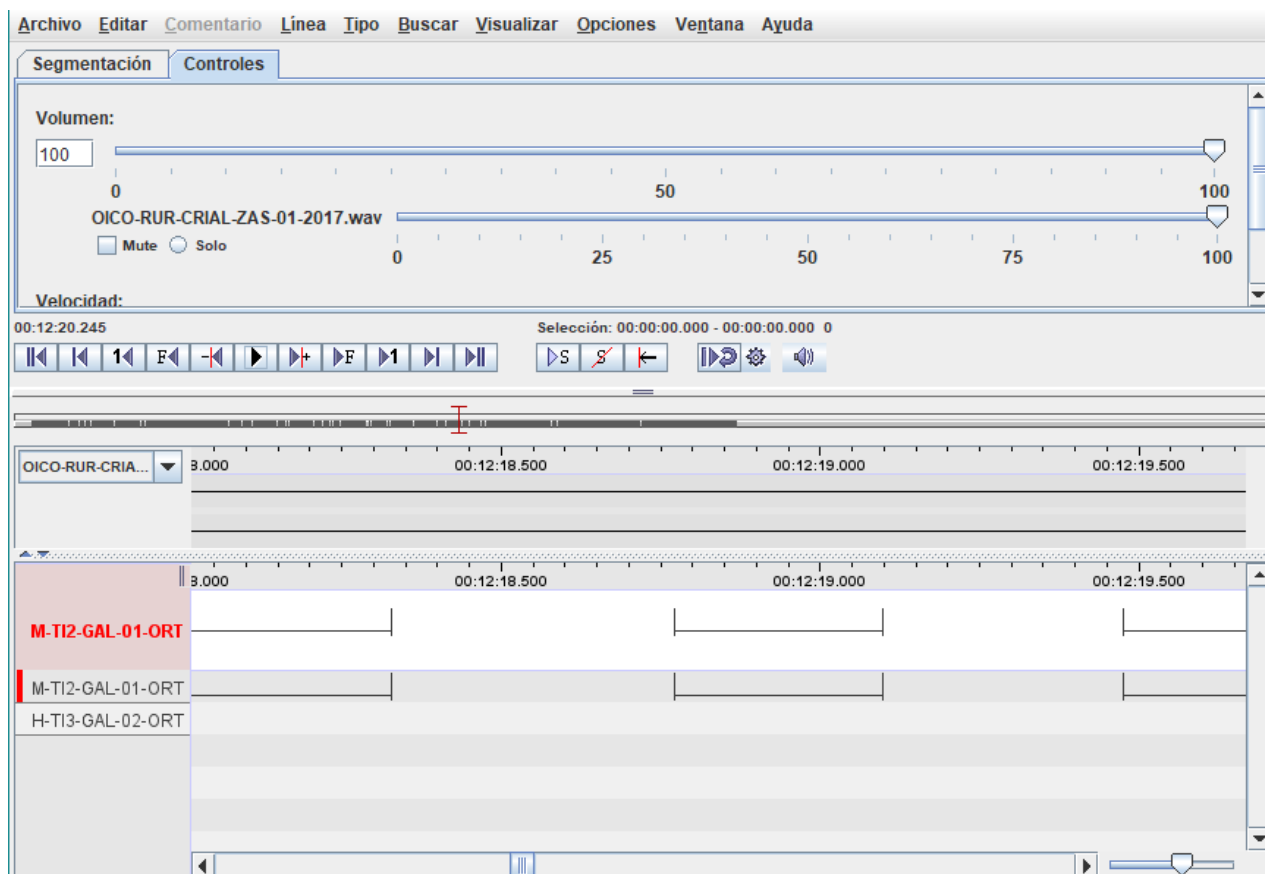
5.1. Segmentación en Modo anotación

- Seleccionar un intervalo de reproducción > botón dereito > novo comentario aquí
- Recomendacións previas:
 - Zoom da onda (clic dereito > zoom vertical) axuda a distinguir na liña da onda os enunciados dos silencios.
 - Se queremos que se creen comentarios nas liñas dependentes á vez que na independente: Editar > Preferencias > Edita preferencias > Edición > seleccionar 'Crear comentarios novos nas liñas dependentes cando se cree un comentario novo' // Outra opción: Liña > Copiar Tira // Outra opción: facelo manualmente seleccionar o comentario da liña nai e clic dereito na liña que nos interesa > Novo comentario aquí e xa nolo duplica.

- Comprobar que seleccionamos o fragmento que queremos = 
- Eliminar comentarios
- Mover comentarios: seleccionas o comentario > Ctrl + Alt + arrastrar o rato
- Solapamentos entre varios falantes.
- Crear comentarios nos espazos baleiros [Liña]
- Movernos ao comentario seguinte ou anterior dentro da mesma liña  
- Para saltar de liña  

5.1. Segmentación en Modo segmentación

- Creación rápida da segmentación
- Coa tecla **Intro** márcase o inicio e o fin do segmento





5.2. Transcripción

5.2.1. Diseñar os criterios de transcripción

5.2.2. Comezar a transcribir

- a) Transcripción en modo anotación
- b) Transcripción en modo transcripción

5.2.1. Diseñar criterios de transcripción individuais

Partimos do proxecto de CORILGA e vemos cales son os seus criterios e convencións á hora de transcribir.

Ideas:

- Final segmento ||
- Pausa |
- <comentarios transcriptor>
- + pronome
- - 2ª forma do artigo
- MAIÚSCULAS para a énfase

5.2.2. Transcribir en modo Anotación

Dobre clic nun segmento previamente creado > escoito  > transcribo

*Neste modo, tamén podemos ir segmentando e transcribindo ó mesmo tempo

*Multiplicar información (silencios, tokens, solapamentos...)

*Liña que ten CV: crear comentario > botón dereito > Novo comentario > salta o despregable > seleccionar a opción

5.2.2. Transcribir en Modo Transcripción

Cadro de diálogo > selecciono a liña ORT en que quero transcribir > Aplicar > Clic en cada liña > escoito > transcribo

Revisión I

Cambios básicos: **Clic no comentario > botón derecho > Opciones**


- Zoom
- Tamaño de Fuente
- Anotación activa negrita
- Altura de línea reducida
- Barra de desplazamiento horizontal visible
- Barra temporal visible
- Conectado
- Modo marcador
- Nuevo comentario aquí
- Modificar el valor del comentario
- Modificar categoría de datos del comentario Alt+Shift-M
- Combinar con el comentario siguiente
- Unir con el comentario precedente
- Dividir el comentario
- Borra el valor del comentario Ctrl-Delete
- Modificar tiempo de anotación Ctrl-Enter
- Modify Annotation Time... Ctrl+Shift-M
- Borrar comentario
- Copiar comentario
- Copiar grupo de comentario
- Desplazar Comentario activo Ctrl+Shift-Enter

Ou no menú **Liña:**

| Liña | Tipo | Buscar | Visualizar | Opciones | Ventana | Ayuda |
|---|------|--------|------------|----------|---------|------------|
| Añadir línea nueva... | | | | | | Ctrl-T |
| Cambiar propiedades de la línea | | | | | | |
| Restituir la tira como madre... | | | | | | |
| Borrar línea | | | | | | Ctrl+Alt-T |
| Importar Tiras | | | | | | |
| Añadir un participante nuevo... | | | | | | |
| Extraer tokens... | | | | | | |
| Filtrar tira... | | | | | | |
| Copiar tira | | | | | | |
| Copy Annotations from Tier to Tier... | | | | | | |
| Combinar líneas... | | | | | | |
| Merge Tiers...(Classic) | | | | | | |
| Combinar grupo de líneas... | | | | | | |
| Crear comentarios en las líneas dependientes | | | | | | |
| Crear comentarios de las superposiciones... | | | | | | |
| Crear comentarios de las superposiciones... (Classic) | | | | | | |
| Crear comentarios a partir de la diferencia... | | | | | | |
| Crear comentarios en los espacios vacíos | | | | | | |
| Crea comentarios normales... | | | | | | |
| Borrar comentarios o valores... | | | | | | |
| Convert Annotation Values to Tiers... | | | | | | |
| Etiqueta y numera los comentarios... | | | | | | |
| Cambiar los comentarios a mayúscula/minúscula... | | | | | | |

Revisión II

Cambios básicos no menú **Liña**

| <u>L</u> ínea | <u>T</u> ipo | <u>B</u> uscar | <u>V</u> isualizar | <u>O</u> pciones | <u>V</u> entana | <u>A</u> yuda |
|---|--------------|----------------|--------------------|------------------|-----------------|---------------|
| <u>A</u> ñadir línea nueva... | | | | | | Ctrl-T |
| Cambiar propiedades de la línea | | | | | | |
| Restituir la tira como madre... | | | | | | |
| <u>B</u> orrar línea | | | | | | Ctrl+Alt-T |
| <u>I</u> mportar Tiras | | | | | | |
| <u>A</u> ñadir un participante nuevo... | | | | | | |
| <hr/> | | | | | | |
| <u>E</u> xtraer tokens... | | | | | | |
| <u>F</u> iltrar tira... | | | | | | |
| <u>C</u> opiar tira | | | | | | |
| Copy Annotations from Tier to Tier... | | | | | | |
| <u>C</u> ombinar líneas... | | | | | | |
| Merge Tiers...(Classic) | | | | | | |
| Combinar grupo de líneas... | | | | | | |
| Crear comentarios en las líneas dependientes | | | | | | |
| Crear comentarios de las superposiciones... | | | | | | |
| Crear comentarios de las superposiciones... (Classic) | | | | | | |
| Crear comentarios a partir de la diferencia... | | | | | | |
|  Crear comentarios en los espacios vacíos | | | | | | |
| Crear comentarios normales... | | | | | | |
| Borrar comentarios o valores... | | | | | | |
| Convert Annotation Values to Tiers... | | | | | | |
| Etiqueta y numera los comentarios... | | | | | | |
| Cambiar los comentarios a mayúscula/minúscula... | | | | | | |

... Retomamos os criterios de transcripción

Partimos do proxecto de CORILGA e vemos cales son os seus criterios e convencións á hora de transcribir.

CORILGA

Corpus Oral Informatizado da Lingua Galega

Por se vos interesa... engádovos os criterios para nomear os arquivos e tamén para nomear as liñas en Corilga.

1. Nomear os arquivos
2. Nomear das liñas de anotación
3. Convencións de transcripción

1. Nome dos arquivos

Cada gravación inclúe un arquivo de audio asociado a un *eaf* co que comparte o mesmo nome. A súa codificación recolle información abreviada sobre a tipoloxía textual, hábitat, córpora oral de procedencia, lugar e ano da gravación.

1.1. Tipo de texto

Distinguimos tres grandes categorías en función de se a gravación constitúe un texto informal, formal ou dos medios de comunicación. Alén diso, incorporamos varias subcategorías asociadas a unha abreviatura que se inclúe no nome do arquivo.

É posíbel que un mesmo texto poida presentar varias tipoloxías. Neste caso, aparecerán todas recollidas nunha base de datos de Microsoft Access mais só reflectiremos no nome do arquivo o tipo de texto predominante.

| Oralidade informal | |
|---------------------|------|
| Conversa | OICO |
| Monólogo | OIMO |
| Lectura | OILE |
| Entrevista dirixida | OIED |

| Oralidade formal | |
|------------------|------|
| Discurso oral | OFDO |
| Discurso lido | OFDL |
| Poesía | ODPO |

| Medios de comunicación | |
|------------------------|------|
| Entrevista | MEEN |
| Debate | MEDE |
| Conversa | MECO |
| Redes sociais | MERS |

1.2. Hábitat

Distinguimos tres categorías en función do número de habitantes: rural, semiurbano e urbano. Para clasificar as gravacións, tomaremos como referencia as cifras ofrecidas polo INE.

A referencia ao hábitat só se incluírá no código cando os textos pertencen á oralidade informal. Non sinalamos o hábitat nos exemplos de oralidade formal ou dos medios de comunicación porque non achegaría información sociolingüística relevante.

| | |
|-------------------------------|------|
| Rural (-5000 hab) | RUR |
| Semiurbano (5.000–10.000 hab) | SURB |
| Urbano (+10.000 hab) | URB |

1.3. Córpora de procedencia

Os arquivos de gravación do proxecto CORILGA pertencen a córpora diferentes. No código incluímos información sobre a súa procedencia empregando as siglas de cada un deles.

| | |
|---|------|
| Atlas Lingüístico Galego | ALGA |
| Corpus do arquivo etnógrafo do profesor Gustav Henningsen | CHEN |
| Corpus de Reyes Rodríguez | CREY |

1.4. Concello de gravación ou apelido das e dos participantes

- Na base de datos recóllese o concello, lugar e parroquia de cada informante e tamén o concello, lugar e parroquia de gravación seguindo as normas que ofrece o Nomenclátor de Galicia.

Ex: OICO-URB-CREY-LUGO-01-2013

1.5. Número de gravación

Distingue as gravacións para evitar que se repita un mesmo nome.

1.6. Ano de gravación

Información en formato numérico completo sobre o ano no que foi realizada a gravación

MERS-URB-CREY-OURENSE-02-2022

[Tipo de texto – Hábitat - Córpora de procedencia - Concello de gravación ou – Nº de gravación - Ano de gravación]

2. Nome das liñas de anotación

Cada liña de anotación presenta un nome que inclúe información abreviada sobre sexo, orde de intervención e franxa de idade que corresponde a cada falante. Alén diso, incorpora información adicional sobre o tipo de transcripción.

2.1. Xénero

Distinguimos dous xéneros: home e muller.

| | |
|--------|---|
| Muller | M |
| home | H |

2.2. Tramo de idade

Non incluímos datos numéricos exactos senón que adscribimos cada falante a un grupo de idade:

| | |
|---------|-----|
| - 30 | T11 |
| 30 - 65 | T12 |
| + 65 | T13 |

2.3. Idioma da gravación

| | |
|------------------|-----|
| Galego | GAL |
| Español | ESP |
| Galego e español | GE |

2.4. Número de participantes

Distingue as e os falantes dependendo da súa orde de aparición na gravación para evitar casos nos que se repita un mesmo nome.

2.5. Entrevistadora ou entrevistador

Nos exemplos de entrevistas dirixidas distinguiremos a fala do entrevistador ou entrevistadora engadindo un “E” no nome da súa liña.

2.6. Contido da liña

No programa Elan, cada falante está asociado a varias liñas de transcripción en función do contido. Por este motivo, empregaremos o seu correspondente código e incorporaremos unha das seguintes abreviaturas en cada caso.

| | |
|-----------------------|-----|
| Anotación ortográfica | ORT |
| Anotación fonética | FON |
| Anotación sintáctica | SIN |

Exemplo:

M-TI2-GAL-01-ORT

[Trataríase dunha liña onde se inclúe a transcripción ortográfica dunha muller menor de 25 anos que foi a primeira en comezar a falar en lingua galega].

Outros tipos de liñas en CORILGA:

Existen dous tipos de liñas que non presentan a codificación antes sinalada:

| | |
|-------------------------|-----|
| Anotación tipo de texto | TTX |
| Anotación temas | TEM |
| Ánotación lingua | LIN |

Neste caso, non están asociadas a cada falante senón ao conxunto da gravación. Por este motivo creamos unha liña independente que leva como código exclusivamente a abreviatura de tipo de texto (TTX) ou anotación temática (TEM). Exemplos:

Listaxe de temas:

- <1. Historias de vida (infancia, moceo, matrimonio, morte...)>
- <2. Lingua e comentarios sociolingüísticos>
- <3. Literatura>

Listaxe de tipos de texto:

- Conversa OICO <OICO>
- Monólogo OIMO <OIMO>
- Lectura OILE <OILE>

*Ser rigorosos para poder recuperar a información

3. Convencións de transcripción

En xeral, as palabras escríbense seguindo a súa forma ortográfica normal, acentuación, contraccións que se representan na escrita, etc., aínda que con algunhas particularidades.

a.1.) Consideracións xerais e símbolos empregados



1. Secuencias comunicativas

| | |
|--|--|
| 1.1. Secuencias non-verbais | Non se sinalan. |
| 1.2. Secuencias verbais | <ul style="list-style-type: none"> - Transcrición ortográfica segundo as <i>Normas ortográficas e morfolóxicas do idioma galego</i>. Non obstante: <ul style="list-style-type: none"> - Só se empregan as maiúsculas no caso de nomes propios. - Sen signos de puntuación porque poden inducir a erros derivados dos hábitos de lectura. - A acentuación é regular segundo as <i>Normas</i>. Non se sinalan timbres abertos e pechados aínda que as variacións na colocación do acento si que aparecen marcadas. Ex: ¡<i>hostiá!</i> ou ¡<i>hombré!</i> Así mesmo, algunhas contraccións tamén aparecerán marcadas cun acento diacrítico para diferencialas doutras formas coincidentes na grafía. Ex: <i>pa</i> (<i>para</i>), <i>pá</i> (<i>para+a</i>). - O texto dispónse en secuencias aliñadas á marxe esquerda e encabezadas pola identificación das interlocutoras e interlocutores. |
| 1.3. Secuencias en máis de unha lingua | <p>Non se traducen fragmentos de linguas diferentes ao galego nin tampouco se usa unha tipografía diferente. En consecuencia, os castelanismos, en xeral, escríbense coa grafía do castelán: <i>ejemplo, jabalí, congelador, mayores</i></p> <p>Non obstante, as formas que amosan influencia do castelán mais que son xerais na fala da que se tratan poden manter a grafía galega. Ex: <i>plaiá</i>.</p> |
| 1.4. Discurso directo reproducido | “texto afectado” |

2. Símbolos relativos ás quendas de palabra

| | |
|--|-----------------------|
| 2.1. Continuación da quenda de palabra en casos de presenza simultánea de varias voces. Este factor é decisivo para establecer a diferenza entre exemplos de solapamento ou de interrupción. | & seguido dun espazo. |
| 2.2. Enlace entre quendas. Rapidez extrema na sucesión das quendas de palabra. | % seguido dun espazo. |
| 2.3. Solapamento | [texto afectado] |

3. Símbolos prosódicos I

| | |
|---|---|
| 3.1. Márcanse algúns aspectos indispensábeis para a comprensión do texto. | |
| 3.2. Separación de palabras | espazo en branco |
| 3.3. Alongamento dun son | : :: :: seguido dun espazo Exemplo: texto afectado:: |
| 3.4. Corte abrupto no medio dunha palabra | # sen espazo precedente. Exemplo: texto afect# <u>Importante: Cando o corte se producir cando só se chega a emitir un son da palabra (por exemplo, p#), é necesario colocalo entre <> como comentario do transcriptor (“<p corte>”), para desta maneira evitar problemas no aliñamento.</u> |
| 3.5. Entoación interrogativa | ¿texto afectado? |
| 3.6. Entoación exclamativa | ¡texto afectado! |
| 3.7. Fin de unidade prosódica | precedida e seguida dun espazo. Exemplo: texto afectado |
| 3.8. Pausa breve ou mediana | precedida e seguida dun espazo. Exemplo: texto afectado |
| 3.9. Pausa de longa duración | <segundos> precedida e seguida dun espazo. Exemplo: texto <0.5> afectado |
| 3.10. Énfase | Maiúsculas Exemplo: TEXTO AFECTADO |
| 3.11. Intensidade | |
| Intensidade forte “forte” | {{(F) texto afectado}} |
| Intensidade moi forte “fortísimo” | {{(FF) texto afectado}} |
| Intensidade suave “piano” | {{(P) texto afectado}} |
| Intensidade moi suave “pianísimo” | {{(PP) texto afectado}} |
| 3.12. Tempo | |
| Tempo rápido “acelerado” | {{(AC) texto afectado}} |
| Tempo moi rápido | {{(ACC) texto afectado}} |
| Tempo lento “desacelerado” | {{(DC) texto afectado}} |
| Tempo moi lento | {{(DCC) texto afectado}} |
| 3.13. Risas | |
| Risas simultáneas con | {{(risas) texto afectado}} |

3. Símbolos prosódicos II

| | |
|---------------------------------------|---|
| 3.14. Inhalación | (INH) |
| 3.15. Exhalación | (EXH) |
| 3.16. Sons paralingüísticos | Escritura ortográfica precedida dun comentario da/o transcritor onde se recolle o fenómeno. Empregamos o formato que se apunta no punto 5.2. {< fenómeno >}. Asentimento {<asentimento> hum / ahá / ha ha} Dúbida {<dúbida> m: / e:} <dúbida nasal> Apelación {<apelación> chist} Apelación {<apelación> eh /ei /oi} Apelación {<apelación> ho} Demanda de silencio {<silencio> sh} Desacordo {<desacordo> puf / uf / buf} Valoración {<valoración> ps /buf} Sorpresa {<sorpresa> oh /oi /buff} Desagrado {<desagrado> puag} Cansazo {<cansazo> ah} Recordo {<recordo> ah} Onomatopeas {<onomatopea> patapumba} |
| 3.17. Outros fenómenos | |
| Non simultáneos con elementos verbais | (descrición do fenómeno) |
| Simultáneos con elementos verbais | {(fenómeno) texto afectado} |

4. Regularizacións, comentarios e fragmentos conflitivos

| | |
|---|---|
| 4.1. Regularizacións do texto | <ul style="list-style-type: none">- En liñas xerais, n on se adaptan os textos orais á variedade estándar excepto no tocante á ortografía.- Non se suprimen nin se modifican segmentos do texto oral. Porén:<ul style="list-style-type: none">- As contraccións realizadas por fonética sintáctica que son esperábeis polo contexto fonético non se transcriben, a menos que o estándar da lingua galega fixe a súa realización escrita. |
| 4.2. Comentarios da ou do transcritor | < > |
| 4.3. Fragmentos incomprensíbeis | {<incomprensíbel> texto afectado} Incluiremos a duración en segundos nos fragmentos que sexan moi extensos. |
| 4.4. Suposicións ou fragmentos incertos | {<incerto> texto afectado} |

Recordamos o procedemento para crear un ficheiro ELAN:

1. Subir o audio (vídeo)

2. Decidir que tipos de liña vou necesitar

3. Se necesito CV

1. Xestos: alegría, tristeza, risa
2. Morfoloxía: substantivo, verbo, adxectivo, pronome, determinante...
*Marcaremos o CV cando creemos o Tipo

4. Crear tipos de liña (nai , fillas - estereotipos)

1. Nai: transcripción
2. Fillas: tradución, notas, xestos, tema...
*Entendo que con ter 2 tipos (None e outro Time Subdivision é suficiente, ou un None se son todos independentes)
*Recordar que podemos copiar ou importar as liñas desde outro ELAN.

5. Crear un xogo de liñas por cada participante (ou só unha liña se non preciso máis)

1. Nai: M-TI3-GAL_ORT *Podo ter moitas/todas independentes = Liña nai: **None**
2. Fillas: tradución, xestos, tema

6. Comezar por segmentar as liñas nai (as dependentes aparelladas ou non) = Crear comentarios

1. Modo de anotación: seleccionar – arrastras
2. Modo de segmentación: intro-intro
*Acordarse que mentres segmentamos a liña nai temos a opción de que á vez nos vaia creando comentarios nas liñas que dependen da liña nai.

7. Transcribir

1. Modo de anotación: cada comentario: enchendo en horizontal e vertical
2. Modo de segmentación: un a un en vertical



6. Posibilidades - Recursos

6.1. Búsquedas

6.2. Estadística

6.3. Importaciones

6.4. Exportaciones

Instituto da Lingua Galega



**CURSO DE
TRANSCRIPCIÓN
CON ELAN**

23 de febreiro e 2 de marzo (modalidade en liña)
24 de febreiro e 3 de marzo (modalidade presencial)

Esta actividade é parte do proxecto de I+D+i PID2019-110352GB-I00 financiado por MCIN/AEI/ 10.13039/501100011033/;



USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
INSTITUTO DA LINGUA GALEGA

reyes.rodriguez.rodriguez@usc.es