

MANUAL DE PROCESAMIENTO DE ARCHIVOS ELAN

Adaptado por:
Faustino Montes Castañeda
CIESAS, CDMX
faustino.montes.c@gmail.com

Primera version por:
Clifton Pye
The University of Kansas
pyersqr@ku.edu

4 de febrero de 2019



El “Manual de procesamiento de archivos Elan”, describe los procedimientos que se debe seguir para procesar los datos transcritos en Elan, al programa Qancod.pl, desde una PC Windows.

1. Ordenar y unificar los archivos con los siguientes códigos de descripción.

T-D-03-02-18

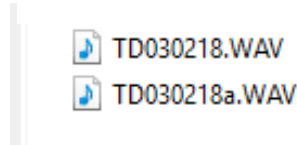
Totonaco-nombre de la niña focal-día de grabación-mes de grabación-año de grabación

Ilustración 1. Codificación de las carpetas



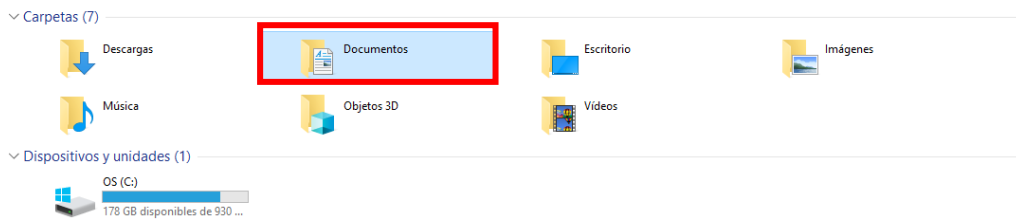
En la carpeta **TD030218** se almacenan los materiales grabados en campo, ya sea desde el formato en audio o video. Los materiales tienen que llevar el mismo nombre de la carpeta (Ilustración 1).¹

Ilustración 2. Codificación de los audios



2. Se sugiere que desde el inicio los materiales (audio y video) estén ubicados en la carpeta **Documentos**. Al estar almacenados en otras carpetas (Vídeo, Escritorio, Imágenes, etc.), el programa **Perl** no reconoce automáticamente la ubicación de otras carpetas.

Ilustración 3. Documentos



¹ Si en un día se graba varias veces, los archivos tienen que tener los mismos códigos, sólo se le agrega una letra distinta del abecedario, comenzando de la **a** a la **Z**, para distinguir cada uno de los archivos (Ver Ejemplo 2).

3. Antes de descargar el programa **Perl**, es importante verificar el sistema operativo de la PC. Algunos equipos manejan el sistema operativo de **32 bits** (depende de la memoria RAM), otras manejan el sistema de **64 bits**.²

El programa **Perl** maneja los dos sistemas operativos de la PC Windows, 32 bits y 64 bits.

Ilustración 4. Sistema operativo de Windows 10

[Ver información básica acerca del equipo](#)

Edición de Windows

Windows 10 Home Single Language
© 2018 Microsoft Corporation. Todos los derechos reservados.

Sistema

Fabricante: ASUSTek Computer Inc.
Procesador: Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz 2.90 GHz
Memoria instalada (RAM): 8,00 GB (7,89 GB utilizable)

Tipo de sistema: Sistema operativo de 64 bits, procesador x64

Lapiz y entrada táctil: La entrada táctil o manuscrita no está disponible para esta pantalla

Compatibilidad con ASUSTek Computer Inc.

4. Descargar e instalar el programa **Perl**³

Ilustración 5. Programa Perl



² **Nota:** Si es la primera que instala el programa Perl en su computadora, y al momento ejecutar no funciona, desinstale e instale nuevamente.

³ **Nota:** En una PC Windows 10, viene preestablecido el programa “SÍMBOLO DEL SISTEMA”. Este programa es similar al programa **Perl**. Sin embargo, para fines del trabajo, ésta no reconoce algunas “ubicaciones específicas” cuando se usa el programa Qancord.pl. **Link para descargar Perl:** <http://strawberryperl.com/>

Símbolo del sistema

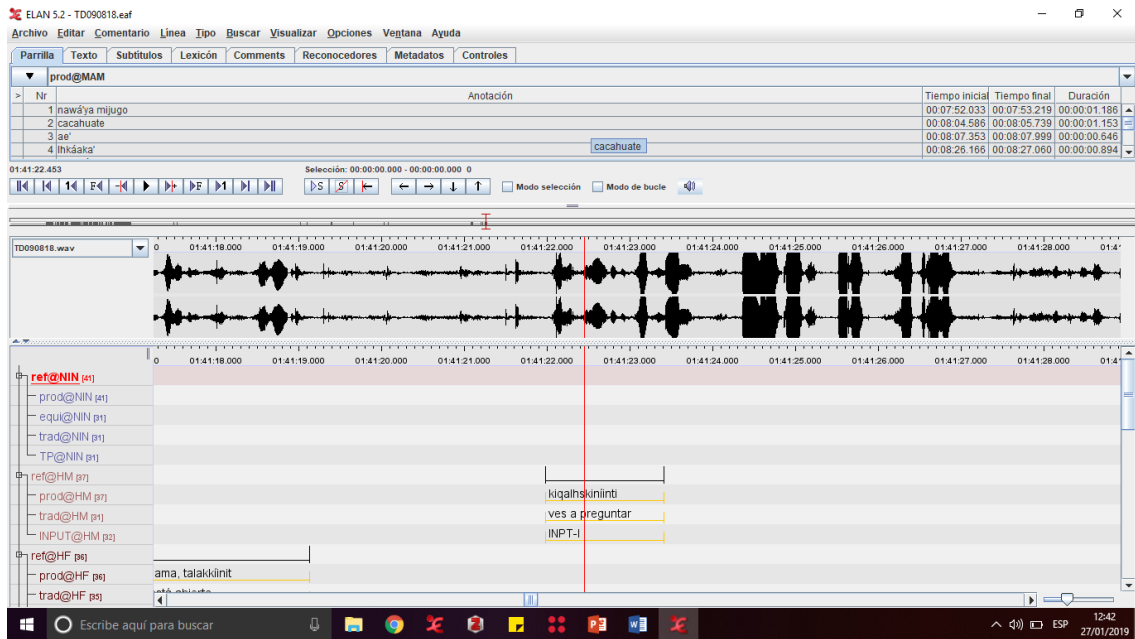
```
Microsoft Windows [Versión 10.0.17134.523]
(c) 2018 Microsoft Corporation. Todos los derechos reservados.

C:\Users\Fausto>cd folder
El sistema no puede encontrar la ruta especificada.

C:\Users\Fausto>
```

5. Una vez realizado los pasos 1, 2, 3 y 4, se abre el archivo de transcripción en ELAN.

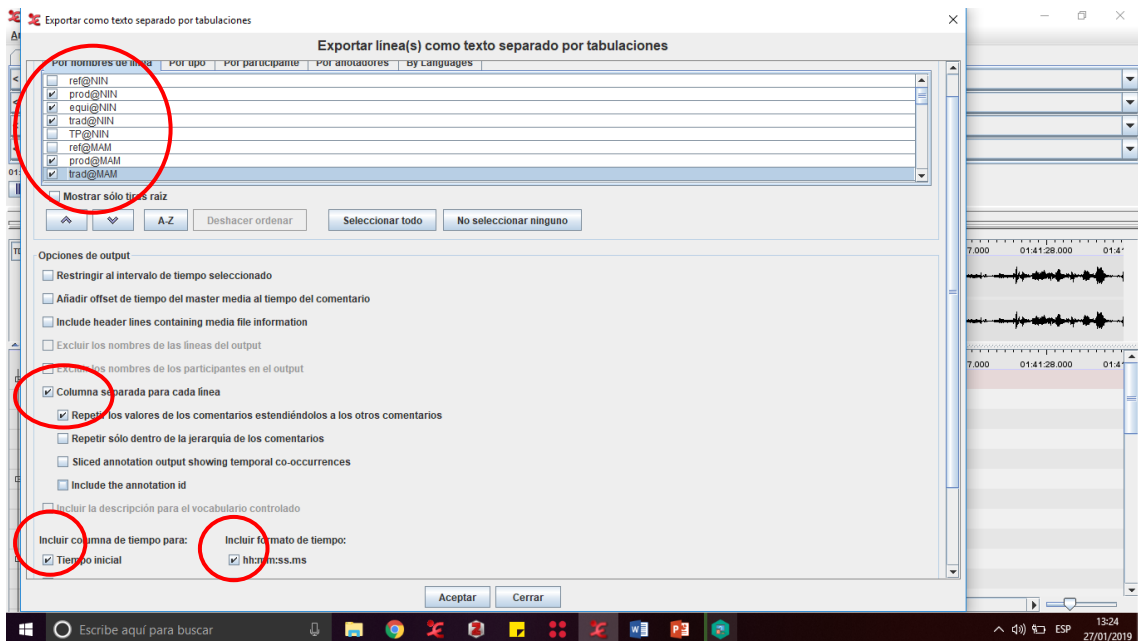
Ilustración 6. Elan



PASOS:

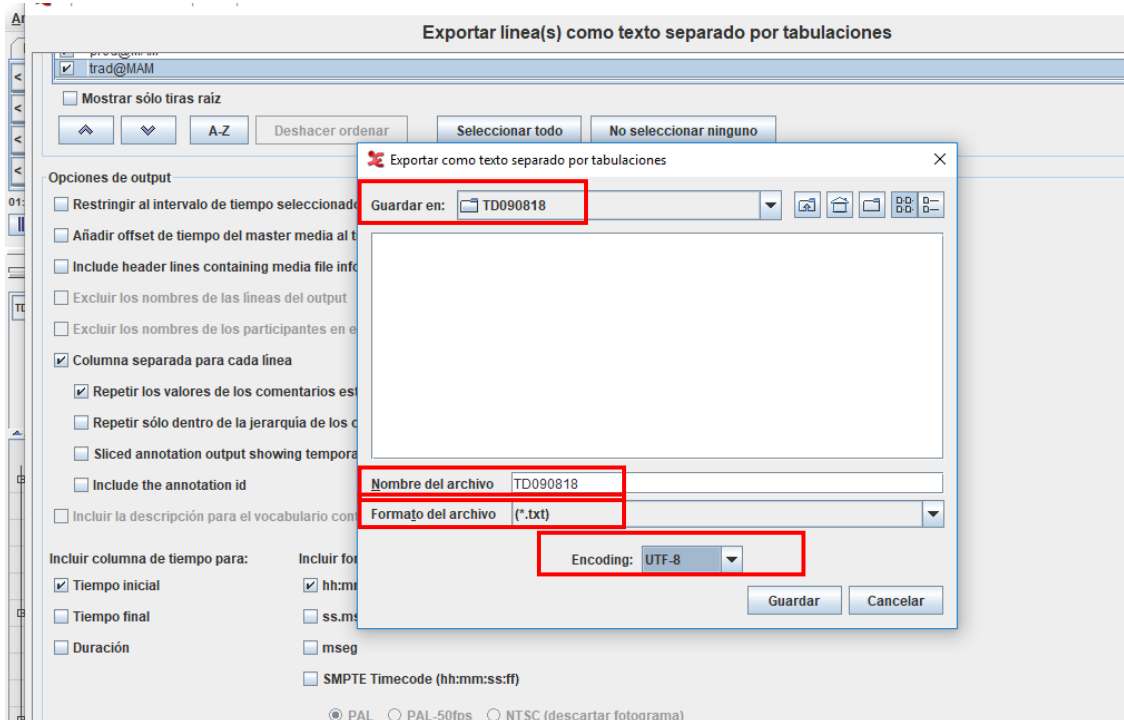
- 5.1. Ir en la pestaña de **Archivo**
- 5.2. **Exportar como** ►
- 5.3. **Texto tabulado** (se puede exportar a varios participantes)

Ilustración 7. Exportar líneas como texto tabulado



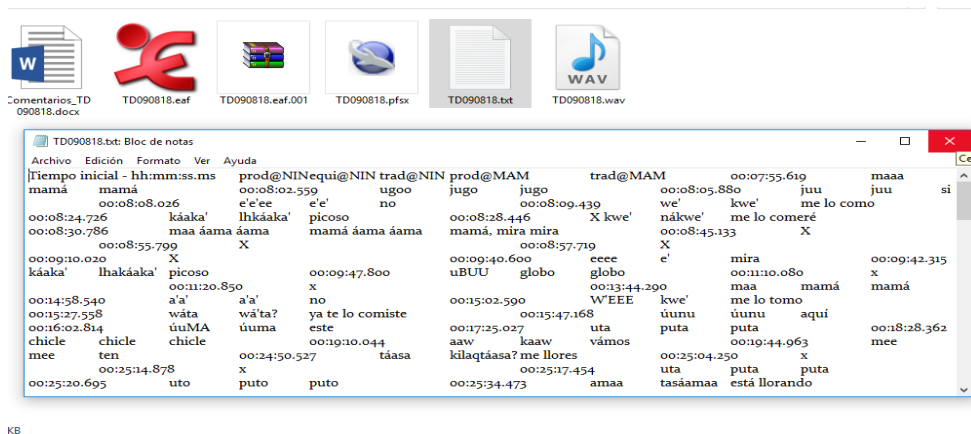
- 5.4. **Guardar en: TD090818** (carpeta de origen).
- 5.5. **Nombre del archivo: TD090818** (se usa el mismo código [nombre] del archivo de transcripción).
- 5.6. **Formato del archivo: (*.txt)**.
- 5.7. **Codificación: UTF-8**.
- 5.8. **Guardar**

Ilustración 8. Guardar líneas como texto tabulado en UTF-8



- 6. Una vez que se guarda el texto en codificación: **UTF-8**. Se abre el archivo **TD090818.txt** con **Bloc de notas**, y se guarda (nuevamente) como archivo **UNICODE⁴** (Ver pasos 6.1-6.5).

Ilustración 9. Guardar archivo UTF-8 a Unicode



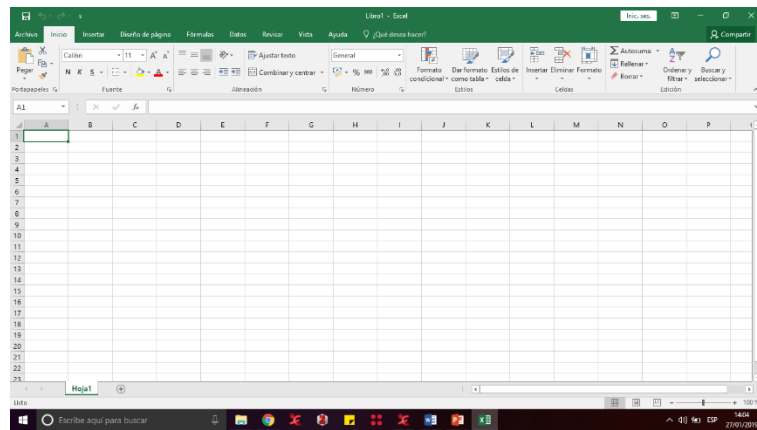
⁴ Es importante abrir el archivo **TD090818.txt** desde **Bloc de notas** y no desde **Word**. Desde un archivo **Word**, no permite guardar en la extensión **UNICODE**.

PASOS:

- 6.1. Archivo
- 6.2. Guardar como
- 6.3. **Nombre:** TD090818
- 6.4. **Tipo:** Documentos de texto (*.txt)
- 6.5. Codificación: **Unicode** (reemplazar el archivo existente)

7. Importar el archivo TD090818 a Excel

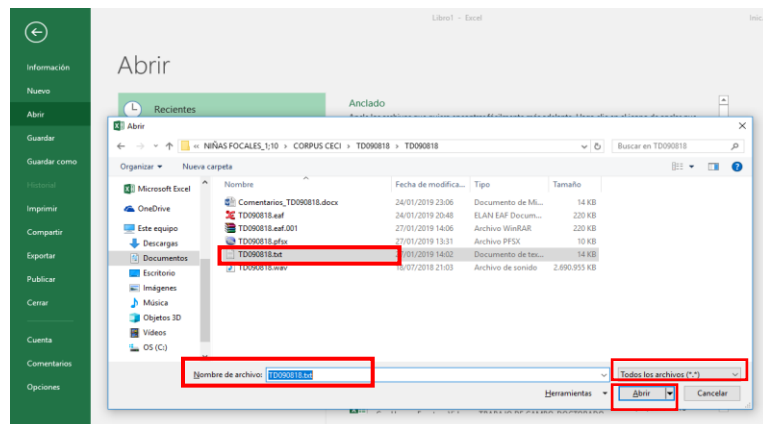
Ilustración 10. Hoja Excel



PASOS:

- 7.1. Archivo
- 7.2. Abrir
- 7.3. Examinar
- 7.4. Nombre de archivo: **TD090818.txt** (en la pestaña que está en el lado derecho, se elige **Todos los archivos (*.*)** para que aparezcan todas las extensiones de los textos).
- 7.5. Abrir

Ilustración 11. Importar archivo TD090818.txt a Excel



7.6. Asistente para importar texto - 1 de 3

Ilustración 12. Paso 1

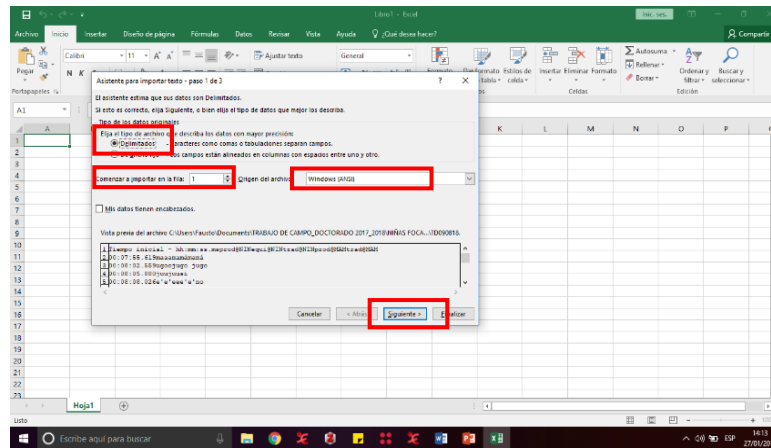


Ilustración 13. Paso 2

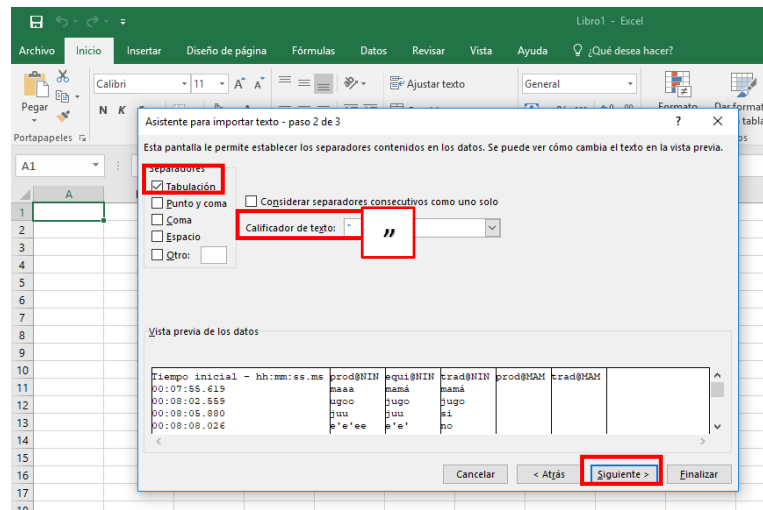
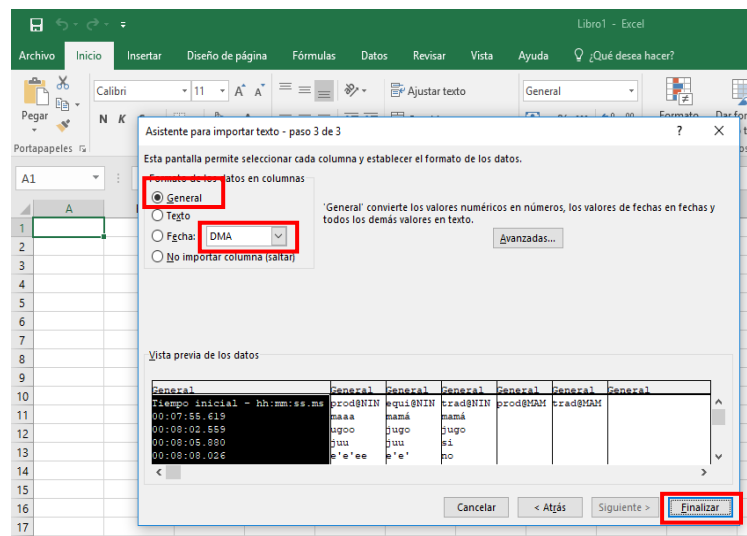


Ilustración 14. Paso 3



8. Las líneas aparecerán en sus propias columnas

Ilustración 15. Transcripción importada a Excel

A1	Tiempo inicial - hh:mm:ss.ms					
A	B	C	D	E	F	G
1	Tiempo inicial	prod@NIN	equi@NIN	trad@NIN	prod@MAM	trad@MAM
2	00:07:55.619	maaa	mamá	mamá		
3	00:08:02.559	ugoo	jugo	jugo		
4	00:08:05.880	juu	juu	si		
5	00:08:08.026	e'e'ee	e'e'	no		
6	00:08:09.439	we'	kwe'	me lo como		
7	00:08:24.726	káaka'	lhakáaka'	picoso		
8	00:08:28.446	X kwe'	nákwe'	me lo comeré		
9	00:08:30.786	maa áama áama	mamá áama áama	mamá, mira mira		
10	00:08:45.133	X				
11	00:08:55.799	X				
12	00:08:57.719	X				
13	00:09:10.020	X				
14	00:09:40.600	eeee	e'	mira		
15	00:09:42.315	káaka'	lhakáaka'	picoso		
16	00:09:47.800	uBUU	globo	globo		
17	00:11:10.080	x				
18	00:11:20.850	x				
19	00:13:44.290	maa	mamá	mamá		
20	00:14:58.540	a'a'	a'a'	no		
21	00:15:02.590	W'EEE	kwe'	me lo tomo		
22	00:15:27.558	wáta	wá'ta?	ya te lo comiste		
23	00:15:47.168	úunu	úunu	aquí		

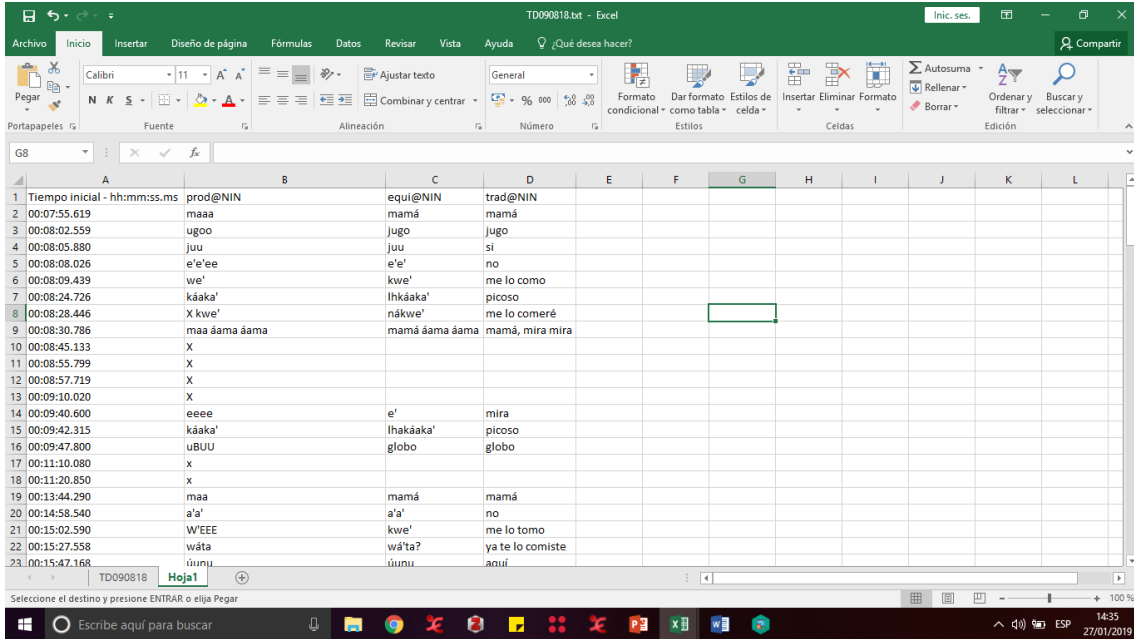
8.1. Seleccionar las columnas para cada hablante (este procedimiento es por hablante)

Ilustración 16. Selección de columnas

A1	Tiempo inicial - hh:mm:ss.ms					
A	B	C	D	E	F	G
1	Tiempo inicial - hh:mm:ss.ms	prod@NIN	equi@NIN	trad@NIN	prod@MAM	trad@MAM
2	00:07:55.619	maaa	mamá	mamá		
3	00:08:02.559	ugoo	jugo	jugo		
4	00:08:05.880	juu	juu	si		
5	00:08:08.026	e'e'ee	e'e'	no		
6	00:08:09.439	we'	kwe'	me lo como		
7	00:08:24.726	káaka'	lhakáaka'	picoso		
8	00:08:28.446	X kwe'	nákwe'	me lo comeré		
9	00:08:30.786	maa áama áama	mamá áama áama	mamá, mira mira		
10	00:08:45.133	X				
11	00:08:55.799	X				
12	00:08:57.719	X				
13	00:09:10.020	X				
14	00:09:40.600	eeee	e'	mira		
15	00:09:42.315	káaka'	lhakáaka'	picoso		
16	00:09:47.800	uBUU	globo	globo		
17	00:11:10.080	x				
18	00:11:20.850	x				
19	00:13:44.290	maa	mamá	mamá		
20	00:14:58.540	a'a'	a'a'	no		
21	00:15:02.590	W'EEE	kwe'	me lo tomo		
22	00:15:27.558	wáta	wá'ta?	ya te lo comiste		
23	00:15:47.168	úunu	úunu	aquí		

8.2. Copiar las columnas y pegar en una nueva hoja de trabajo

Ilustración 17. Columna por hablantes (niño)



8.3. Guardar como archivo de texto Unicode

PASOS:

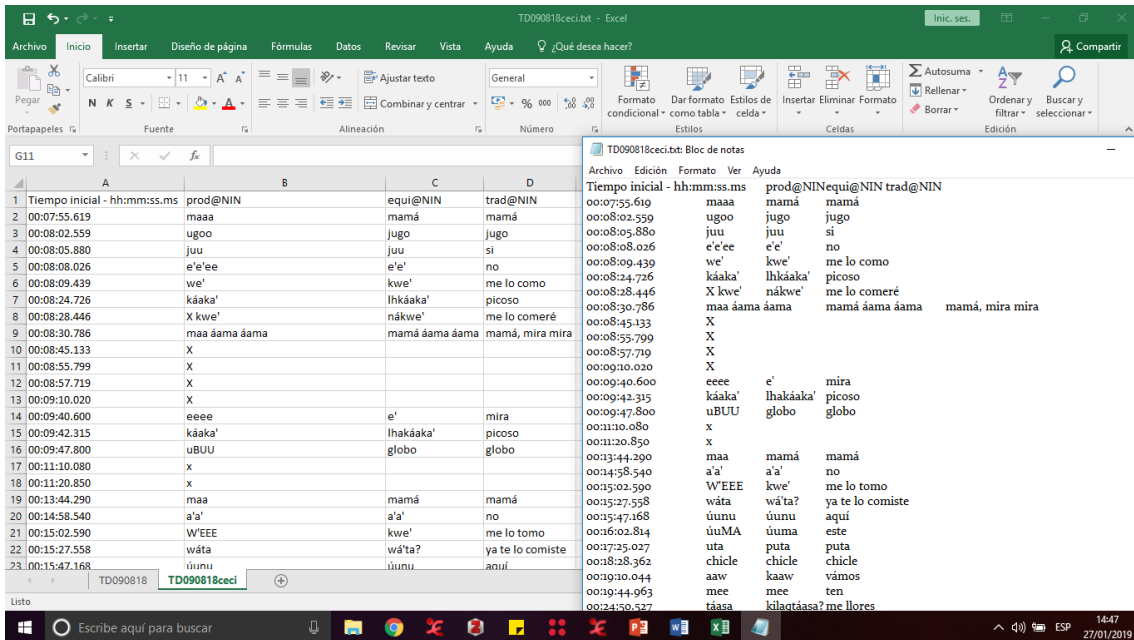
8.4. Archivo

8.5. Guardar como (lugar de origen)

8.6. Nombre del archivo: **TD090818ceci.txt** (NOTA: se le agrega el nombre del niño para distinguir del archivo anterior)

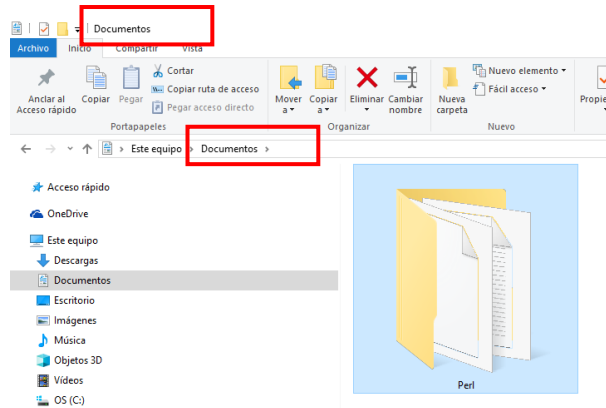
8.7. Tipo: Texto Unicode (*.txt)

Ilustración 18. Guardar archivo en Unicode



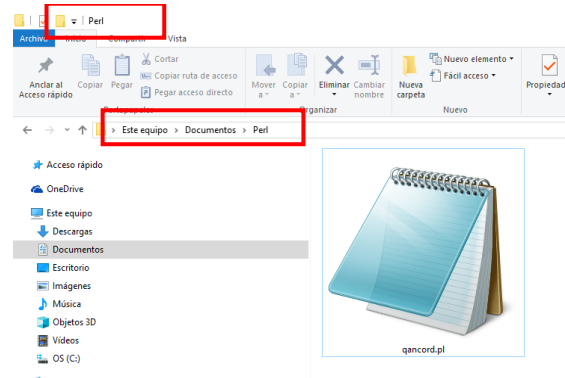
9. En la carpeta **Documentos** se crea una carpeta nueva con el nombre de **Perl**.

Ilustración 19. Creación de la carpeta Perl



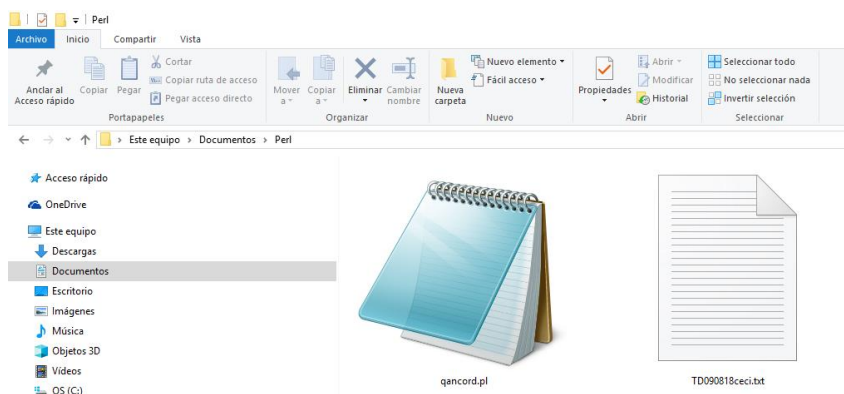
9.1. En la carpeta **Perl** se adjunta el programa **qancord.pl**

Ilustración 20. Programa qancord.pl



10. El archivo que se guardó en el formato **Unicode** con el nombre de **TD090818ceci.txt** (ver apartado 8.3), se convertirá al formato **UTF-8** y se guardará en la carpeta **Perl**.⁵

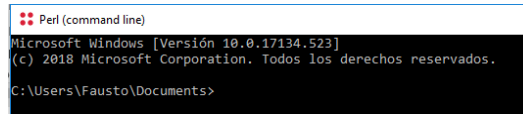
Ilustración 21. Archivo UTF-8



⁵ Sólo se crea una vez la carpeta **Perl**, en esa misma carpeta se van almacenando todos los archivos que se procesen en el programa **qancord.pl**

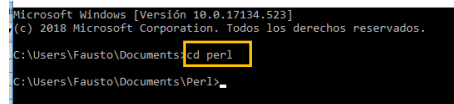
11. Abrir el programa **Perl** (doble click)

Ilustración 22. Abrir Perl



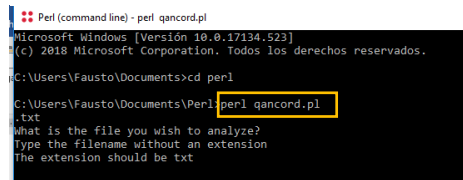
11.1. En la primera línea se escribe: **cd folder** (en minúscula)

Ilustración 23. Insertar cd folder



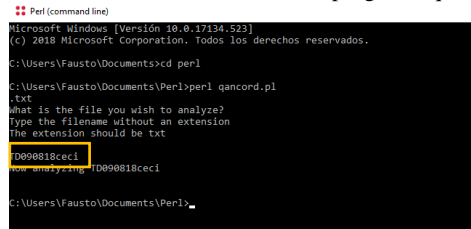
11.2. En la segunda línea se escribe: **perl qancord.pl**

Ilustración 24. Perl qancord.pl



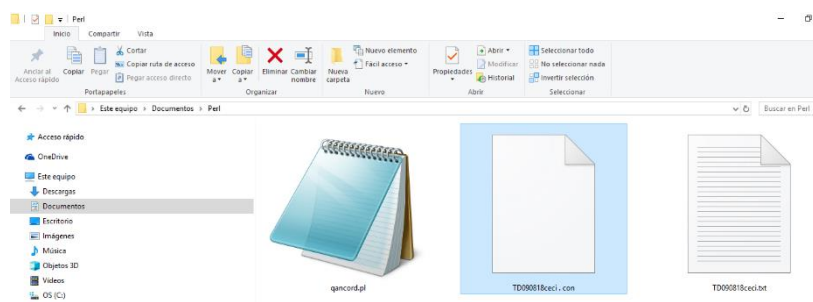
11.3. En la tercera línea se escribe: **TD090818ceci** (el nombre del archivo que procesará el programa **qancord.pl**)

Ilustración 25. Procesamiento del programa qancord.pl



12. Si el procedimiento fue satisfactorio, en la carpeta **Perl** aparecerá el archivo **TD090818ceci.con** (el archivo que procesó en el programa **qancord.pl**)

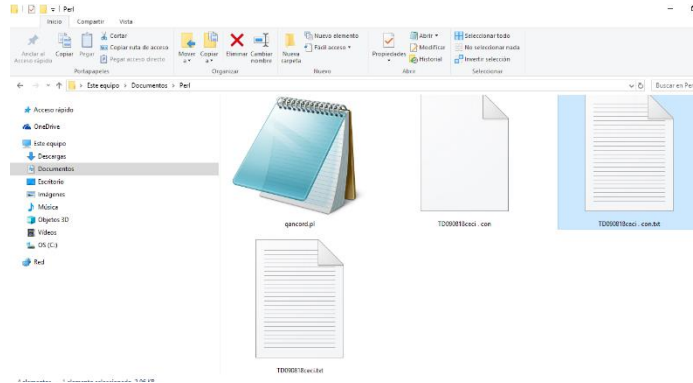
Ilustración 26. Archivo .con



MANUAL DE PROCESAMIENTO DE ARCHIVOS ELAN

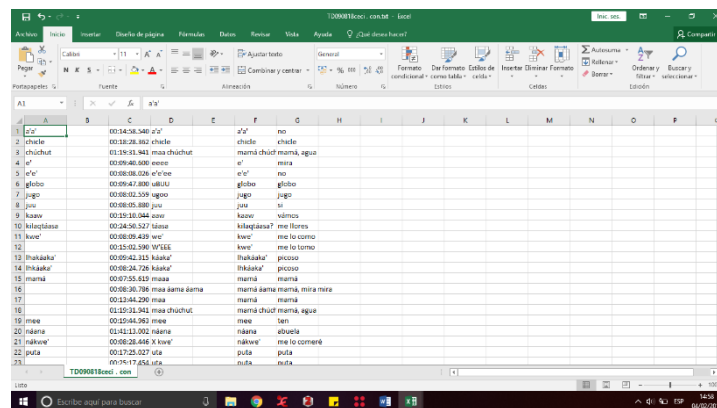
12.1. Abrir el archivo **TD090818ceci.con** con **Bloc de notas** y se guarda en el formato **Unicode**. Cuando el archivo se guarda a **Unicode**, la extensión del archivo aparecerá como **TD090818ceci.con.tx**.

Ilustración 27. Archivo .con.tx



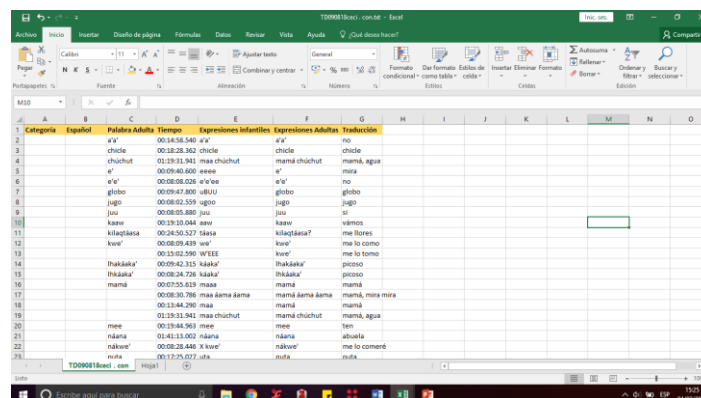
13. Importar el archivo **TD090818ceci.con.txt** a una nueva hoja de trabajo en Excel.

Ilustración 28. Importar archivo .con.txt a Excel



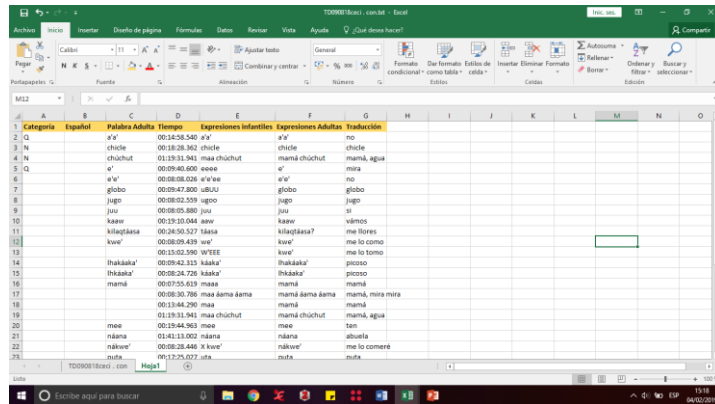
13.1. Agregar el encabezado de cada columna.

Ilustración 29. Encabezado de las columnas



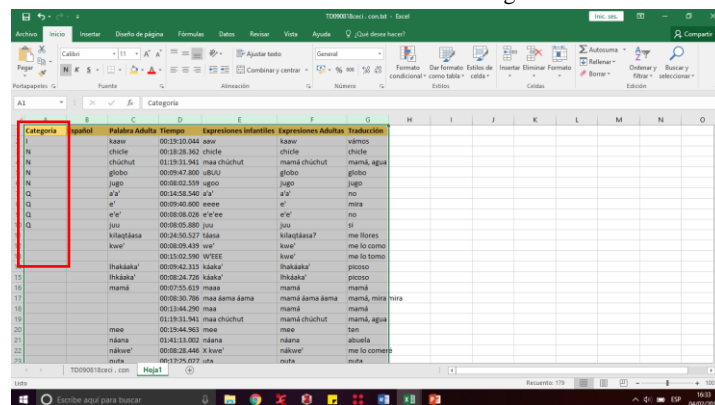
13.2. En la columna **Categoría**, se agregan los códigos que correspondan a las categorías léxicas de la **Palabra Adulta**.

Ilustración 30. Inserción de los códigos



13.3. Clasificación personalizada de los datos en las dos primeras columnas. Seleccione todas las columnas, se va a la pestaña **Ordenar y filtrar**; seleccione **Ordenar de la A a la Z**.

Ilustración 31. Ordenar los códigos



13.4. Guardar el archivo como **Tipo: Libro de Excel (*.xlsx)**

Ilustración 32. Guardar el archivo

