

TAligner 3.0

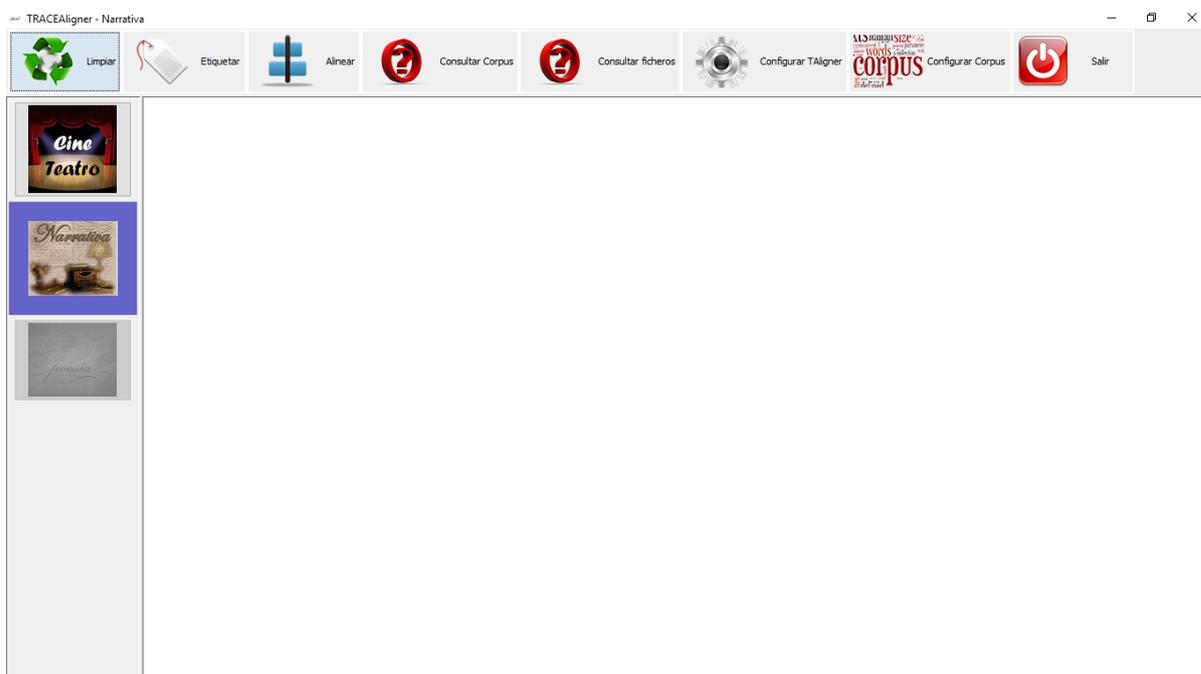
TAligner es una herramienta de alineación de textos desarrollada en el seno del grupo de investigación TRALIMA/ITZULIK (Facultad de Letras, Universidad del País Vasco, UPV/EHU), a partir de TRACE Aligner (TRACE, Universidad de León). Desarrollada en java, incluye una base de datos embebida que permite realizar consultas sobre el corpus de los textos alineados.

<https://addi.ehu.es/handle/10810/34436>

<http://corpusnet.unileon.es/herramientas-tecnicas>.

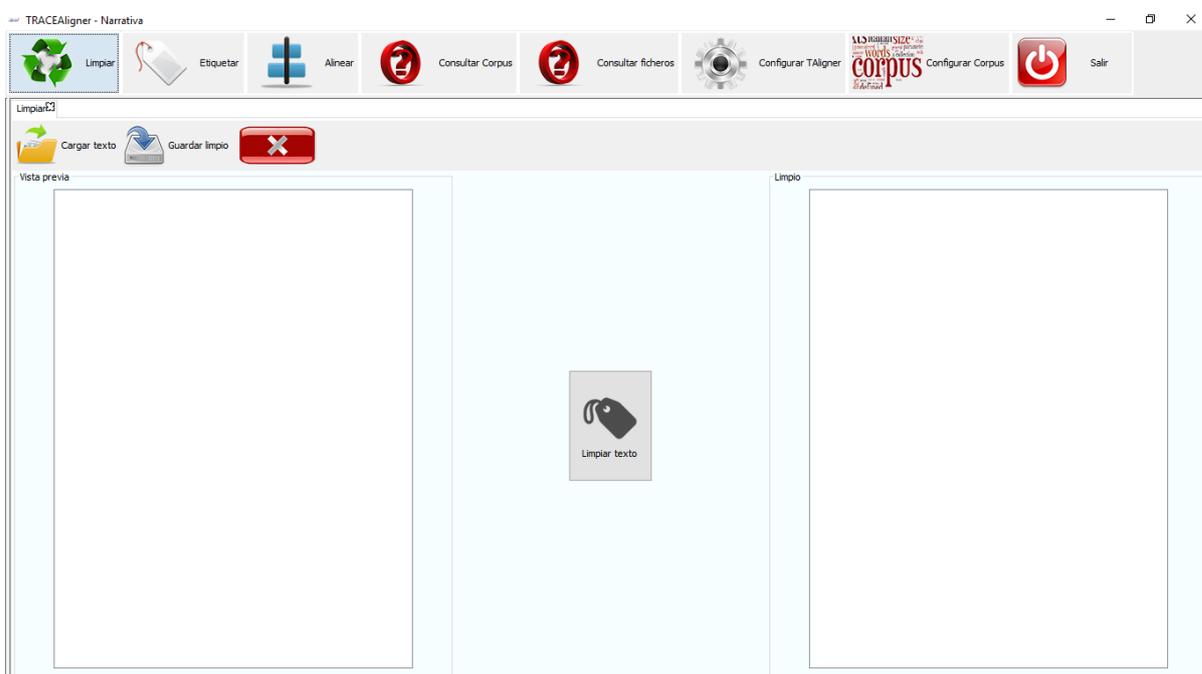
Se muestra a continuación un breve manual de usuario de la herramienta.

1- Selección del tipo de texto



Lo primero que debemos hacer al poner la herramienta en marcha, es decidir con qué tipo de texto vamos a trabajar. Este paso es muy importante, ya que los procesos de etiquetado y alineado son diferentes. Para elegir, en la parte izquierda de la pantalla se muestra el menú de selección de tipo de texto. La primera vez que usemos la herramienta, veremos que, por defecto, está escogida la opción "Narrativa". A partir de esta vez, aparecerá escogida la opción que hayamos utilizado la vez anterior. Este menú sólo se muestra cuando no hay ninguna otra ventana abierta.

2- Limpieza del texto



Es lo primero a hacer cuando vamos a trabajar con algún texto. Es un paso vital, ya que estandariza los textos de forma que luego puedan etiquetarse y alinearse con el menor número de errores posible. El proceso es sencillo:

2.1. Primero debemos pulsar sobre "Cargar texto". En la pantalla que se nos muestre, debemos escoger el texto con el que vamos a trabajar. Se nos mostrará el texto en el cuadro izquierdo de la pantalla.

2.2. En el caso de los textos de teatro, en la parte central se nos muestran tres cuadros de texto para ajustar la limpieza. En ellos, se muestran los caracteres usados por defecto para determinar dónde termina el nombre del personaje y los signos de inicio y fin de comentarios. En caso de que el texto a tratar haga uso de diferentes signos, debemos indicarlos en los cuadros antes de proceder con la limpieza.

Signo final personaje:

Signo inicio comentario:

Signo final comentario:



Limpiar texto

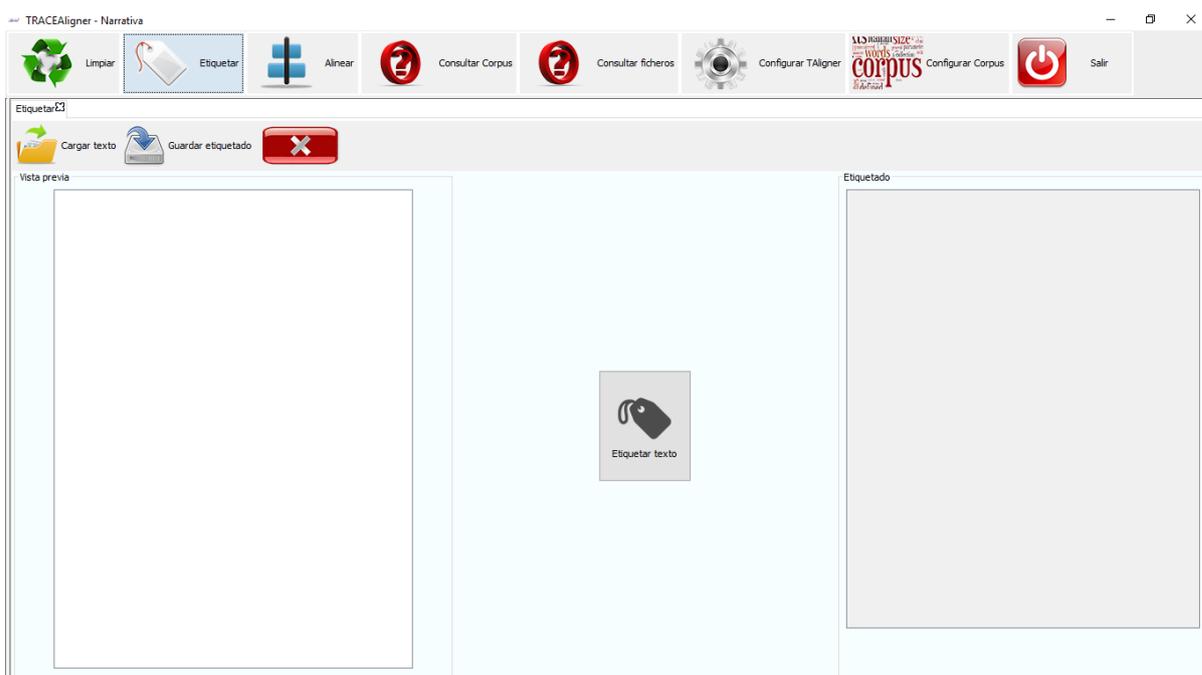
2.3. Una vez cargado, pulsaremos "Limpiar texto". El resultado se muestra en la parte derecha.

2.4. Por último, debemos guardar el texto limpio, para luego poderlo etiquetar. ...

3- Etiquetado del texto

Después de limpiar un texto, se procede al etiquetado. Es un paso necesario para luego poder alinearlo con otros textos. El proceso es idéntico al de la limpieza. Primero debemos cargar el texto, luego etiquetarlo y finalmente, guardarlo.

En la parte izquierda de la pantalla se mostrará el texto escogido y en la parte derecha, el texto una vez etiquetado.



Al guardar un texto etiquetado, se nos muestra una nueva pantalla, en la que se nos pide indicar los metadatos del texto. Estos datos son necesarios para luego poder realizar consultas, por lo que es recomendable rellenar el mayor número de campos posible.

TRACEApp

Autor:

Título:

Editorial/Exp.AGA:

Colección:

Fecha:

Idioma: ----- ▾

Traductor:

Modo: ----- ▾

Código:

SubCorpus: Sin Especificar ▾

Genero: ----- ▾

Guardar Texto

4- Alineación del texto

En este caso, necesitamos tener dos o más textos etiquetados. Comprobaremos en la pantalla que el menú tiene más opciones. Se detallan a continuación:

- **Cargar texto:** Se añade un nuevo texto a la alineación.
- **Cargar alineación:** Se carga una alineación de textos completa, con todos los textos de los que consta.
- **Guardar alineación:** Se guarda la alineación actual, con todos los textos que tenemos cargados.
- **Cargar TMX:** Permite cargar una alineación en formato TMX. Es un formato estándar pero carece de los metadatos de los textos, por lo que, se recomienda el uso del formato propio de TAligner, mediante "Cargar alineación". Cargar TMX es útil si queremos cargar textos alineados con otras herramientas.
- **Guardar TMX:** Guarda la alineación actual en formato TMX, no guardándose los metadatos de los ficheros. Sólo recomendable para exportar para su uso con otras herramientas. Si queremos guardar una alineación para continuarla en otro momento, se recomienda usar la opción "Guardar alineación".
- **Deshacer:** Cada vez que pulsemos, deshace la operación realizada anteriormente. Además



Además de las opciones mencionadas, una vez cargados los textos en la pantalla, podemos realizar las operaciones de alineado. Para ello,

pulsaremos con el botón derecho del ratón sobre el texto en el que queramos llevar a cabo la acción. Se nos mostrará un nuevo menú.

- **Combinar:** Tras seleccionar tantas celdas como queramos, si pulsamos con el botón de "Combinar", todas las celdas se convierten en una.
- **Insertar blanco:** Inserta una línea en blanco sobre la celda seleccionada.
- **Dividir:** Nos da la opción de dividir en dos el contenido de una celda. Para eso, se nos muestra una nueva pantalla con dos celdas por cada celda seleccionada, en las que debemos escribir el contenido tal y como queremos que quede.
- **Editar:** Nos permite modificar el contenido de una celda.
- **Eliminar:** Se elimina la celda seleccionada.
- **Guardar XML:** Guarda en formato XML el texto de la columna sobre la que hemos hecho click. Esto permite guardar un texto de forma individual, con todos los cambios que se hayan realizado.
- **Editar Metas:** Permite modificar los metas de un texto.
- **Añadir a corpus:** Esto sólo debería hacerse una vez finalizada la alineación. Permite añadir al corpus el texto de la columna sobre la que se ha hecho click. Es un proceso intensivo, que podría prolongarse si los textos son grandes. Debe hacerse con cada uno de los textos alineados que quieran añadirse al corpus. Aunque no es necesario, conviene añadir al corpus primero el texto original, para que luego se guarde para las traducciones cuál es su original.
- **Quitar columna:** Nos permite quitar un texto completo de la alineación.
- **Ajustar altura de líneas:** Hace que la altura de todas las líneas se ajuste a su contenido.

5- Consultar corpus

Esta pantalla nos permite realizar consultas sobre los textos que se han ido añadiendo al corpus. En la parte superior es donde indicaremos las condiciones de búsqueda. Posteriormente, pulsaremos el botón "Filtrar", lo que hará que se nos muestren los resultados, resultados que podemos exportar a un archivo Excel pulsando sobre "Exportar".

TRACEAligner - Narrativa

Limpiar Etiquetar Alinear Consultar Corpus Consultar ficheros Configurar TAligner Configurar Corpus Salir

Consultas corpus

Consultar corpus						FILTRAR
Idioma de búsqueda:	-----	Modo de búsqueda:	Exacto	Texto a buscar 1:		SALIR EXPORTAR
Modo de búsqueda:	Exacto	Texto a buscar 2:		Distancia a Texto 1:	Cualquiera	
Modo de búsqueda:	Exacto	Texto a buscar 3:		Distancia a Textos...	Cualquiera Cualquiera	
Autor:	Todos	Traductor:	Todos	Título:	Todos	
Modo:	Cualquiera	Género:	Cualquiera	Buscar traducciones:	<input checked="" type="checkbox"/> No diferenciar Ma... <input checked="" type="checkbox"/>	
Cond. Especial:	Todos	Antes y después:	0	SubCorpus:		

Nº Coincidencias: 0

6- Consultar ficheros

Esta es la pantalla para realizar consultas sobre los ficheros que deseemos. Para ello, sólo debemos ir cargando los ficheros deseados.

The screenshot shows the TRACEAligner - Narrativa software interface. The title bar indicates the application name and standard window controls. The main menu bar includes options: Limpiar, Etiquetar, Alinear, Consultar Corpus, Consultar ficheros (highlighted), Configurar TAligner, Configurar Corpus, and Salir. Below the menu bar, there is a 'Cargar textos' button with a red 'X' icon. The interface is divided into two main sections: 'Ficheros seleccionados' on the left, which contains a list box for files, and 'Resultado' on the right, which displays a table of search results. The table has four columns: 'Palabra', 'Cantidad', 'Palabra', and 'Cantidad'. Below the table, summary statistics are shown: Total palabras: 0, Total palabras diferentes: 0, Media aparición por palabra: 0.0, Palabras que cumplen filtro: 0, and Media: 0. On the far right, there are checkboxes for 'Diferenciar may/min:' and 'Quitar números:', and input fields for 'Mínimo de repeticiones:' and 'Máximo de repeticiones:'. At the bottom right, there are 'Consultar' and 'Exportar' buttons.

Palabra	Cantidad	Palabra	Cantidad
---------	----------	---------	----------

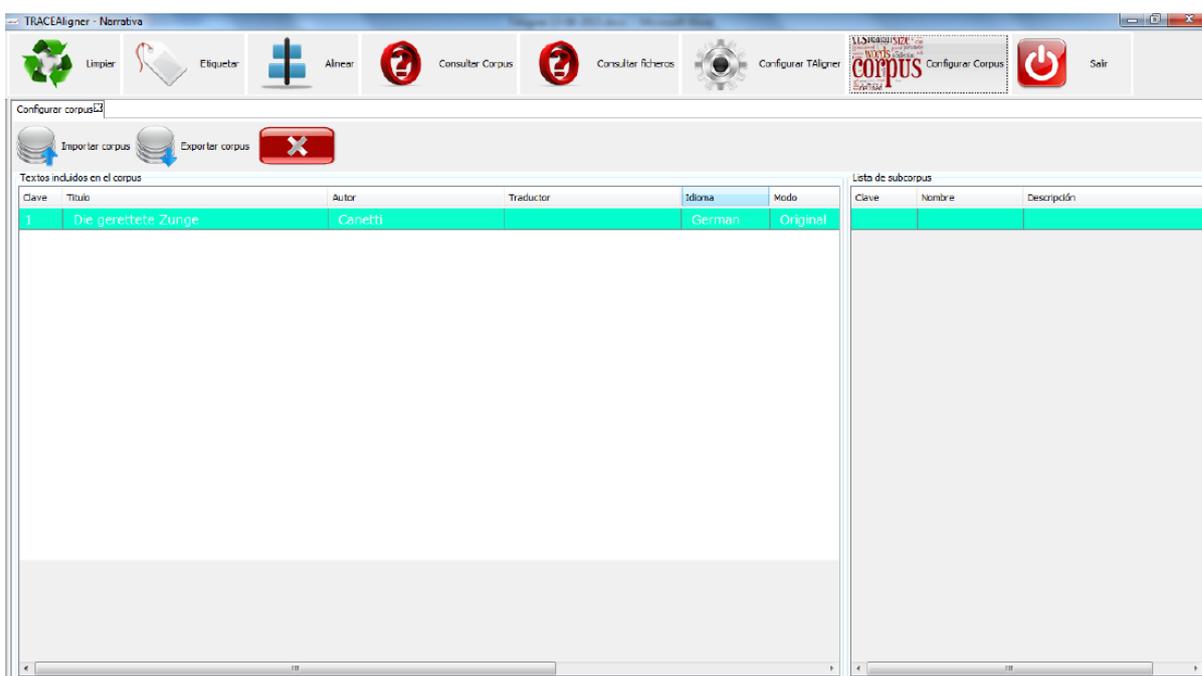
Total palabras: 0
Total palabras diferentes: 0
Media aparición por palabra: 0.0
Palabras que cumplen filtro: 0
Media: 0

Diferenciar may/min:
Quitar números:
Mínimo de repeticiones:
Máximo de repeticiones:

Consultar
Exportar

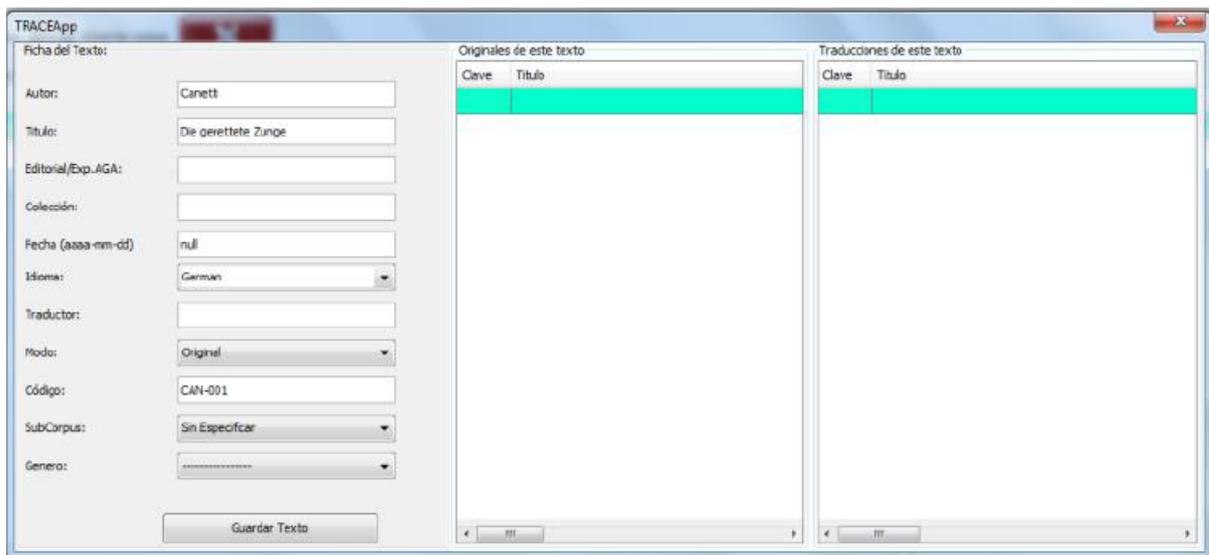
7- Configurar Corpus

En la pantalla de "Configurar corpus" aparecen dos tablas. En la tabla derecha (la pequeña) se ve la lista de subcorpus que hemos creado hasta ahora. Esto nos sirve por si queremos agrupar textos y luego poder filtrar las consultas. La manera de crear o eliminar un subcorpus es haciendo click con el botón derecho del ratón sobre las líneas de la tabla.



En cuanto a la tabla de la izquierda, la más importante, se ven los textos que han sido añadidos al corpus. Los textos se añaden desde la ventana de alineación. Se puede quitar un texto del corpus haciendo click con el botón derecho del ratón sobre el texto y escogiendo la opción de eliminar.

Tanto haciendo doble click sobre un texto como haciendo click con el botón derecho y escogiendo "Editar Metadatos y relaciones", llegamos a una nueva pantalla:



Aquí podemos modificar cualquier metadato de un texto que ya esté en el corpus, sin tener que quitarlo, y volverlo añadir tras modificar los metadatos en la ventana de alineación. Pero además, y esto es lo más importante, desde aquí se indican las relaciones entre los textos.

- Tenemos dos tablas en la ventana. Una con el título "Originales de este texto" y otra "Traducciones de este texto". En la primera podemos añadir (o eliminar) tantos textos como deseemos marcar como originales del texto que tenemos en pantalla. Si está traducido de un solo texto, ponemos uno, si se ha hecho a partir de varios, indicamos todos ellos. Y lo mismo en la segunda tabla. Podemos indicar qué textos se han traducido a partir del que tenemos en pantalla.

- No es necesario repetir las relaciones, es decir, si tenemos dos textos: "Texto1" y "Texto2", y el primero es original del segundo, podemos entrar en la ficha del primero e indicar que el segundo es una traducción suya o entrar en la ficha del segundo e indicar que el primero es un original suyo. Tanto si lo hacemos de una forma como si lo hacemos de la otra, comprobaremos que la relación aparecerá en la ficha de ambos textos.