

¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fB10Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

En esta breve guía explico cómo lo hago yo y qué herramientas utilizo para ello.

Para empezar, conviene aclarar primero algunos puntos básicos:

- Cada oración en el sitio de tatoeba tiene asignada una referencia única, que es lo que permite identificarla y utilizarla para cualquier operación.
- Los archivos de audio, como es lógico, también tienen una referencia única.
- Para poder enlazar una frase con su archivo de voz correspondiente, es necesario que, tanto el nombre de la frase como el archivo de voz, contengan el mismo número de referencia; si no, el enlace habría que hacerlo manualmente y eso llevaría mucho tiempo.
- Así pues, para que puedan conectarse automáticamente ambos archivos: la frase escrita y el audio correspondiente, lo ideal es que ambos se denominen siguiendo el mismo esquema:

[número de referencia] [espacio] [guión] [espacio] [título]

Ej.

Identificación de la frase:	6640144 - Vine en cuanto se me dijo.
Identificación del archivo de sonido:	6640144 - Vine en cuanto se me dijo.flac

Según lo anterior, ¿puedo usar cualquier programa o dispositivo para grabar mis frases?

La respuesta es sí. Se pueden crear archivos de voz para las frases con cualquier programa o dispositivo de grabación, siempre que tanto la frase como su audio correspondiente mantengan la información que se ha dicho (*son preferibles archivos de sonido de alta calidad: WAV, AIFF o FLAC*).

Sin embargo, aunque se puedan usar otros programas para grabar las frases, no con todos los programas es fácil asignar de forma automática la información necesaria para poder enlazarlos después y entonces hay que hacerlo de forma manual (*no resulta muy cómodo renombrar a posteriori manualmente cada uno de los archivos de sonido obtenidos*).

Shtooka Recorder es un pequeño programa lingüístico gratuito, concebido específicamente para la grabación de palabras y frases, que asigna automáticamente todos los datos necesarios a los archivos creados; es por eso que se recomienda su uso.

Una aplicación alternativa, también gratuita, sería Audacity, pero no es tan cómoda para trabajar. Más información en este enlace [<http://www.manythings.org/tatoeba/audacity.html>].

¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBl0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

Bueno, pues aclarado esto, paso a explicaros cómo lo hago yo.

Yo utilizo estos dos programas gratuitos:



-Shtooka Recorder: Con él grabo las frases y las archivo directamente en una carpeta.

Enlace para descarga:

[http://web.archive.org/web/20110621160436/http://shtooka.net/soft/kit_shtooka/kit_shtooka_Install_0.9.8.exe]



-VLC Media Player: Con este otro escucho las frases que ya he grabado y compruebo que están bien antes de publicarlas.

Enlace para descarga:

[<https://www.videolan.org/vlc/index.es.html>]

¿Cómo lo hago y cuánto tarda?

Con el procedimiento que describo detalladamente en las páginas siguientes, habitualmente concluyo la grabación y envío para su posterior publicación de unas 50 frases en aproximadamente 20 minutos.

¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fB10Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

1. Primeramente, preparo una lista con las oraciones que voy a grabar, que, como se ha explicado, debe incluir su referencia numérica en la base de datos de la plataforma, pues es su única parte invariable y la que permitirá identificarla para enlazar posteriormente las grabaciones con su texto correspondiente.

(Este listado puede proporcionarlo CK o se lo puede preparar uno mismo [ver cómo en Anexo II]).

Lista para grabar

6627094 - Saludé al policía.
6629101 - El accidente fue muy grave.
6629141 - Iba por una autopista.
6630302 - Yo que tú estudiaría.
6630337 - Le dijiste que sí.
6632806 - Lo he hecho fatal.
6632810 - Ese vestido te está ideal.
6632850 - Al llegar, no vi a nadie.
6632853 - Se me cayó al cogerlo.
6632874 - Yo no lo veo tan claro como tú.
6632923 - En mi zona apenas hay viento.
6632934 - Un panel solar sólo funciona por el día.

Fíjate en que hay un espacio antes y otro después del guión

Parte 1: Número de referencia Parte 2: Guión de separación
(OJO: guión normal, no rayita larga) Parte 3: Texto de la frase a grabar

Yo suelo trabajar con tandas de 50 o 100 frases, que es un volumen cómodo para detectar errores y para su archivo posterior.

En una sesión normal me hago unas 200 o 300 frases en total.

¿Cómo poner voz a las frases en Tatoeba?

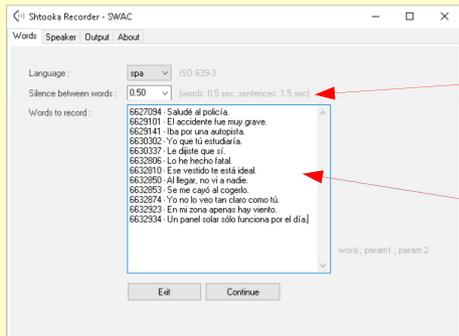
-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBl0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

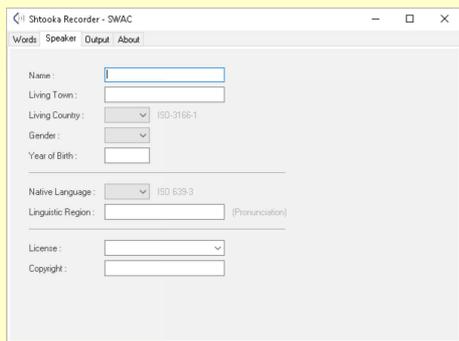
2. Inicio el programa Shtooka.

La máscara de este programa, al abrirlo, ofrece las siguientes 4 pestañas de selección.

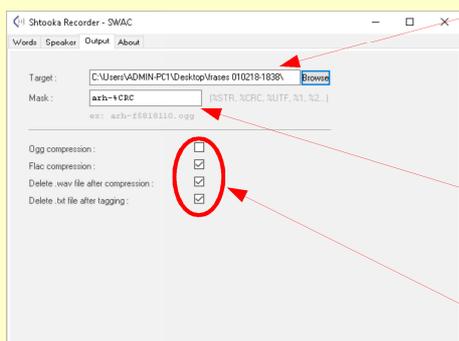


● En la sección 'Words' podemos cambiar el tiempo deseado de silencio entre palabras. El programa usa esta información para detectar cuándo se ha acabado de enunciar la frase (salvo que el texto sea muy largo, 0,50 está bien y no hará falta modificarlo).

● En su caja central 'Words to record' es donde colocaremos las palabras o frases que deseamos grabar. Se pueden escribir una a una, o también pegarlas directamente desde el archivo de texto donde esté el listado que queremos grabar.



En esta sección 'Speaker' no necesitamos modificar nada.



● La sección 'Output', en su caja 'Target', nos permite buscar y elegir, con su botón 'Browse', la carpeta donde se guardarán los archivos de voz que se generan (yo uso una carpeta de destino distinta para cada lote que grabo, y la nombro con un nombre sencillo que informa del contenido, por ejemplo la fecha: 010218-1838, porque cuando se van acumulando las grabaciones, me resulta más fácil para su localización y control posterior).

● En su caja 'Mask' debemos cerciorarnos de que ponemos nuestro nombre de usuario en Tatoeba ante la secuencia (-%CRC), pues servirá posteriormente para saber quién ha grabado la frase, que puede ser una persona distinta a quien la escribió originalmente.

● Por último, como el proceso de grabación genera diferentes archivos que no necesitamos, conviene mantener marcadas las tres casillas tal cual aparecen aquí.



Esta última sección 'About' es meramente informativa y no necesitamos modificar nada.

¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

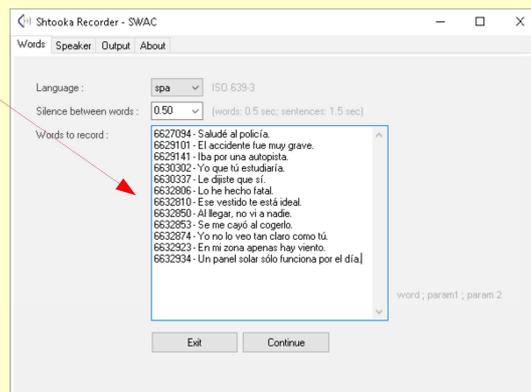
Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

3. Una vez que he modificado los parámetros necesarios, copio las frases de mi lista (*directamente, con el botón derecho del ratón*) y las pego en la caja principal 'Words to record' de la pestaña 'Words'.

Lista para grabar

6627094 - Saludé al policía.
6629101 - El accidente fue muy grave.
6629141 - Iba por una autopista.
6630302 - Yo que tú estudiaría.
6630337 - Le dijiste que sí.
6632806 - Lo he hecho fatal.
6632810 - Ese vestido te está ideal.
6632850 - Al llegar, no vi a nadie.
6632853 - Se me cayó al cogerlo.
6632874 - Yo no lo veo tan claro como tú.
6632923 - En mi zona apenas hay viento.
6632934 - Un panel solar sólo funciona por el día.



¿Cómo poner voz a las frases en Tatoeba?

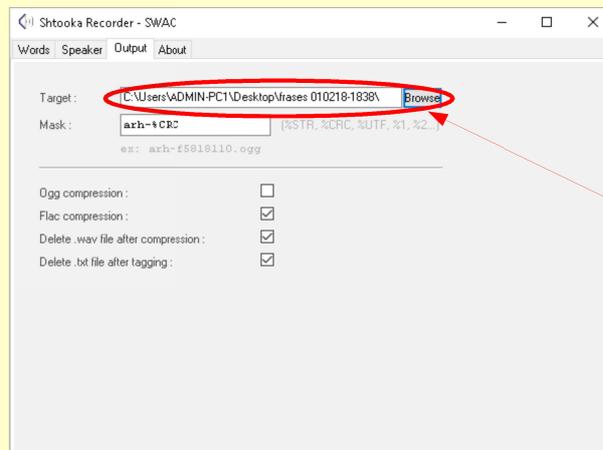
-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

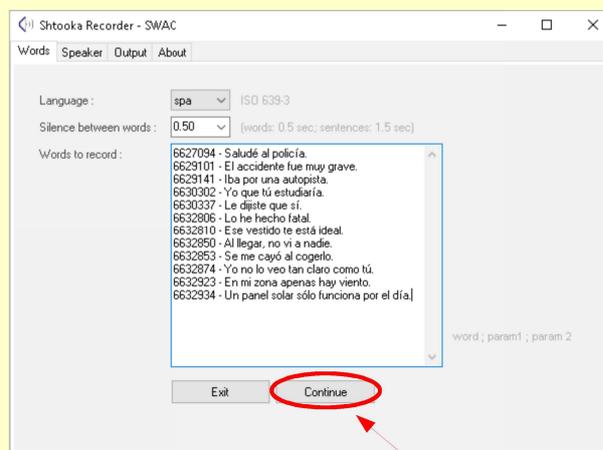
4. Luego, en la pestaña '**Output**', usando el botón '**Browse**', busco y selecciono la ubicación de mi carpeta donde deseo que se guarden los archivos que vaya generando al grabar.

(Yo guardo cada lote de 50/100 frases en una carpeta distinta, porque, si no, se irían acumulando en la misma carpeta todas las tandas sucesivas de grabación y el archivo acabaría siendo incontrolable).



A esta carpeta irán a parar los archivos de audio que genera el programa, hasta que yo cambie el destino.

Cuando tengo seleccionada la carpeta, vuelvo a la pestaña de '**Words**' y, pulsando el botón inferior '**Continue**', se abre automáticamente la herramienta propia de grabación que incluye el programa. (Si no ves estos dos botones inferiores, es que necesitas ampliar la carátula hacia abajo para verlos)



Pulsando el botón 'Continue' se abre la herramienta de grabación.

¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

Al abrirse esta herramienta, aparecen dos nuevas máscaras:

-'Input graph', que nos ofrece un gráfico del sonido, conforme grabamos:

Márgenes temporales previo al inicio y posterior a la grabación de la frase. Deben respetarse para que no resulte truncada la grabación.

Importante comprobar el volumen de grabación.

Conforme grabamos la frase, se visualiza el gráfico de nivel.

Es importante que el nivel de grabación sea adecuado, ni muy alto ni muy bajo, pues el programa detecta automáticamente, mediante las variaciones de la voz, cuándo se empieza a hablar y cuándo se acaba y, si el volumen es muy bajo, puede terminar la grabación antes de tiempo.

-Y 'Record list', que nos muestra todas las frases y nos permite adecuar algunos parámetros mediante el botón 'Record / Recording settings':

Botón de grabación; se activa con la barra espaciadora.

La frase activa para grabar se muestra en color diferente.

Valores de grabación recomendados.

Este contador nos indica el número de orden de la frase que estamos grabando.

Los parámetros que aquí aparecen serán adecuados para la mayoría de las grabaciones y no hará falta cambiarlos en cada nueva sesión.

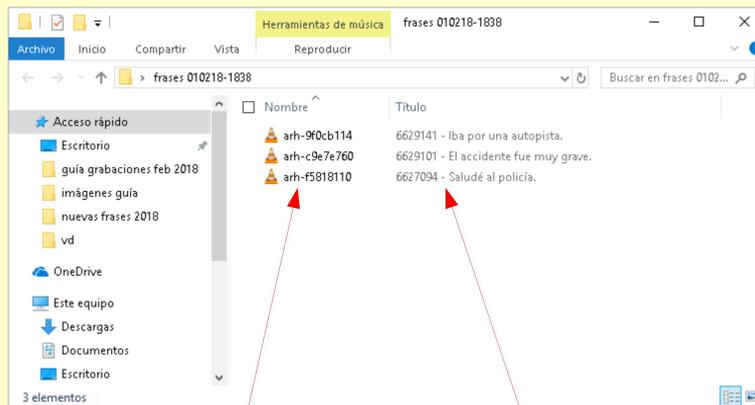
¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

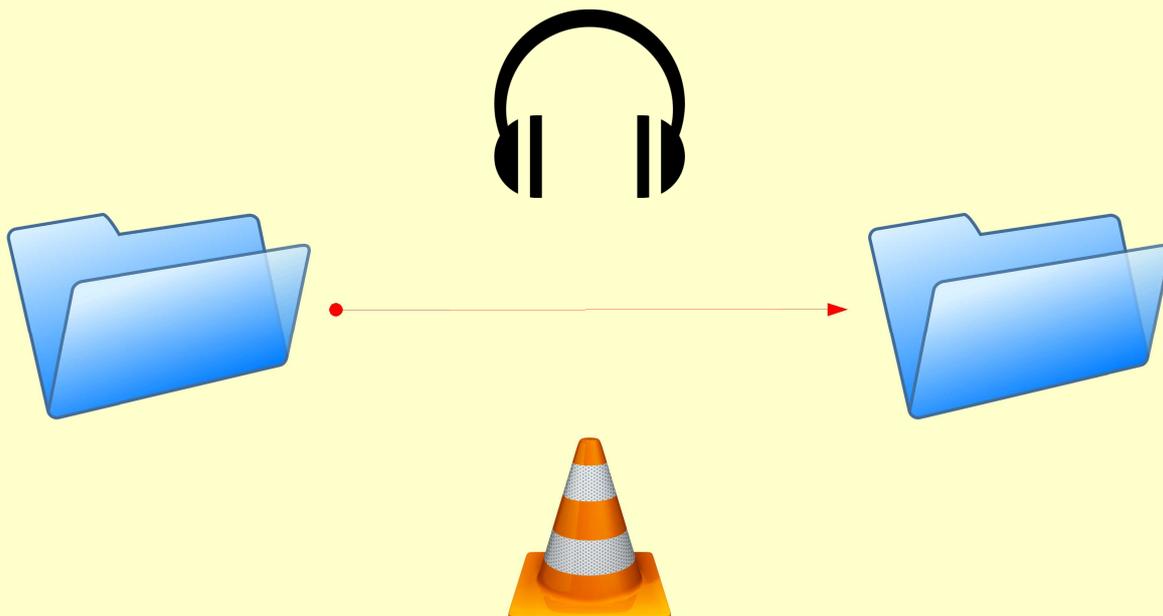
Conforme voy concluyendo cada frase, el correspondiente archivo de audio ya aparece en la carpeta de destino que elegí.



Los archivos de audio ya incorporan mi nombre de usuario.

Los archivos de audio creados se corresponden con la oración e incluyen su número de referencia en la base de datos de Tatoeba, con el esquema
[número de referencia] [espacio] [guión] [espacio] [título]

5. A continuación, yo escucho cada uno de los archivos de audio y paso a una carpeta definitiva los que voy comprobando que están correctos. Para ello uso el programa VLC. Si detecto algún problema, lo borro y simplemente vuelvo a grabar la frase.



¿Cómo poner voz a las frases en Tatoeba?

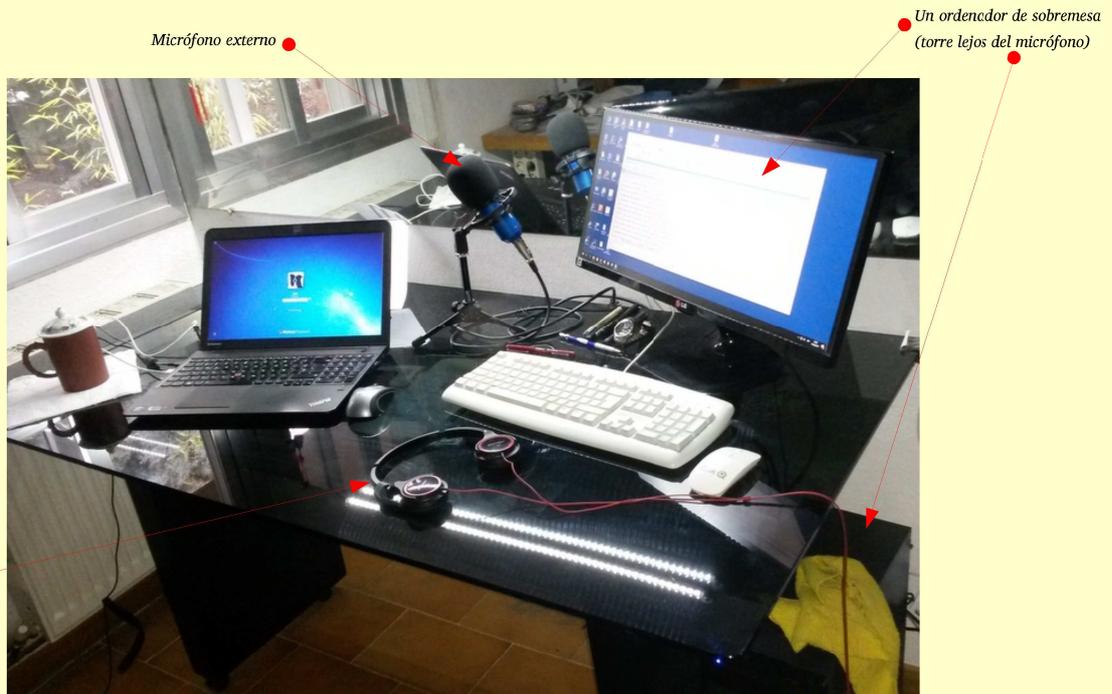
-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fB10Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

6. Finalmente, comprimo la carpeta en formato .rar 🗄️ o .zip 📦 y se la envío a CK para que la publique.

Éste es el equipo que utilizo para la grabación



¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBl0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

- ANEXO I -

Algunas situaciones que pueden darse y sus posibles soluciones:

<i>Problema</i>	<i>Posibles soluciones</i>
Ruido de fondo perceptible en las grabaciones.	<p>Distancia al micrófono:</p> <ul style="list-style-type: none">-Separar el micrófono del ordenador y comprobar el resultado. <p>Ajustes de sonido en el ordenador:</p> <ul style="list-style-type: none">-En función del equipo y sistema operativo, es posible ajustar diversos parámetros en el ordenador. Probar con ajustes distintos y comprobar el resultado. (Ver imágenes de ajustes)
Grabación muy baja, que me obliga a elevar el nivel del sonido y entonces tengo ruido de fondo.	<p>Orientación del micrófono al grabar:</p> <ul style="list-style-type: none">-La membrana, que suele ser plana y no se ve, al estar oculta tras la rejilla metálica o la espuma paravientos, tiene que estar prefentemente en posición perpendicular a la fuente de sonido.-Girar y colocar el micrófono en distintas posiciones y comprobar el resultado. (Ver imágenes)-Ajustar parámetros de sonido en configuración de micrófono del sistema operativo. (Ver imágenes)
Soplidos al grabar.	<ul style="list-style-type: none">-Colocar una esponja o paravientos.-Alejarse del micrófono.
Se truncan mis frases al grabarlas.	<ul style="list-style-type: none">-Vigilar el nivel de sonido durante la grabación. Un nivel bajo desactiva la grabación.-Comprobar que se respetan los márgenes antes y después de la grabación.
La grabación se para antes de tiempo con frases largas.	<ul style="list-style-type: none">-Modificar los parámetros de 'Final silence'.
En frases con ';' no veo la parte que sigue a este símbolo al grabar.	<ul style="list-style-type: none">-El programa entiende este símbolo ';' como un límite para la visualización de la frase y por eso no la

¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

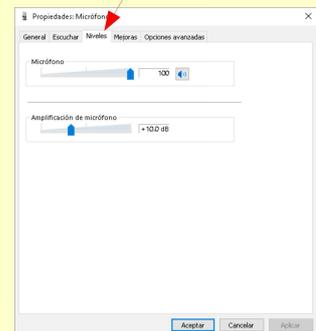
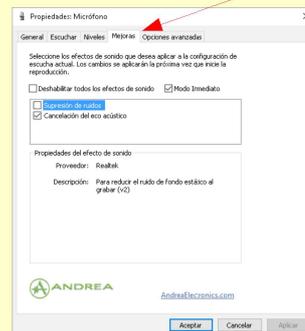
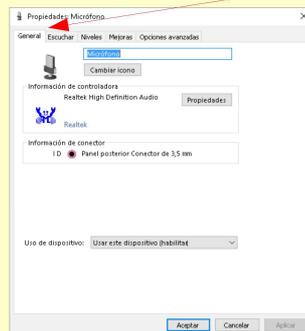
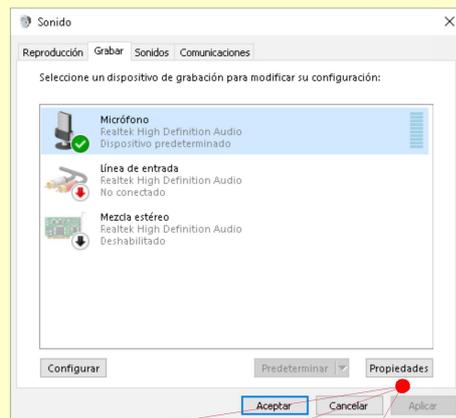
Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

muestra. Aunque no se vea esta parte, puede grabarse normalmente toda la frase (habrá que leer el resto de otra fuente).

Los micrófonos para grabar voz suelen ser direccionales y tienen una membrana plana dentro de la rejilla. Hay que procurar hablar perpendicularmente a dicha membrana para que capte bien el sonido.



El ordenador permite realizar una serie de ajustes para mejorar el rendimiento de los dispositivos de sonido como los micrófonos conectados a él, mediante las opciones en propiedades (la nomenclatura puede variar según el sistema operativo).



¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

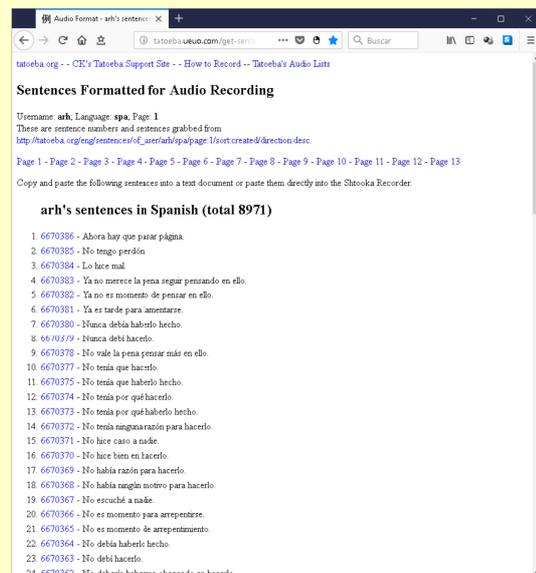
Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBl0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

- ANEXO II -

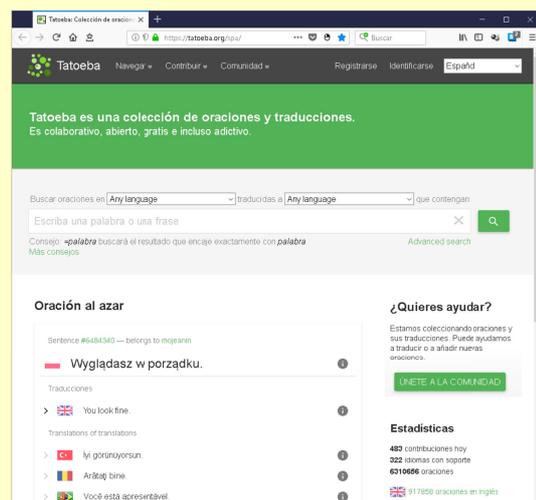
¿Cómo obtengo una lista de mis frases para grabar?

La forma más fácil, sin duda, es pedirle a CK que nos la prepare y así tendremos acceso a nuestras frases a través de una página similar a ésta:



Sin embargo, también podemos hacernos nosotros nuestro propio listado de frases siguiendo unos cuantos pasos sencillos que no llevan mucho tiempo, especialmente si se trata de una lista con sólo unas pocas frases para grabar.

Primeramente, accedemos a la página principal de Tatoeba (tatoeba.org)



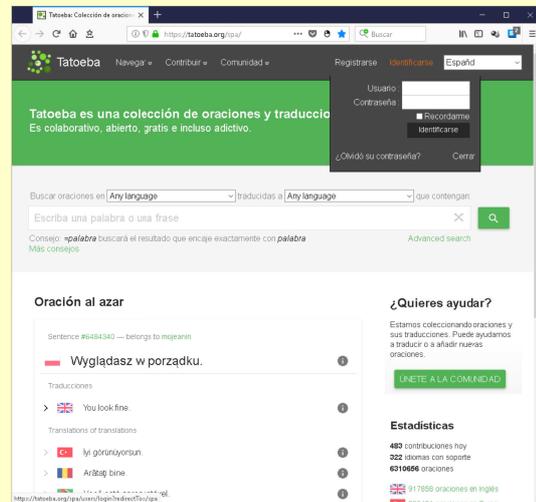
¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

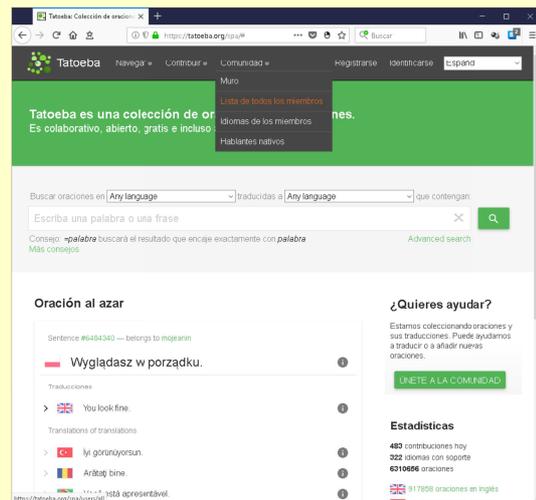
Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBl0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

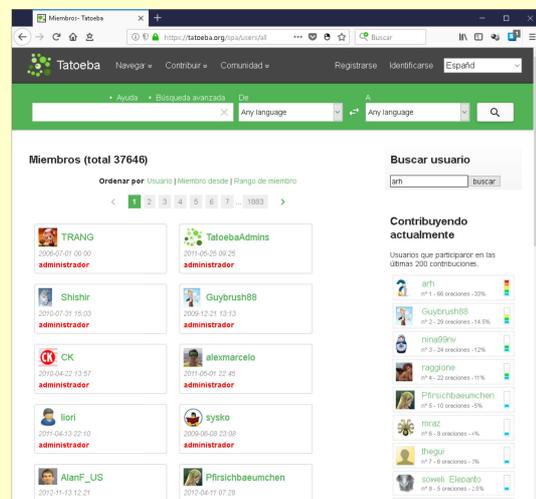
Para obtener este listado, **NO** nos vamos a identificar, porque cuando entramos como usuarios al sitio, ya no se ven los números de referencia de las frases.



Podemos acceder a nuestras frases de otra forma, a través del listado de miembros. Para ello, vamos a la pestaña 'Comunidad' y seleccionamos la opción 'Lista de todos los miembros'.



En esta otra página introducimos nuestro nombre de usuario en la caja de búsqueda y pulsamos en 'buscar'.



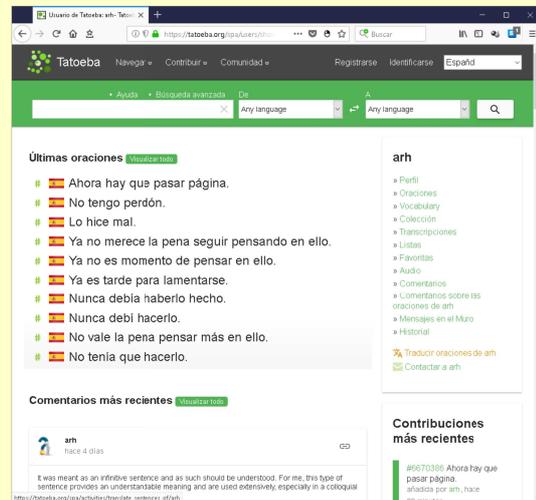
¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

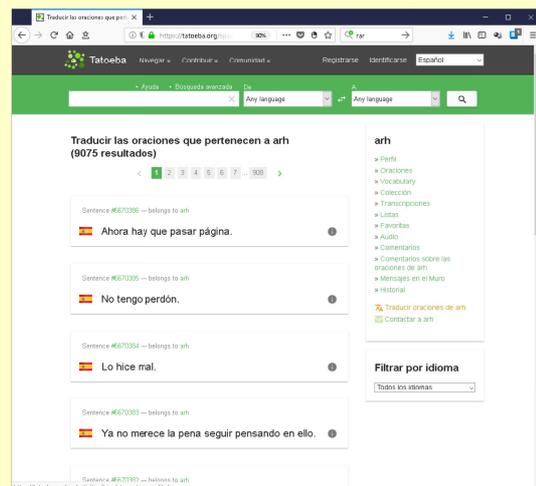
Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

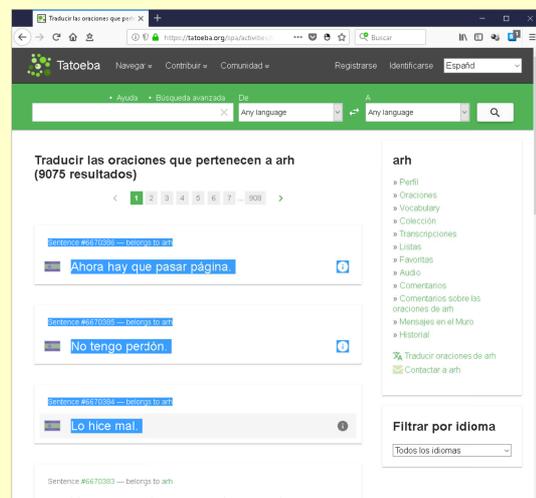
Llegamos, así, a nuestra página personal o a la del usuario a cuyas frases queremos poner voz. Aquí seleccionamos la opción 'Traducir oraciones de ...'



Ya podemos ver el listado completo de nuestras frases, con su correspondiente número de referencia en la base de datos. Aunque, como se puede apreciar, hay más información de la que necesitamos y la eliminaremos posteriormente.



Seleccionamos todas las frases que deseemos recuperar en cada página y las copiamos (con el botón derecho del ratón). No nos olvidemos de incluir el número de referencia de cada frase.



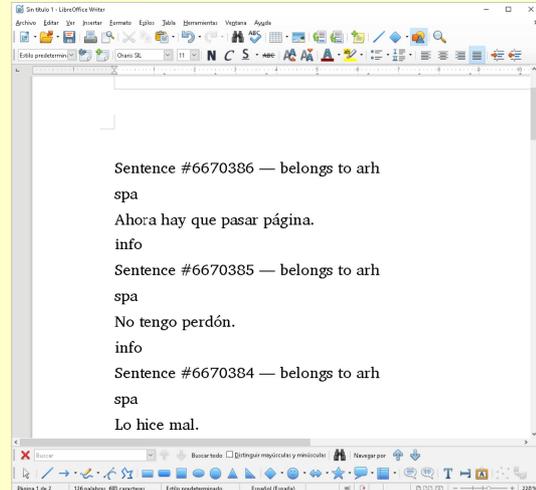
¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

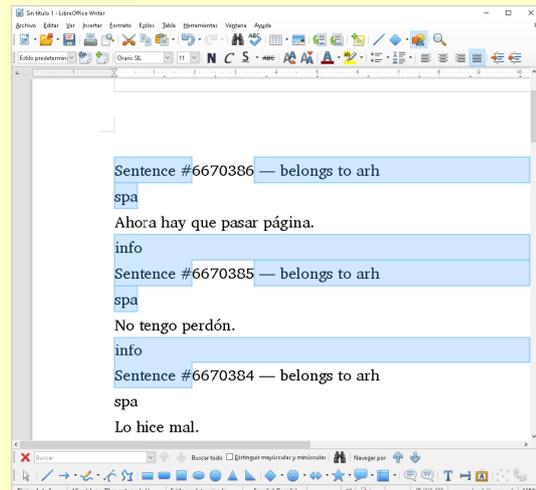
Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

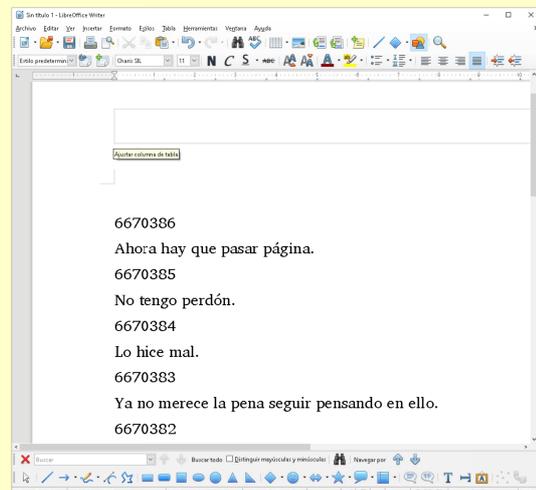
*Pegamos el texto que copiamos en el procesador de textos que estemos utilizando (mejor pegar sin formato).
Vemos, como habíamos previsto, que nos sobra información, por lo que eliminaremos lo que no nos interesa.*



Marcamos toda la información sobrante y la borramos directamente.



*Así nos queda el listado al borrar lo que no necesitamos.
Sólo tenemos ahora que alinear cada código con su frase e intercalar el guión con los espacios necesarios.
Recuerda que el programa para enlazar automáticamente cada frase con su archivo de audio respectivo utiliza la secuencia:
[número de referencia] [espacio] [guión] [espacio] [título].*



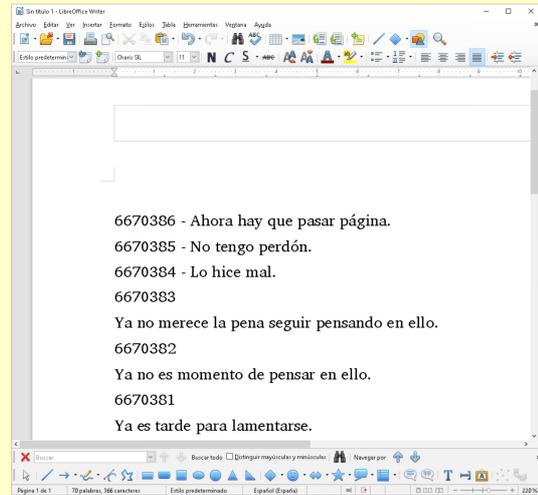
¿Cómo poner voz a las frases en Tatoeba?

-Guía preparada por arh (febrero 2018)

Comprueba la última versión aquí: https://mega.nz/#F!9cEixIAL!A5Al-5OWAzB_VTh0fBI0Ig

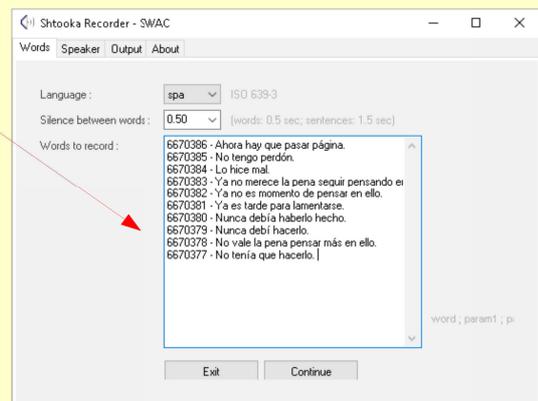
Consulta también esta otra explicación de CK (en inglés): <http://study.aitech.ac.jp/tatoeba/shtooka/>

Así que lo hacemos para cada frase y ...



6670386 - Ahora hay que pasar página.
6670385 - No tengo perdón.
6670384 - Lo hice mal.
6670383 - Ya no merece la pena seguir pensando en ello.
6670382 - Ya no es momento de pensar en ello.
6670381 - Ya es tarde para lamentarse.
6670380 - Nunca debía haberlo hecho.
6670379 - Nunca debí hacerlo.
6670378 - No vale la pena pensar más en ello.
6670377 - No tenía que hacerlo.

... ya tenemos nuestro listado listo, que podemos copiar y pegar directamente en la caja 'Words to record' para iniciar la grabación.



Si te animas a traducir estas instrucciones a tu lengua nativa, encontrarás los archivos originales (textos e imágenes) en el mismo enlace que este documento.