

Okamy Synth W

(Devkit documentación W1 ver.)

[Herramientas](#)

1.- Generacion del Diccionario:

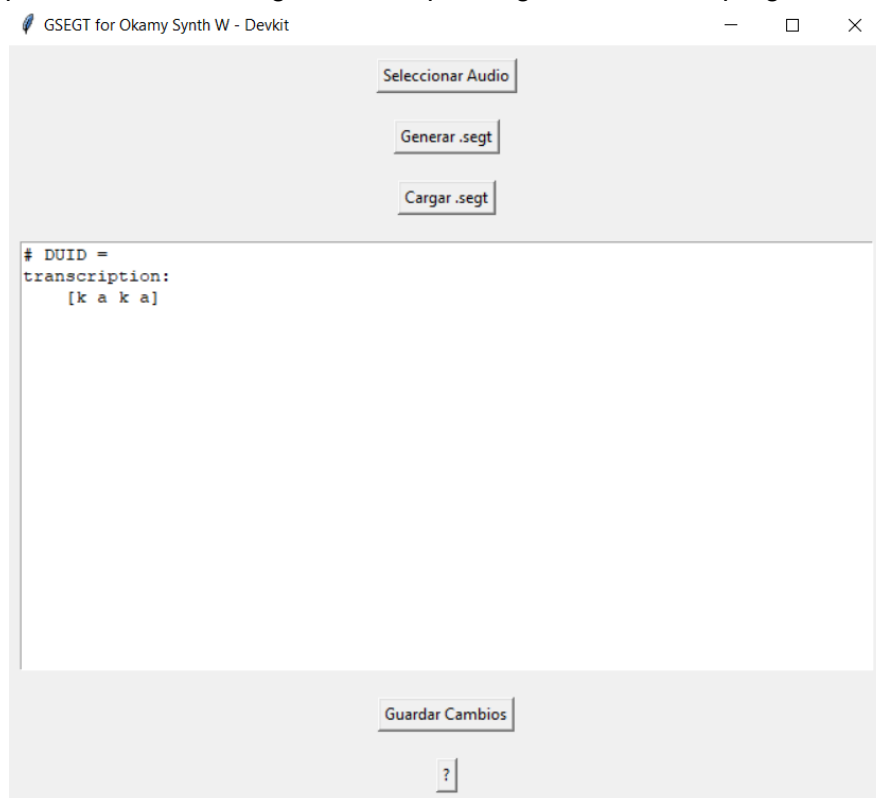
Al tener ya las herramientas, tendrás que crear una carpeta llamada “voicebank” y otra “Devkit”. Dentro de “Devkit” pondrás las 5 herramientas proporcionadas en el .rar. Ejecutarás el “Dictionary_Generator.exe” y seleccionamos las opciones a tu gusto, se puede editar el diccionario pero el DUID no se debe cambiar.

2.- Generación del modelo de banco de voz:

Cuando tengas el diccionario listo, ejecutaras el “SingerMaker.exe” y rellenaras todas las opciones, primero de todo generarán el “PUID”, la carpeta de destino tiene que ser dentro de la carpeta “voicebank” y generarán el banco de voz. Tiene un singer.inf que puedes editar manualmente.

3.- Generar Transcripciones:

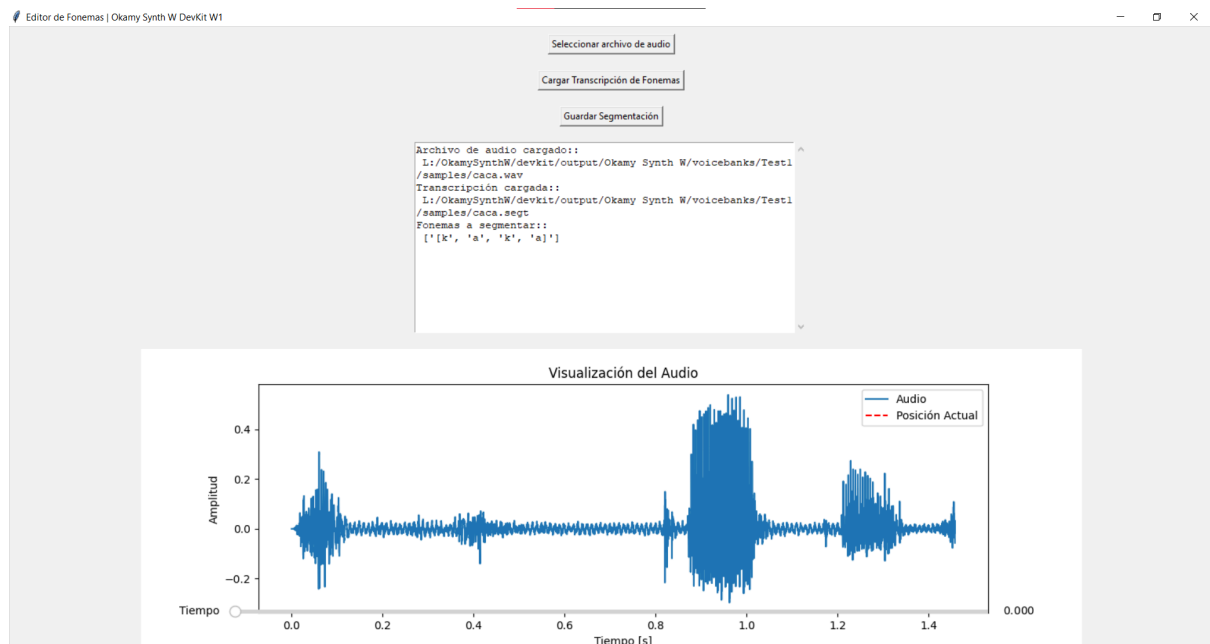
Dentro de la carpeta del banco de voz, abra una carpeta llamada “samples” dentro de ahi pondras todos los fonemas(aún en desarrollo el multipitch, perdon chicos soy novato) y con la herramienta “GSGT_Generator.exe” generarán los .segt(los archivos de transcripción) para cada audio, luego tendrás que cargarlos desde el programa.



Se verá así, el “# DUID =” tienes que poner el DUID del diccionario y en los [] entrará la transcripción, los fonemas tienen que estar en el diccionario.

4.- Segmentar fonemas:

Después de la transcripción, ejecutaremos “DevKit_UI.py” y cargaremos el audio junto a su .seg:



De momento no he llegado a poner las etiquetas pero tienes que clicar donde empieza el fonema y donde acaba, se puede ampliar pero aún se está improvisando. Se segmenta por orden, si tienes por ejemplo [k a k a] primero se segmenta la “k” y luego la “a” y así progresivamente.

Después de segmentar tendrás que darle a “Guardar segmentación” y así se guardaran los datos en el .seg.