

Introducción

En nuestros días, la mayor parte de la sociedad necesita de la tecnología para el desarrollo de la vida cotidiana. La comunicación, educación, noticieros, bancos, compras y ventas electrónicas se han vuelto aspectos útiles y en algunos casos necesarios. Para interactuar con esto se necesitan habilidades como la lectura, escritura, visualización y conocimiento de lo que se puede y quiere hacer, ¿Pero qué pasa con las personas que padecen de alguna incapacidad que no les permita interactuar con los servicios electrónicos? Para esto sería necesario algún tipo de herramienta que facilitara su uso, ya que no cuentan con las mismas capacidades que la mayor parte de las personas. Es por esto que se busca el desarrollo de la voz artificial, ya que de esta forma tendríamos el beneficio de acceder e interactuar con los sistemas electrónicos que hoy en día utilizamos.

En el caso de la educación, los niños comienzan a desarrollar sus habilidades mediante los sentidos (tacto, gusto, vista, olfato y oído). Uno de los principales sentidos utilizados es el oído, por medio de éste, los bebés escuchan a sus padres y los lleva a la imitación de sonidos para generar palabras, hasta llegar a frases y posteriormente ideas completamente bien estructuradas. Por lo tanto, el hablar y escuchar son factores importantes para el desarrollo del lenguaje. A nivel primaria, los niños aprenden el abecedario, sílabas, palabras, frases y reglas gramaticales para lograr una lectura clara y uniforme. Tomando en cuenta la naturaleza misma que el niño tiene para aprender escuchando, podemos decir que podría ser un buen medio que el niño aprendiera a leer también a través del oído, así mismo se podría tener una mejor calidad de enseñanza y se fomentaría el hábito de la lectura.

En el caso del estudio de la voz es necesario adaptar los sonidos al lenguaje español mexicano, y se busca realizar un sintetizador de voz en el que se podrán expresar las palabras que utilizamos en la enseñanza de niños de nivel primaria y en el lenguaje cotidiano. En México, esto se requiere para las aplicaciones que actualmente se están desarrollando en TLATOA, para la enseñanza de la lectura y escritura a los niños mexicanos.

La síntesis de voz es el proceso de transformar el texto a sonido, esto nos sirve para la creación de voz artificial. Dadas las palabras, frases o números escritos, el sintetizador (TTS) se encarga de pronunciar adecuadamente lo que el usuario desea escuchar. Un TTS no solo produce palabras aisladas, sino también forma palabras con fonemas para su reproducción. Actualmente, existen sintetizadores en varios idiomas por ejemplo: inglés, español, francés, alemán, entre otros. La mayoría de los sintetizadores en español están diseñados para palabras más complejas de las que un niño puede entender. Por ejemplo, en México los libros de texto a nivel primaria utilizan palabras de los dialectos náhuatl, maya, zapoteca e incluso palabras en diminutivo que los sintetizadores no pueden expresar con claridad, por lo que nos vemos en la necesidad de desarrollar un sintetizador de voz tomando en cuenta el caso de sistemas para niños, en español mexicano.

Se pretende desarrollar un sintetizador de voz para el idioma español en México, basado en sintetizadores existentes que comprenden sonidos específicos para una pronunciación ideal en el contexto de la enseñanza para niños mexicanos.

Los objetivos específicos de este trabajo son:

1. Evaluar los sintetizadores existentes.
2. La búsqueda de sonidos característicos del español en México.
3. Diseñar el corpus de grabación que represente de manera completa todos los sonidos y las combinaciones posibles de éstos.
4. El desarrollo del protocolo de grabación.
5. La grabación del corpus para el sintetizador de voz con una voz clara y uniforme.
6. Etiquetación manual de los archivos de sonido.
7. La adaptación y evaluación del sintetizador con herramientas desarrolladas en Java [Flores, 2001].

El proyecto constará de cinco capítulos en los que se describirá paso a paso el proceso y el marco teórico sobre la síntesis de voz. En el capítulo uno se habla sobre la comunicación y lenguaje, el cual es importante mencionar para tener las bases sobre este tema. En el siguiente capítulo se explica lo que son los sintetizadores de voz, donde se mencionan los tipos de sintetizadores, sus características y algunos de los que se utilizan en la actualidad, también se da una pequeña introducción de lo que es un corpus de voz, tipos y características. En el capítulo tres se describe el desarrollo del corpus de voz llamado *Gama*, desde la simbología del lenguaje hasta tener el documento completo, tratando de abarcar todos los posibles sonidos del Lenguaje Español hablado en México. En el capítulo cuatro se habla del protocolo de grabación y asimismo de la etiquetación de voz. Por último, en el capítulo cinco retomamos la tesis de licenciatura de Leonardo Flores realizada en el año 2001 para las pruebas con el nuevo corpus, utilizando el programa desarrollado para síntesis de voz con el método de Unit Selection [Flores, 2001].

