

Conclusiones

El desarrollo de una nueva voz en español mexicano para el sistema de texto a voz Festival implicó una investigación muy completa pues involucró la inter-relación de aspectos muy diferentes. Entre ellos se encuentra la lingüística y gramática del español mexicano, acústica, comprensión de procesamiento digital de señales dentro de la rama de la ingeniería electrónica, programación orientada objetos en C/C++, programación lógica y funcional en Lisp y Scheme, así como un alto nivel de conocimiento en la configuración de sistemas operativos Unix (en este caso Linux).

Además de la investigación necesaria para realizar esta tesis, se requirió una gran paciencia y pericia técnica para materializar la investigación, principalmente resolviendo diversos problemas de compilado y configuración. La cooperación entre investigadores y desarrolladores a través de foros fue vital para poder realizar este trabajo. La participación activa del Dr. Alan Black en la corrección de un error dentro del código de Festival identificado en el desarrollo de este trabajo, lo cual fue necesario para poder crear la nueva voz en español, me resultó muy gratificante. Mostró que mi investigación pudo tener una importancia global dentro del desarrollo de voz, pues se corrigió el código de Festival desde su punto de distribución en el servidor de Carnegie Mellon, además me enseñó que pude comunicarme con otros investigadores vía remota de una manera clara y eficiente.

A lo largo de este documento se trató de mostrar varios puntos de vista que engloban la producción de una voz en un sistema texto a voz. En el primer capítulo se dió una visión general a lo que es un sistema TTS, sus diferentes tipos y el estado actual de la materia, lo cual nos

permitió entender las generalidades de estos sistemas y resultó muy útil para analizar el caso de esta tesis en particular. En el segundo capítulo se describió de manera técnica cómo poder instalar toda la infraestructura necesaria para desarrollar una nueva voz en Festival; de esta manera se expuso una parte muy importante que ayudará a investigaciones futuras. En el tercer capítulo se explicó la arquitectura de Festival y una descripción detallada de su estructura, para poder mostrar cómo programar cada uno de los módulos que definen las características de un idioma, poniendo énfasis en aquellas partes susceptibles de modificarse para adecuar al español como lenguaje dentro del sistema. En el capítulo siguiente se describió el proceso requerido para la síntesis de la voz, exponiendo el proceso de creación de una base de datos de difonemas como fuente de la voz en español. Por último, en el quinto capítulo se mostraron los resultados de las pruebas realizadas a esta nueva voz para efecto de evaluación. Estas pruebas fueron hechas sobre aspectos concretos de la nueva voz de manera que pueda ser evaluada objetivamente. Finalmente, se proponen posibles trabajos futuro dentro de esta línea de investigación del procesamiento automático de voz, donde se profundice la adecuación de algunos aspectos del sistema TTS para el español.

Esta tesis fue realizada con el objetivo de retomar una investigación muy importante y adaptarla al presente. Existe un inmenso trabajo a futuro dentro del desarrollo de una voz artificial para el español mexicano. El trabajo para la afinación (*tuning*) es infinito, pues siempre existirán mejoras a los módulos con objetivo de emular cada vez mejor la voz humana. Esta tesis vuelve a poner sobre la mesa las herramientas necesarias para poder realizar una voz artificial para nuestro idioma. Espero que estas herramientas sean útiles y futuras investigaciones las utilicen como base para poder acercarse cada vez más a la naturalidad e inteligibilidad de la voz humana en nuestra lengua.