

Capítulo 1

1.1 Introducción

Existen varias investigaciones acerca de los reconocedores de voz y cada vez se pretende que este sea más eficiente, pues su principal objetivo es reconocer la voz con las menores restricciones posibles, para poder lograr esto el corpus utilizado al momento de crear el reconocedor debe contener todas las posibles conjugaciones fonéticas.

Actualmente estamos lejos de crear un sistema completo que pueda comprender cualquier mensaje oral tal como lo podría hacer cualquier humano. Esto por que el vocabulario que se tiene es muy amplio con varios modismos y diferencias en significado por la zona geográfica, cultura, etc.

En México no existe mucha investigación y no hay mucho desarrollo sobre reconocimiento de voz esto debido a la carencia de un corpus lingüístico del español hablado en México.

Es el objetivo de este trabajo, el generar un recurso que apoye las investigaciones en tecnologías del habla para el español hablado en México. Para ello se creará un corpus de voz de propósito general de alta calidad, completamente etiquetado a nivel texto, palabra y fonema.

El dominio de este corpus debe ser rico fonéticamente, es decir, debe contener todas las posibles combinaciones del idioma, para que este pueda representar lo mejor posible al español hablado en México.

1.2 Objetivos

1.2.1 Objetivo General

El propósito general de esta tesis es crear un corpus balanceado fonéticamente en Español de México hablado por adultos y grabado con calidad micrófono (16 KHz.).

1.2.2 Objetivos Específicos

Diseño del protocolo de grabación para maximizar la cobertura de las combinaciones fonéticas del lenguaje. Y también verificar que este bien balanceado fonéticamente para es español mexicano.

Grabación de por lo menos 50 locutores por medio de una aplicación elaborada en CSLU Toolkit y cada una de las grabaciones será validada para que este en el rango de 16 KHz.

Limpieza, transcripción y etiquetado de cada unas de las grabaciones que se hicieron y después de que el corpus este depurado se tratara entrenar una red neuronal con CSLU Toolkit.

1.3 Planteamiento del Problema

El área en aplicaciones del habla ha avanzado demasiado rápido así como los progresos en el lenguaje y computadoras. A medida que pasa el tiempo y la tecnología avanza, se acrecentando las necesidades de la gente, lo que conlleva a crear servicios más eficientes y cada vez más sencillos. Creando así un sistema que funciona reconociendo las palabras que pronuncia el usuario, interpretando la secuencia de las mismas para obtener un significado en términos de la aplicación del sistema y proporcionando una respuesta apropiada para el mismo. El objetivo es hacerlos sencillos pues debe ser capaz

de reconocer voz con las menores restricciones. El designio del reconocimiento de voz es la conversión de los sonidos del habla a texto. [Rodríguez, 2002]

Para tal investigación es necesario utilizar los recursos adecuados para los usuarios. Pues en la actualidad no existe un reconocedor de voz para el vocabulario del Español de México. Este utiliza un sistema fonético diferente a cualquier otro idioma incluso con el mismo español de España. Existen muchos reconocedores que reconocen Español de España por lo que surge la necesidad de crear un corpus que apoye al crecimiento de un reconocedor de voz para el Español de México.

El problema que se aborda en esta tesis es la creación de un Corpus de voz con habla espontánea y la comparación con el mismo corpus pero eliminando el habla espontánea, ya que en el habla espontánea entran sonidos o palabras incompletas que al momento de hacer la prueba del reconocedor este tiene emisiones de algunas palabras que no logra encontrar.

Debemos tomar en cuenta detalles como son: Tener que calibrar el micrófono antes de comenzar a hacer las grabaciones, ya que la calidad de estas va a depender del timbre de cada locutor para que el corpus que necesitamos cuente con una mejor grabación y calidad.

Resumen

Este capítulo nos habla sobre el problema al cual está enfocada esta tesis que consiste en la creación de un corpus balanceado fonéticamente de propósito general calidad micrófono y la importancia que este tiene para el entrenamiento de los reconocedores de voz para el español mexicano. Nos explica los objetivos tanto generales como específicos, y el planteamiento del problema.