

Capítulo 1: Introducción

1.1 Discapitados en México

Según el XII Censo General de Población y Vivienda 2000, una persona con discapacidad "Es aquella que presenta una limitación física o mental de manera permanente o por más de seis meses que le impide desarrollar sus actividades en forma que se considera normal para un ser humano". Las estadísticas del INEGI (Instituto Nacional de Estadística Geográfica e Informática) en el 2000 se identificaron 1 millón 795 mil personas con discapacidad, las cuales representaron 1.8% de la población total. La siguiente grafica nos muestra una distribución porcentual de la población según el tipo de discapacidad (Figura 1) [INEGI, 2004].

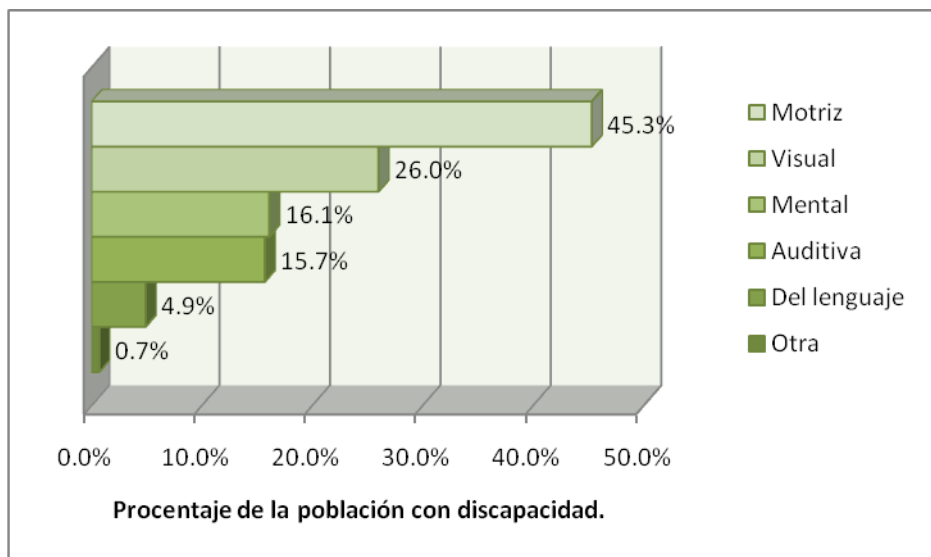


Figura 1. Distribución porcentual de la población según el tipo de discapacidad. (La suma de los distintos tipos de discapacidad puede ser mayor a cien por la población que presenta más de una discapacidad.) Datos tomados de [INEGI, 2004].

Es importante destacar a las personas con algún tipo de discapacidad que se encuentran en nuestro país. Porque estas personas son las que más necesitan ayuda en todos los sentidos, pues debido a su discapacidad tiene una vida más complicada.

Aunque en nuestro país poco a poco se está creando una cultura de ayuda a las personas discapacitadas, esta no es del todo completa. Una persona discapacitada, tiene los mismos derechos que los demás a acceder y conocer la información que en la Web se ofrece. En la actualidad las personas con discapacidades no cuentan con aplicaciones suficientes que les permitan acceder a la información de la Web cumpliendo con todas sus necesidades. Es por esto que esta tesis se enfoca en esa parte de la población mexicana. Especialmente personas con discapacidades visuales que conforman el 26% de la población con discapacidad.

1.2 Importancia de Internet

Internet ha revolucionado el mundo de las computadoras y las comunicaciones. Es un canal de transmisión en todo el mundo, un mecanismo de difusión de la información y un medio de colaboración e interacción entre individuos y sus computadoras sin importar su localización geográfica [Leiner et al, 2003]. Una de las características importantes de Internet, es la de facilitar el acceso a la información a las personas.

En la actualidad internet cuenta con millones de usuarios alrededor de todo el mundo. Basados en los reportes arrojados por la página electrónica “Internet World Stats” la cantidad de usuarios es de 1’173’109’925 hasta finales del mes de junio de 2007. Los cuales se dividen en regiones (Figura 2 y 3) [Argaez, 2002].

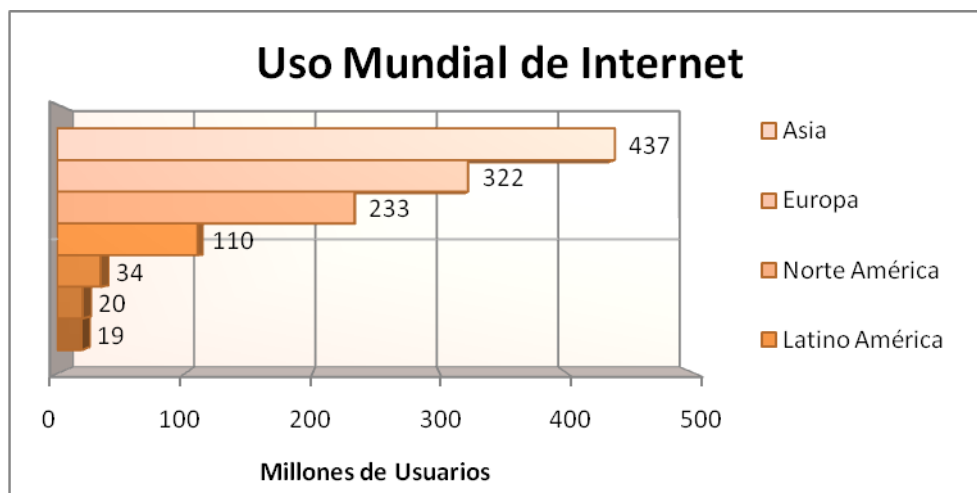


Figura 2. Uso mundial de internet dividido por regiones. Datos tomados de [Argaez, 2002].

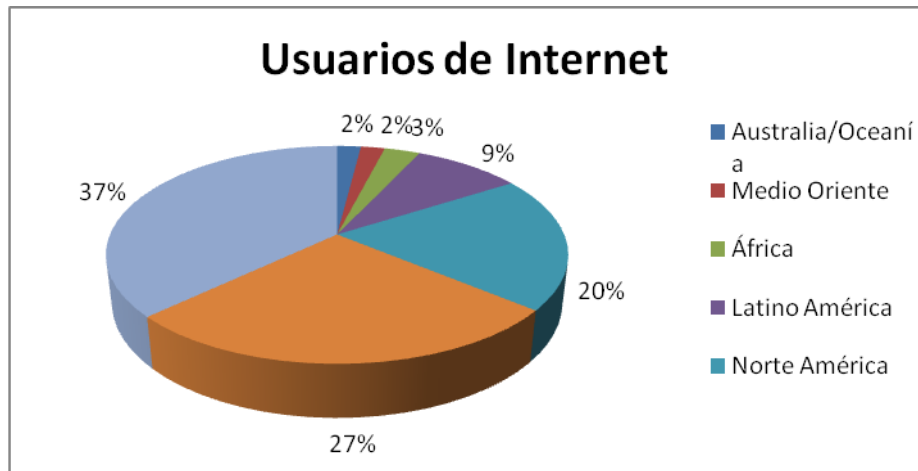


Figura 3. Porcentaje mundial de usuarios en Internet. Datos tomados de [Argaez, 2002].

Las gráficas muestran que en la región de Asia, es donde se concentran la mayor cantidad de internautas, enseguida Europa y Norte América. Latino América ocupa el 9% de la población mundial de usuarios. Las graficas siguientes, se dividen en base al idioma de los usuarios (Figura 4) [Argaez, 2002].

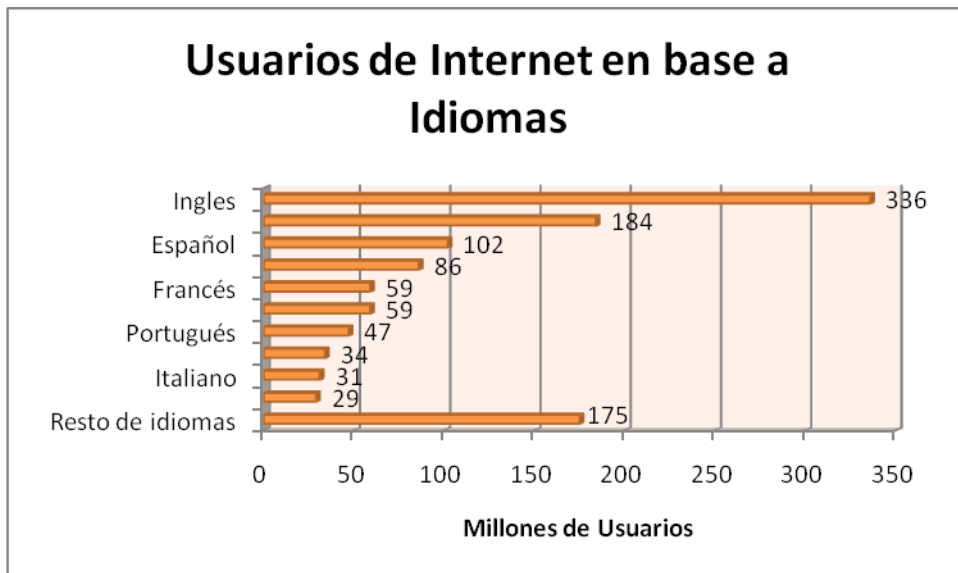


Figura 4. Uso mundial de Internet en base al idioma de los usuarios. Datos tomados de [Argaez, 2002].

La figura 4 muestra el número de usuarios en internet que hablan español en todo el mundo. Es evidente la importancia de estos, pues solo se encuentran debajo del idioma inglés y chino. Comparados con el resto del mundo los ingleses integran el 31.2%, los chinos el 15.7% y los españoles el 8.7% [Argaez, 2002].

Apoyados en los estudios realizados por la Asociación Mexicana de Internet (AMIPCI). En el 2007 existen 14.8 millones de computadoras personales (PC's) en México de las cuales 8.7 cuentan con acceso a Internet. Como resultado el 59% de las computadoras están conectadas a Internet. De estas, el 55% están instaladas en hogares, el 45% restante se encuentran en las empresas. México cuenta con 22.7 millones de internautas mayores de 6 años, más del 20% de los usuarios de todo latino América. Google es el buscador más usado por los internautas mexicanos, pues es utilizado por el 85% de los usuarios.

1.3 Definición del problema

El problema que se busca resolver es facilitar a los usuarios invidentes y con visión reducida que hablan español, específicamente español mexicano, la búsqueda de información y recursos en la Web, de manera gratuita, sencilla y natural. El desarrollo de esta aplicación estará basado en el sistema Linter-Vox [Elizalde, 2006]. Linter-Vox ofrece la interfaz que se conecta al motor de búsqueda Google y lee los resultados pero carece de funciones de reconocimiento del habla que le permitan hacer búsquedas con la voz como entrada, mismas que serán desarrolladas dentro de esta tesis, para facilitar su uso a personas invidentes.

El problema técnico al que se enfrenta este proyecto es la integración de un reconocedor y un sintetizador de voz a una aplicación desarrollada en Java. Pues tanto los reconocedores como los sintetizadores de voz que se encuentran disponibles están programados en otros lenguajes distintos a Java. En cuanto al reconocedor de voz se debe buscar una herramienta que sea gratuita, que ofrezca una buena precisión en sus resultados y que su funcionamiento este enfocado al idioma español. Además se debe definir una estrategia que permita que al vocabulario del reconocedor se le agreguen

nuevos términos de manera dinámica, definir un contexto para el vocabulario y no permitir que este incremente demasiado.

De igual manera esta aplicación se debe integrar a la ya existente que lleva por nombre Linter-Vox para lograr que ambas trabajen de forma coordinada y transparente. Lo cual es una tarea sumamente difícil pues implica modificar la aplicación Linter-Vox para que esta reciba los resultados obtenidos por el reconocedor.

1.4 Propuesta de solución

En respuesta al problema planteado se propone la creación de Conquiro-Vox¹, una aplicación prototipo que permita a los usuarios realizar búsquedas en la Web, comunicándose con la computadora por medio de voz, para que esta, en base a la información dictada realice la búsqueda, lea los resultados y permita navegar a través de los mismos. Para realizar estas funciones la aplicación integrará un reconocedor y un sintetizador de voz los cuales funcionaran especialmente enfocados al español, con el fin de hacer más fácil y cómodo el acceso a la información en la Web a personas invidentes. La aplicación prototipo permitirá a los usuarios utilizar el motor de búsqueda Google de una manera fácil, sencilla y natural. Pues la consulta la hará utilizando su voz, para después escuchar los resultados que serán leídos por la máquina. Esta aplicación será gratuita y basada en español mexicano.

1.5 Objetivos

1.5.1 Objetivo general

Crear una herramienta que permita hacer búsquedas en la Web utilizando la voz como entrada.

¹ El nombre Conquiro-Vox proviene del latín. Conquiro que significa buscar y Vox que es equivalente a voz.

1.5.2 Objetivos específicos

- Determinar la arquitectura adecuada del sistema.
- Determinar el tipo de reconocedor apropiado.
- Definir el contexto y vocabulario inicial.
- Diseñar una estrategia para poder adaptar el vocabulario dinámicamente.
- Realizar pruebas de funcionalidad y usabilidad.

1.6 Alcances

- Utilizar herramientas de reconocimiento de voz y salida de voz, decidiendo cuales basado en un análisis previo.
- Programar y evaluar el sistema sobre un contexto predefinido.
- Buscar diseñar una estrategia que permita ampliar el vocabulario del reconocedor.

1.7 Limitaciones

- Se podrán reconocer palabras aisladas, no habla continua.
- Se utiliza un reconocedor existente de propósito general sin tratar de mejorarlo o re-entrenarlo.
- Se determinó un contexto de aplicación (no se pretende que sea independiente del contexto).
- No se pretendió remplazar por completo los medios de captura tradicionales (teclado y ratón), si no complementarlos.

1.8 Herramientas

1.8.1 Hardware

Para correr esta aplicación es necesario contar con una PC, la cual debe de contar como mínimo con un procesador Intel o compatible de 100 MHz, 16 MB de memoria RAM, 50 MB de espacio libre en el disco duro para la instalación de componentes, tarjeta de sonido compatible con Sound Blaster, micrófono, bocinas y conexión a Internet. Como se mencionó al principio estos son los requerimientos mínimos en cuanto al Hardware, los cuales son fácilmente superados por cualquier computadora desarrollada en nuestros tiempos.

1.8.2 Software

En cuanto al software necesario este quedará dividido en dos partes: software necesario para el usuario y software necesario para el desarrollo y construcción de esta aplicación.

Dentro del software para el usuario se necesita tener un sistema operativo Windows 95, 98, NT, ME, 2000, XP Home and XP Professional, tener instalado el java JDK 1.5 o inferior, al igual que el CSLU Toolkit 2.0, instalar los agentes de Microsoft, sustituir el script `genrecog_3.1.tcl` por el ofrecido en el directorio de esta aplicación y por ultimo agregar a la librería de java los siguientes archivos: `googleapi.jar`, `itclsrc.jar`, `jaclsrc.jar`, `janino.jar`, `janinosrc.jar`, `tcljavasrc.jar`, `tjc.jar`, `tjcsr.jar` y `tcljava.jar`.

El software para el desarrollo necesita además de contar con el software del usuario, el API de Google para conectarse al motor de búsquedas, instalar el JCreator Pro 4 en el cual se desarrollará el código de la aplicación, las librerías Jacl 1.4.0, instalar el Altova UModel para realizar los diagramas de clases, al igual que el SmartDraw 7.

1.9 Descripción del documento

Este documento se encuentra dividido en cinco capítulos los cuales recopilan todo el proceso de estudio, análisis, diseño, construcción, pruebas, así como la metodología que fue implementada para cumplir todos los objetivos previamente establecidos.

En el capítulo 1 se define el problema que se ataca, al igual que los objetivos que se deben de cumplir y las herramientas utilizadas para desarrollar de forma exitosa este proyecto.

Dentro del capítulo 2 se presenta una introducción al reconocimiento de voz y se describe el proceso de análisis de la aplicación, se analizan todos los elementos que conforman a Conquiro-Vox, se estudiaron todas las herramientas existentes y se eligieron las más adecuadas para el desarrollo de esta aplicación. Se diseña la arquitectura de la aplicación y como se espera que funcione basándose en las tendencias actuales de los motores de búsqueda.

El capítulo 3 se refiere al diseño y construcción de la aplicación, aquí se explica de manera detallada cómo funciona el sistema y la forma en que el usuario puede interactuar con ella esto con la ayuda de diagramas de secuencia y de clases. Además se explica sobre los errores y excepciones que pueden ocurrir durante la ejecución del sistema y como resolverlos.

La construcción y aplicación de las pruebas de uso y funcionalidad se encuentran en el capítulo 4. Este capítulo explica los métodos que se siguieron para la elaboración y aplicación de las pruebas, así como el análisis de los resultados y los errores que fueron detectados gracias a estas.

Por último el capítulo 5 presenta las conclusiones que se generaron al término de esta tesis además de detallar las actividades que serán realizadas a futuro con el fin de mejorar y complementar la aplicación.

Dentro de este capítulo se define el problema que ataca esta tesis, se plantea una solución y se describen los objetivos que se deberán de cumplir para lograrlo. De igual manera se muestran los alcances y limitaciones, se justifica el uso de herramientas tanto de hardware como de software. En el siguiente capítulo se muestra la teoría referente al reconocimiento de voz, al igual que el análisis de la aplicación.