

## Capítulo 2. Marco Teórico

---

<b>2.1 Tecnología computacional y personas con capacidades diferenciadas</b>	<b>19</b>
<b>2.2 Estadísticas en México sobre discapacidad</b>	<b>21</b>
<b>2.3 Delphi: el lenguaje de programación elegido</b>	<b>25</b>
<b>2.4 SAPI. Speech Application Programming Interface</b>	<b>26</b>
<b>2.4.1 API para sistemas TTS (Text to speech)</b>	<b>29</b>
<b>2.4.2 API para reconocimiento de voz</b>	<b>30</b>

No hay fracaso que exista en mi mente  
Sólo quiero luchar con valor  
Aunque se que he caído mil veces  
Levantarme ha sido lo mejor  
[Ramírez, 2003]

## Capítulo 2. Marco Teórico

---

El proyecto MexVox es un sistema que requiere del apoyo de diferentes herramientas computacionales, principalmente de software. En este capítulo se presenta una breve descripción de cada una de estas herramientas de software así como del trabajo realizado y la necesidad de colaboración de todos y cada uno de nosotros para las personas con capacidades diferenciadas.

### **2.1 Tecnología computacional y personas con capacidades diferenciadas**

El uso de la computadora ha revolucionado y facilitado la vida de las personas en gran parte de las actividades cotidianas: trabajo, entretenimiento, investigación, hogar, estudio etc. Sin embargo, para las personas con capacidades diferenciadas es más difícil poder acceder a la tecnología y a los beneficios que esta ofrece. Teclear un mensaje, navegar por Internet, jugar o aprender con programas multimedia, leer textos en pantalla o simplemente desplazar el ratón son actividades tan sencillas y hasta rutinarias para la mayoría de los videntes, que por lo general, desconocemos la realidad de los discapacitados en nuestro país, quienes necesitan costosas herramientas de hardware y software para poder desarrollar la misma clase de actividades.

#### **Adaptadores.**

Para que una persona con capacidades diferenciadas pueda hacer uso de la tecnología requiere de adaptadores que compensen sus carencias, ya sea motriz, cognitiva, visual o

auditiva. Claudia Miramón, educadora diferencial de la fundación Alter Ego en Chile [TID, 2002], desempeña su trabajo en niños con parálisis cerebral y conoce de los avances que ellos han logrado gracias al uso de la computación. “Tenemos pacientes que sufren daños motores pero poseen buenos niveles intelectuales, por lo tanto, el computador les ha permitido desarrollar habilidades de escritura, lenguaje, matemáticas, juegos, búsqueda y manejo de información” [TID, 2002].

Los adaptadores para capacidades diferenciadas pueden ser de tres tipos: hardware, software y adaptadores para las personas.

1. *Adaptadores de Hardware*: son dispositivos propios del hardware de computadora que han sido modificados de alguna manera con el objetivo de cubrir las necesidades que el usuario requiere. Algunos de estos son:
  - *Intellikeys* Es un teclado alternativo con láminas intercambiables para distintos usos. Se pueden hacer plantillas en lenguaje Braille para adecuarla a personas ciegas.
  - Pantalla táctil o *Touch Window* es un monitor que permite a la persona interactuar con el computador a través de un toque manual de la pantalla.
2. *Adaptadores de Software*: Software que ha sido desarrollado con el fin de que una persona con capacidades diferenciadas pueda realizar cualquier tipo de tarea por medio de la computadora. Este tipo de software se desarrolla en base a las limitaciones de una capacidad diferente. En algunas ocasiones este tipo de software tiene un papel mediador entre el usuario y el software comercial.

- **Globus:** Se trata de un visualizador fonético. A través de la utilización del micrófono de la computadora, es posible visualizar el sonido de distintas maneras y hacer pequeñas actividades de imitación de los sonidos emitidos por el profesor, o de competir con el ordenador, simulando en un sencillo juego de carreras.
  - **Speech Viewer III - IBM:** se trata de una herramienta que transforma palabras o sonidos hablados en atractivos gráficos. Asimismo, incrementa la efectividad de la terapia de lenguaje y habla en las personas que presentan problemas de lenguaje, habla y audición.
3. *Adaptadores para persona:* Son herramientas que permiten a la persona estar en una posición cómoda y a su vez interactuar con el hardware.
- **Bipedestador:** Adaptador que se utiliza para que las personas puedan estar de pie sin la necesidad de una segunda persona.

## **2.2 Estadísticas en México sobre discapacidad**

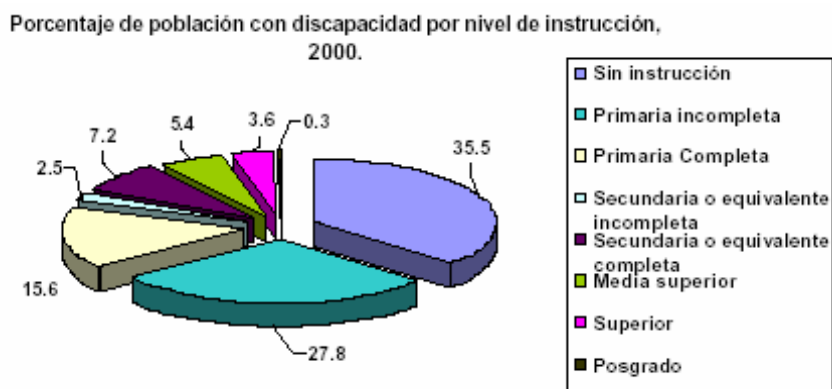
Según información de la Organización Mundial de la Salud por cada persona que presenta una discapacidad dos personas más se encuentran completamente involucradas en la discapacidad.

**Discapacidad:** de acuerdo al censo efectuado por el INEGI en el año 2000, la población total en México era de 97 483 412 habitantes [Discapacidad, 2003] de los cuales 1.8% tienen alguna discapacidad.

Las discapacidades más frecuentes en el país son:

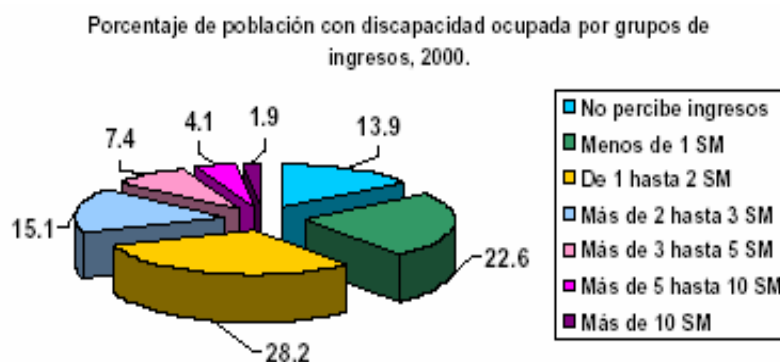
Motriz	45.3 %
Visual	26.0%
Auditiva	15.7%
Mental	16.1%
Lenguaje	4.9%

**Educación y discapacidad:** El 56.5% de la población con discapacidad de 8 a 14 años sabe leer y escribir. En México de cada 100 personas con discapacidad mayores de 15 años, 33 no saben leer y escribir. El 63.8% de la población con discapacidad de 6 a 29 años no asiste a la escuela y 34.5% aún lo hace; el análisis entre hombres y mujeres muestra que para ambas situaciones los varones tienen mayores concentraciones; es decir, van más hombres a la escuela (56%) [DIPD, 2003].



**Figura 2.1** Porcentaje de población con discapacidad por nivel de instrucción [DIPD, 2003]

**Trabajo y discapacidad:** La tasa de participación económica nacional para la población con discapacidad es de 25%; resalta que la participación de los hombres es tres veces mayor (36.4%) que la de mujeres (12.6%). La población no económicamente activa se concentra más en las mujeres (55.7%) que en hombres (44.3%). Estos datos muestran que la población con discapacidad en edad de trabajar no ha logrado integrarse al mercado laboral, situación que acentúa su vulnerabilidad. La población que no está empleada formalmente se dedica a distintas actividades no económicas, 24.6% únicamente hacen las tareas de su hogar, 11.4% están incapacitados permanentemente para trabajar, 10.7% son jubilados o pensionados, 4.8% estudian y 48.5% realizan otro tipo de actividad no económica [DIPD, 2003].



**Figura 2.2** Porcentaje de población con discapacidad ocupada en grupos de ingresos, 2002 [DIPD, 2003]

**Asociaciones en México para personas con discapacidad:** en el país existen 988 asociaciones de y para personas con discapacidad, el mayor número de ellas se concentra en el Distrito Federal, Coahuila, Jalisco, Guanajuato y Nuevo León. Del total de asociaciones, 57% atiende a población con discapacidad neuromotriz, 56% mental, 41%

problemas de lenguaje, 37% afecciones visuales y 36% auditivas. Cabe señalar que algunas asociaciones atienden más de un tipo de discapacidad [DIPD, 2003].

**ACRIP (Asociación Cultural y Recreativa de Invidentes de Puebla):** Fundada el 16 de marzo de 2001 por el profesor Alejandro Ramírez Campos con la finalidad de dar un trato digno al invidente. Actualmente cuenta con 37 miembros activos que trabajan y participan de actividades diversas organizadas por la asociación.

Dentro de las actividades que la Asociación ofrece a sus integrantes se encuentra:

- Cursos de sistema braille
- Clases de música
- Manualidades (macramé, tejido)
- Movilidad y manejo de bastón para invidentes
- Asesorías para obtener perros guías
- Deportes
- Capacitación, rehabilitación y adaptación de ciegos recientes
- Cursos de computación por medio de sistemas parlantes: MexVox

Se pretende lograr que los medios de comunicación se interesen en difundir el apoyo a invidentes para ellos puedan demostrar con educación, respeto y seguridad las capacidades que se tienen y que por medio de ellas puedan mejorar su calidad de vida y un digno trato social [Ramírez, 2004].

**Otras asociaciones:** ONCE (Organización Nacional de Ciegos Españoles). Fundada el 13 de diciembre de 1938 por un Decreto del Jefe del Estado [ONCE, 2003]. El objetivo fundamenta fue integrar en una sola organización a los ciegos españoles que se encontraban dispersos en más de una treintena de asociaciones. Además de su unión en una sola organización, consiguieron el reconocimiento y protección del Estado para el Cupón Pro-Ciegos: lotería con cuya venta se financia la labor social que realiza, además de proporcionar una actividad laboral a muchos de sus afiliados. Esta labor la realiza sin ánimo de lucro y no está dirigida únicamente a las personas con problemas visuales sino que de igual manera brinda su apoyo a personas con alguna otra capacidad diferenciada.

### **2.3 Delphi: el lenguaje de programación elegido**

Borland Delphi es una herramienta de desarrollo de aplicaciones en Windows [Delphi, 2002]. La forma de programación es muy fácil y puede llegar a tener grandes alcances como C++. El lenguaje de programación es una versión extendida de Pascal, pero basada en la Programación Orientada a Objetos.

Todas las partes comunes de una interfaz gráfica en Windows están incluidas en Delphi en forma de componentes, lo cual facilita la creación de nuevas aplicaciones, pues no es necesario escribir todo el código sino que únicamente se pegan los componentes en la forma principal, es como un programa para dibujar: *drag and drop*. Además, se pueden programar funciones y procedimientos específicos para cuando suceda un determinado evento.



Existen algunas tareas en las que los componentes estándares no son suficientes para poder realizar alguna aplicación. Para esto, existe un gran número de componentes avanzados disponibles en Internet como *freeware* o *shareware*. Cuando se tiene un poco más de experiencia en la programación de Delphi se pueden crear componentes propios con las características deseadas. Una de las ventajas que esto ofrece es que todos estos componentes pueden ser reutilizables, lo cual ahorra gran cantidad de tiempo, que puede ser utilizado en detalles de la aplicación. Delphi es un lenguaje de programación que no solamente agiliza el desarrollo de aplicaciones sino que también da una mejor perspectiva de la aplicación mientras esta se desarrolla.

## **2.4 SAPI. Speech Application Programming Interface**

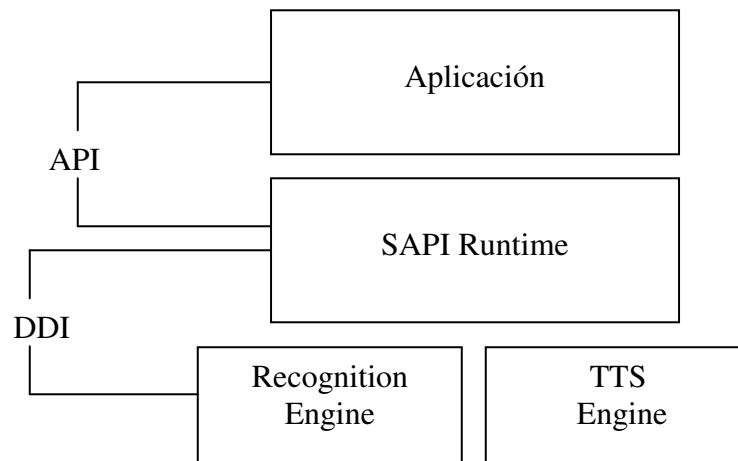
Hace un tiempo atrás el hecho de hacer hablar a una aplicación o que ésta pudiera entender lo que el usuario decía únicamente podía ser tocado desde el punto de vista de la ciencia ficción [8]. Afortunadamente, ahora podemos realizar aplicaciones que “hablen” y “escuchen” sin la necesidad de gran poder computacional. A través del tiempo la implementación de tecnología de habla dentro de las aplicaciones ha ido aumentando considerablemente y con ello, la creación de diferentes motores de síntesis y reconocimiento de voz. Desafortunadamente, no se ha seguido un estándar por lo que cada uno de los motores tiene su propio API (Application Program Interface) por lo que poder intercambiarlos era algo imposible.

A finales de 1995 Microsoft introdujo su Speech API (SAPI) como parte de la arquitectura de servicios abiertos de windows (Windows Open Services Architecture WOSA) con el fin

de simplificar el problema de compatibilidad [SAPI, 2002]. La figura 2.3 muestra a SAPI como mediador entre aplicaciones y maquinas de síntesis o reconocimiento.

El Speech API es de gran utilidad ya que reduce dramáticamente la cantidad de código que utiliza una aplicación encargada del reconocimiento y síntesis de voz. SAPI es una herramienta de programación que hace más accesible y robusta la implementación de la síntesis y/o reconocimiento de voz en un gran número de aplicaciones [11].

SAPI ofrece una interfaz de alto nivel que funciona como mediador entre una aplicación y los motores de voz. Utiliza todos los recursos de bajo nivel necesarios para el control y la administración en tiempo real de los diferentes motores de voz.



**Figura 2.3** Función de SAPI como mediador entre la aplicación y maquinas de síntesis y reconocimiento de voz [SS&SR, 2003]

Los dos tipos de **máquinas** de proceso que SAPI utiliza son:

1. Sistemas del tipo Text-to-speech (TTS): sistema que transforman cadenas y archivos de texto en sonido, por medio de voces artificiales.
  - Usos: este tipo de sistemas se utiliza para la lectura dinámica de texto, notificar al usuario de algún evento, proveer retroalimentación audible y para desarrollar aplicaciones dirigidas a personas con capacidades diferenciadas [SS&SR, 2003].
  - Limitaciones: dentro de las desventajas que se encuentra en TTS encontramos la mala calidad de la voz que produce pues es carente de emoción o énfasis. Hay algunos errores en la pronunciación de ciertas palabras o frases lo cual puede dar como resultado un ligero cambio en el sentido de la oración. Por otra parte se han desarrollado excelentes productos TTS diferentes a SAPI, sin embargo, su costo es demasiado elevado [SS&SR, 2003].
  
2. Reconocedor de voz: consiste en convertir una señal acústica de voz en una secuencia de palabras.
  - Usos: dado que el habla es el principal medio de comunicación entre los seres humanos se pretende utilizar para la comunicación con las computadoras. Dentro los usos que se han dado encontramos el dictado automático, sistemas automovilísticos, sistemas bancarios y para el desarrollo de sistemas para personas con capacidades diferenciadas [SS&SR, 2003].

- Limitaciones: dentro de las características que limitan al reconocimiento de voz encontramos el tamaño de vocabulario. Este problema impide que se pueda manejar cualquier dominio de discurso. De igual forma, la forma de hablar de los usuarios y los ruidos de medio ambiente restringen el rango posible de los hablantes. Cada uno de ellos podría tener algunos cambios en la voz como el acento, tonalidad diferente, hablar de forma continua o espontánea, situaciones que aún limitan a esta tecnología [SS&SR, 2003].

#### **2.4.1 API para sistemas TTS (Text to speech)**

SAPI nos permite crear aplicaciones que controlan *text-to-speech* (TTS) por medio de la interfaz de ISpVoice Component Object Model (COM). Dicha interfaz permite que las aplicaciones puedan realizar operaciones de síntesis de voz; pueden hablar cadenas y archivos de texto o incluso reproducir archivos de audio a través de la interfaz. Estas operaciones pueden ser realizadas de manera síncrona o asíncrona [SAPI, 2002].

Una vez que la aplicación ha creado un objeto ISpVoice, la aplicación solamente necesita a ISpVoice.Speak para generar una salida de audio en base al texto recibido. De igual forma, la interfaz ISpVoice ofrece algunos otros métodos que permiten cambiar de voz o modificar algunas propiedades como el volumen, tonalidad, la voz misma, etc.

SAPI se comunica con las aplicaciones por medio del envío de mensajes, *callback process* o eventos de Win32. Para TTS, los eventos en su mayoría son usados para una *output speech*. Con SAPI se puede crear léxico propio ya que las aplicaciones pueden ser programadas para realizar una pronunciación especial de alguna palabra en específico por

medio de los métodos que el API ofrece. Finalmente, existe una interfaz que nos permite modificar la salida de audio y enviarla a una dirección específica como el teléfono o cualquier otro tipo de hardware [SAPI, 2002].

#### **2.4.2 API para reconocimiento de voz**

Así como ISpVoice es la interfaz principal para la síntesis de voz, ISpReconContext es la principal interfaz para el reconocimiento de voz.

Una aplicación tiene la opción de elegir entre dos tipos de máquinas de reconocimiento. Un reconocedor compartido que posiblemente puede ser utilizado por otra aplicación de reconocimiento de voz, lo cual, es recomendado para la mayoría de las aplicaciones. Al crear una aplicación de este tipo se ejecutan diversos y variados métodos por medio de los componentes. Una vez creada la aplicación esta se encuentra a la espera de información que pueda procesar. Cuando la voz llega a la aplicación inmediatamente se procesa por medio del reconocedor para que posteriormente esta información sea traducida en un documento de texto.