

REPÚBLICA BOLIVARIANA DE VENEZUELA
LA UNIVERSIDAD DEL ZULIA
FACULTAD EXPERIMENTAL DE CIENCIAS
DIVISIÓN DE PROGRAMAS ESPECIALES
LICENCIATURA EN COMPUTACIÓN

**Arquitectura para la Edición de Mensajes de Texto en Dispositivos Móviles
Utilizando Reconocimiento Automático de Voz**

**Proyecto de Trabajo Especial de Grado presentado como
requisito para optar al título de Licenciado en Computación**

Autor: Gerardo José Curiel Orosco

Tutor: MSc. Gerardo Pirela

Cotutor: MSc. Carlos Rincón

Maracaibo, junio de 2012

**Arquitectura para la Edición de Mensajes de Texto en Dispositivos Móviles
Utilizando Reconocimiento Automático de Voz**

Gerardo José Curiel Orosco
CI. No.: 17.684.547
Teléfono: 58-261-7573725
Urb. La Paz
Correo electrónico: gcuriel@gmail.com

MSc. Gerardo Pirela
C.I. No.: 12.404.565
Correo electrónico: gepirela@fec.luz.edu.ve

MSc. Carlos Rincón
C.I. No.: 12.590.157
Correo electrónico: crincon@fec.luz.edu.ve

Curiel Orosco, Gerardo José. **“Arquitectura para la Edición de Mensajes de Texto en Dispositivos Móviles Utilizando Reconocimiento Automático de Voz ”**. Trabajo Especial de Grado. La Universidad del Zulia. Facultad Experimental de Ciencias. División de Programas Especiales. Maracaibo, Venezuela. 2012. 9p.

RESUMEN

El objetivo de esta investigación consistió en desarrollar una arquitectura para la edición de mensajes de texto utilizando reconocimiento automático de voz, que permita redactar correos y mensajes de texto de una manera mucho más rápida y eficiente, además de permitir una interacción intuitiva y accesible para invidentes y personas que no han aprendido a usar esta tecnología para comunicarse. La metodología para la elaboración del sistema es un modelo de desarrollo por fases siguiendo los lineamientos de los modelos de procesos incrementales. Principalmente, se revisó la literatura sobre motores de reconocimiento automático de voz determinando que Pocketsphinx, herramienta para reconocimiento continuo de voz en tiempo real, resultó la mas adecuada a los requerimientos de la investigación. Se diseñó y construyó la arquitectura del lado del servidor que procesaría las peticiones de reconocimiento de voz. Seguidamente, se desarrolló una interfaz móvil que se comunica con el componente servidor para enviar peticiones de reconocimiento de voz, caracterizada por ser intuitiva y extensible, capaz de asistir en el proceso de redacción de mensajes de texto usando reconocimiento de voz.

Palabras claves: Tecnología Inalámbrica, Móvil, Reconocimiento de Voz, Interfaces, Arquitectura

Correo electrónico: gcuriel@gmail.com

Curiel Orosco, Gerardo José. **“Arquitectura para la Edición de Mensajes de Texto en Dispositivos Móviles Utilizando Reconocimiento Automático de Voz”**. Trabajo Especial de Grado. La Universidad del Zulia. Facultad Experimental de Ciencias. División de Programas Especiales. Maracaibo, Venezuela. 2012. 9p.

ABSTRACT

The objective of this research is to develop a prototype interface for SMS editing using automatic speech recognition that allows writing emails and text messages quickly and efficiently, and allows a much more intuitive and accessible interaction for blind people and people who haven't learned how to use this technology to communicate. The system development methodology is a phased development model following the guidelines of incremental process models. An initial documental research was carried out to select the automatic speech recognition engine, which pointed to Pocketsphinx, a real-time continuous speech recognition system, as the most suited according to this investigation's requirements. A server-side architecture for speech recognition was designed and developed. Next, an intuitive, extensible mobile interface for sending speech recognition request to the server component was created, capable to assist in the SMS editing process using speech recognition.

Key Words: Wireless Technology, Mobile, Speech Recognition, Interfaces, Architecture

Email: gcuriel@gmail.com

ÍNDICE DE CONTENIDO

	Pág.
Resumen	3
Abstract	4
Introducción	5
Capítulo 1. El Problema	
Planteamiento del Problema	7
Alcance del Problema	8
Objetivos	8
Objetivo General	8
Objetivos Específicos	8
Justificación de la Investigación	8
Delimitación de la investigación.	9

INTRODUCCIÓN

El mundo de las telecomunicaciones ha ofrecido desde sus comienzos una serie de beneficios tecnológicos que se han aprovechado correctamente; sin embargo, con el transcurso del tiempo han surgido nuevas necesidades y nuevos requerimientos, entre ellos la movilidad y flexibilidad en las comunicaciones.

Actualmente los mercados y las tecnologías se encuentran en un cambio constante, las comunicaciones móviles han sido clave en esta revolución comercial y estratégica, trayendo movilidad a las comunicación en dos vías.

Uno de los métodos de comunicación que presentan los dispositivos móviles, complementaria a la comunicación vía voz, son los mensajes de texto. Sin embargo, la interacción con el teléfono para utilizar este método es ineficiente y lenta. Las empresas fabricantes de teléfonos tampoco han desarrollado mejoras o alternativas de interacción para la composición de mensajes de texto.

Afortunadamente, la lingüística computacional ha generado desarrollo científico teórico y aplicado que pretende incorporar habilidades lingüísticas a las computadoras, permitiendo la comunicación entre maquinas y personas a través de medios mas naturales, por ejemplo mediante una comunicación oral. Particularmente, los reconocedores de voz permiten a un computador transformar automáticamente una entrada de voz en su correspondiente forma textual, efectivamente posibilitando el desarrollo de un método alternativo de entrada de información a un dispositivo móvil, mas natural y lo mas parecida posible a como lo haría un ser humano.

Esta investigación se encuentra estructurada en siete capítulos, conformados cada uno de ellos de la siguiente manera:

El Capítulo I identifica el problema objeto de estudio mediante su debido planteamiento y señala los objetivos, justificación y alcance establecidos en esta investigación.

El Capítulo II está conformado por los antecedentes y los fundamentos teóricos sobre los cuales se asentó la investigación, ofreciendo de esta manera una mejor comprensión del tema abordado.

El Capítulo III describe la metodología empleada para el desarrollo de la investigación, así como los pasos para dar respuesta a los objetivos planteados.

En el Capítulo IV se definen las herramientas empleadas para la ejecución del sistema objeto de la investigación, así como los criterios de selección respectivos y los métodos empleados más sobresalientes.

El Capítulo V muestra la arquitectura del lado del servidor, los detalles de implementación y la integración con el motor de reconocimiento de voz. El Capítulo VI muestra la interfaz de usuario del sistema, gráfica y funcionalmente según su interacción con el usuario.

El Capítulo VII muestra el análisis de los resultados de la investigación mediante las pruebas efectuadas a las herramientas y a la arquitectura en general.

Finalmente, se plantean las conclusiones y recomendaciones obtenidas al culminar esta investigación, brindando la información y aportes necesarios a considerarse en investigaciones similares efectuadas a futuro.

CAPÍTULO 1

El Problema

1. Planteamiento del Problema

El uso de dispositivos móviles revolucionó la manera en que la humanidad se comunica, haciendo posible el realizar y recibir llamadas en cualquier lado, conectando y acercando a las personas cada vez más. Para muchos, en un principio, las llamadas por dispositivos móviles estaban fuera de su alcance, por el alto costo de esta nueva tecnología, en comparación con las tarifas que llevaban las compañías de telefonía fija. Es por esto que con la invención del envío de mensajes de texto, un formato reducido de comunicación, una creciente población adquirió teléfonos móviles para tener todos los beneficios de la movilidad de los celulares pero con bajo costo.

Sin embargo, la interacción necesaria para el envío de un mensaje de texto es ineficiente. La redacción en los teclados pequeños de los teléfonos actuales es inherentemente lenta e inconveniente, además, es una tarea que requiere de toda la atención del usuario; por ejemplo, crear un mensaje mientras se conduce un automóvil se convierte en una tarea extremadamente peligrosa, ya que el conductor podría no prestar toda la atención a la vía mientras redacta el mismo (1).

Las empresas productoras de teléfonos móviles y software, aún cuando han desarrollado innovaciones en el área de interfaz de usuario de estos dispositivos como el marcado rápido y discado de voz, se han quedado atrás en el tan popular envío de mensajes de texto, haciendo que la manera de relacionar al usuario con el dispositivo no haya mejorado desde que el mismo salió el mercado.

Con la realización de este trabajo se buscó el desarrollo de una interfaz y arquitectura alternativa de redacción de mensajes de texto, utilizando reconocimiento de voz, que reduzca al mínimo o incluso elimine la interacción necesaria del usuario para la redacción y envío de mensajes de texto utilizando dispositivos móviles.

Se requirió que la arquitectura fuese robusta y el servicio expuesto fuese fácil de consumir por dispositivos móviles, los cuales actualmente no tienen la capacidad de procesamiento necesaria para llevar a cabo el proceso de reconocimiento de voz por si solos.

2. Alcance del Problema

En investigación estuvo delimitada a la creación de una arquitectura para la redacción de mensajes de texto usando reconocimiento automático de voz y un componente móvil, el cual consume el servicio de reconocimiento de voz ofrecido por esta arquitectura.

3. Objetivos

3.1. Objetivo General

Construir una arquitectura para la edición de mensajes de texto en dispositivos Móviles utilizando reconocimiento automático de voz.

3.2. Objetivos Específicos

- Revisar literatura acerca del tema y aplicaciones existentes para el reconocimiento de voz e interacción con tales sistemas.
- Realizar un análisis comparativo de los motores de reconocimiento automático de voz aplicables.
- Seleccionar la herramienta de reconocimiento automático de voz.
- Diseñar la arquitectura del sistema.
- Diseñar la interfaz gráfica del sistema.
- Implementar el protocolo de comunicación entre la interfaz gráfica y el reconocedor automático de voz.
- Probar la funcionalidad de la aplicación.

4. Justificación de la Investigación

El interés original que motivó esta investigación es el de darle capacidades de reconocimiento de voz a sistemas de lingüística computacional desarrollados en los centros de investigación de La Universidad del Zulia. Estos esfuerzos de investigación requieren de un sistema de procesamiento de emisiones de voz con la capacidad de integración con múltiples tipos de interfaces de usuario, como por ejemplo un dispositivo móvil e incluso otros sistemas lingüísticos.

La capacidad de respuesta textual basada en una entrada de habla proporcionaría otra forma de comunicación entre el dispositivo de cómputo y el hombre que podría ser útil incluso para personas con necesidades especiales (invidentes) y para facilitar en general el uso de tecnología, sustituyendo las interfaces gráficas basadas en dispositivos físicos de entrada por interfaces naturales basadas en voz.

Entre los beneficios que ofrecen sistemas de este tipo están la creación de sistemas de redacción de texto, sistemas del tipo “asistente personal”, sistemas tipo kioscos de información interactivos, interfaces de usuario para personas invidentes, entre otros.

La integración de reconocimiento de voz en sistemas de lingüística computacional permite aumentar la interactividad con el usuario y mejorar la accesibilidad, consiguiendo que el contenido e incluso las órdenes sean reconocidas, lo que posibilita la construcción de sistemas con los que los usuarios pueden interactuar de manera natural.

5. Delimitación de la investigación.

La investigación se llevó a cabo en la Unidad de Lenguajes y Modelos Computacionales de la Facultad Experimental de Ciencias de la Universidad del Zulia y su duración fue de 16 meses. Se inició en el mes de noviembre de 2011 y culminó en marzo de 2012. La línea de investigación a la cual estuvo anexa fue: Lingüística Computacional.