

**UNIVERSIDAD NACIONAL DE SAN ANTONIO
ABAD DEL CUSCO**

**FACULTAD DE INGENIERÍA ELÉCTRICA, ELECTRÓNICA,
INFORMÁTICA Y MECÁNICA**

**ESCUELA PROFESIONAL DE INGENIERÍA INFORMÁTICA Y DE
SISTEMAS**



TESIS DE GRADO

**“CONVERSOR DE TEXTO ESCRITO A VOZ POR
CONCATENACIÓN DE DIFONEMAS PARA EL IDIOMA
QUECHUA”**

PRESENTADO POR: Br. Gabriela Sánchez Espirilla.

PARA OPTAR AL: Título Profesional de Ingeniero Informático y de Sistemas.

ASESOR: Ing. Luis Beltrán Palma Ttito.

CO – ASESORES: Ing. Antonio Jesús Bonafonte Cávez.
Lic. Juan Antonio Cruz Tello.

Cusco – Perú

2016

Resumen

La presente tesis tiene como fin implementar un conversor de texto a voz, herramienta tecnológica, utilizando la técnica de concatenación de unidades básicas “difonemas”, para la síntesis de voz del idioma¹ quechua con alfabeto pentavocálico y frecuente en la región Cusco. Se identificaron dos momentos claves para llevar a cabo el trabajo.

Primero, se construyó el corpus de voz y la base de datos para dicho idioma. Para lo cual, inicialmente, se identificaron los difonemas válidos y propios del idioma quechua; para ello, previamente, se realizó una aproximación al estudio de la lingüística del idioma. Seguidamente, se realizaron las grabaciones de las frases y palabras portadoras de dichos difonemas; esto para el etiquetado y extracción de los mismos. La información obtenida en este punto sirvió para la construcción de la base de datos de difonemas para el idioma quechua.

Segundo, se construyó el sintetizador de voz. Para lo cual, se identificaron dos bloques, del procesamiento del lenguaje natural, y del procesamiento digital de la señal. El primer bloque comprende la extracción y el análisis de la oración para la elaboración del fichero de transcripción fonética de los difonemas que componen dicha oración. El segundo bloque se encarga de procesar la transcripción obtenida en el bloque anterior con la ayuda de diferentes algoritmos para convertirla en voz. Este utilizó una técnica llamada TD-PSOLA y otras para suavizar la unión de difonemas.

La creación de la herramienta en mención; así como, su uso y futura mejora; aprovecha las ventajas de las tecnologías del habla y conversión de texto a voz para aspirar ser un ejemplar de aquellas escasas tecnologías de información y

¹ Para la presente tesis, los vocablos “idioma” y “lengua” tienen el mismo significado, según la Lingüista Zavala Virginia en sus trabajos de investigación.

comunicación para idiomas confinados al hogar y a generaciones mayores² como el quechua. Este busca democratizar el acceso a la información y los avances tecnológicos a personas quechua hablantes buscando desarrollar capacidades. Además, ésta busca rechazar la idea que el idioma quechua resulta inadecuado para transmitir los avances y desarrollos de las ciencias y las artes. Dicho de esta manera, es notable la ausencia de herramientas tecnológicas para la enseñanza del idioma quechua en diferentes campos donde podría complementar los aprendizajes, para ejemplo están los proyectos de alfabetización y educación intercultural bilingüe, sistemas de lectura para invidentes, entre otros.

En conclusión, se implementó un conversor de texto escrito a voz utilizando la estrategia de concatenación de difonemas para el idioma quechua de la familia lingüística denominada Cusco – Collao.

Palabras clave: Conversor de texto a voz, concatenación de difonemas, síntesis de voz, quechua, corpus de voz, transcripción fonética, TD-PSOLA.

² Rivera, Cecilia. Colaboradora del Proyecto RunasimiNet. “El quechua y sus hablantes: En la Pontificia Universidad Católica de Perú”.