

Traductor de texto y voz a lengua de señas ecuatoriana a través de un avatar implementado para dispositivos Android

Translator of text and voice to Ecuadorian sign language through an avatar implemented for Android devices

Myriam Carguacundo^a, Patricia Constante^b

^{a,b} Departamento de Ciencias Energía y Mecánica, Universidad de las Fuerzas Armadas ESPE Extensión Latacunga

Correspondencia Autores: mpcarguacundo@espe.edu.ec; pnconstante@espe.edu.ec

Recibido: 25 de septiembre del 2018, **Publicado:** xxxxxxxx

Resumen— El traductor tiene como finalidad constituirse en una herramienta de aprendizaje del lenguaje de señas ecuatoriano y así ayudar a las personas sordomudos a que puedan comunicarse dentro de este núcleo familiar y su entorno. Este sistema tiene una interfaz amigable con un avatar de fisonomía humana el cual realiza las señas y los movimientos respectivos de 120 palabras de uso frecuente; este traductor ha sido desarrollado en Android Studio para ser ejecutado en dispositivos Android. En la pantalla principal del menú de la interfaz se tiene dos opciones de ingreso por teclado de una, dos o tres palabras; o por reconocimiento de voz de una frase que contenga tres palabras, en ambos casos se ha limitado la frase a 25 caracteres; se pretende entonces brindar una herramienta para contribuir al aprendizaje de personas sordomudas tanto niños como adultos y principalmente de quienes sin tener ningún tipo de discapacidad están interesados en aprender lenguaje de señas ecuatorianas LSEC de una manera entretenida y sencilla.

Palabras Claves— Lenguaje de señas, avatar, aplicación Android.

Abstract— The purpose of the translator is to become a tool for learning the Ecuadorian sign language in order to help families of deaf people to communicate inside this familiar nucleus and environment. This system has a friendly interface with an avatar of human appearance which makes the respective signs and movements of 120 words of frequent use; this translator has been developed in Android Studio to be executed in Android devices. In the main screen of the interface menu there are two input options, by keyboard of one, two or three words, or by voice recognition of a phrase containing three words, in both cases the phrase has been limited to 25 characters; therefore it is intended to provide a tool to contribute the language learning for deaf-mute people with talk and listen disabilities and mainly those who without any disability are interested in learning sign language in an entertaining and simple way.

Keywords— Sign language, avatar, Android application.

I. INTRODUCCIÓN

El lenguaje de señas es un sistema lingüístico [1] se basa en las expresiones faciales, movimientos corporales y gestuales que usan las personas sordomudas para comunicarse [2]. En el Ecuador actualmente existen instituciones con docentes especializados que prestan ayuda para facilitar la

comunicación en el lenguaje de señas [3] entre personas con discapacidad y su entorno. Estas instituciones no tienen tecnologías de comunicación y debido a su infraestructura puede enseñar lenguaje de señas a un número limitado de ellas, según el Consejo Nacional de Discapacidades el total de personas con alguna discapacidad es de aproximadamente 447.294 de las cuales el 14.13% que equivale a 63.121 son personas con discapacidad auditiva [4].

Los lenguajes de señas no son universales, cada lenguaje de signos es un lenguaje distinto, existen por ejemplo el Lenguaje de Señas Americano (LSA) cuyo pariente más cercano es el Lenguaje de Señas Francés (LSF), la Lengua de Signos Brasileña (LSB) como la Lengua de Señas Quebequense (LSQ) y la Lengua de Señas Mexicana (LSM), que están genéticamente relacionadas [5], así también existe la Lengua de Señas Ecuatoria (LSEC) en cuyo diccionario [6] se basa este trabajo. En la actualidad la tecnología ha ido evolucionando rápidamente y se utiliza todo tipo de dispositivos electrónicos [7] como tecnología asistida que permite a los individuos con discapacidades verse involucrados en las mismas actividades de sus pares sin discapacidad utilizando una herramienta electrónica [8]. A nivel nacional se ha desarrollado diferentes tipos de guantes que realizan señales de señas para discapacidad auditiva [9], en [10] se construye un prototipo electrónico capaz de traducir el lenguaje de señas del abecedario de una persona sordomuda al lenguaje de letras; en [11] se ha implementado el lenguaje de señas elementales traducido a lenguaje escrito mediante software libre. También se han desarrollado aplicaciones como [12] para dispositivos móviles Android utilizando redes neuronales; en [13] implementaron un sistema traductor, mediante visión artificial en un ambiente controlado utilizando software Labview; y [14] realizó un sistema de codificación de señales de voz a texto usando Matlab mediante series de

Fourier, esto evidencia el interés en el desarrollo de tecnología al servicio de la comunidad, donde es primordial un ambiente favorable para la comunicación entre personas sordas y personas oyentes [15] para convivir en sociedad [16].

La tecnología y aplicaciones desarrolladas para discapacidad auditiva se han centrado en traducir una seña en letra, pero no tiene la traducción de una letra a una señal, con lo cual se lograría una comunicación bidireccional. Se pretende entonces diseñar un sistema traductor de texto y voz a señales de señas utilizando un dispositivo Android que ayude a la comunicación y aprendizaje de las personas con discapacidad auditiva entre ellas con sus familias y su entorno, se ha seleccionado dispositivos Android ya que el 31.6% de las personas tienen este dispositivo [17], el cual permitirá instalar la aplicación para poder usarla.

II. MATERIALES Y MÉTODOS

Los materiales que se utiliza en este sistema traductor son:

- Software de Android Studio que sirve para la programación del sistema.
- Software Daz Studio 4.9 Pro es para la realización del avatar con sus respectivas señas de manos para cada palabra que se encuentra en sistema.
- GOOGLE NOW para reconocimiento de voz.
- Un celular Samsung S5.

Lenguaje de señas

El lenguaje de señas se basa en las expresiones faciales, movimientos corporales y gestuales para comunicarse. Todo tipo de lenguaje tiene claramente definido las estructuras gramaticales, al igual que el sistema de lenguaje [1], como cada lenguaje tiene variaciones tanto en léxico como en gramática [5], es importante el desarrollo de herramientas de LSEC.

Avatar

El avatar es una animación tridimensional de la representación de un ser humano el cual ejecutará el movimiento de las señas según corresponda a las palabras o frases ingresadas.

Ergonomía visual y los colores

Los colores transmiten emociones e influyen en la percepción de la realidad de cada persona. La diferente gama de colores con un determinado tipo de letra, tienen el poder de transmitir un mensaje o emoción concreta [18], se ha optado por el uso de colores neutros.

Procesamiento de la voz

En el preprocesamiento de voz (fig. 1) se debe

condicionar la señal de entrada para que sea procesada en el reconocimiento mediante redes neuronales artificiales [19].

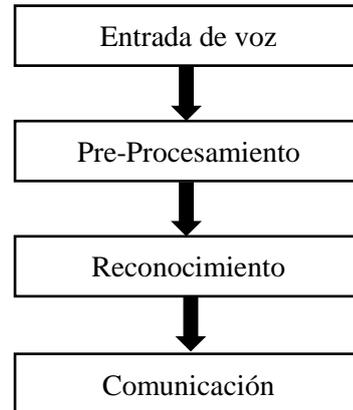
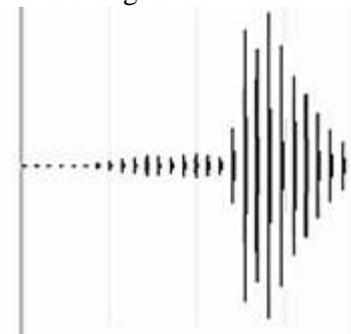


Fig. 1. Componentes del procesamiento de voz [17].

- Preprocesamiento: Convierte la entrada de voz a una forma que el reconocedor pueda procesar.
- Reconocimiento: Identifica lo que se dijo mediante redes neuronales artificiales.
- Comunicación: Envía la palabra al sistema que lo requiere.

a) Entrada de la voz: Se activa el motor de reconocimiento de voz para ingresar la palabra. La palabra de la señal ingresa de forma análoga al sistema como se indica en la Fig. 2.



b) Pre-procesamiento: La señal ingresada se convierte en señal digital. Luego pasar al dominio del tiempo y al dominio de la frecuencia. En la figura 3 se observa la representación en el dominio del tiempo y frecuencia respectivamente. Se debe trocear el espectrograma de la señal en varias partes para luego enviar a una red neuronal artificial.

c) Reconocimiento: En esta fase recibe la señal del preprocesamiento y se envía a una red neuronal recurrente que utiliza GOOGLE NOW que consta 16000 redes de varias capas las cuales tienen retroalimentación que realizan diferentes acciones como por ejemplo una capa intenta encontrar los elementos que le componen como vocales, consonantes

y a continuación otra capa identifica los sonidos que están en cada grupo para obtener finalmente la palabra [20].

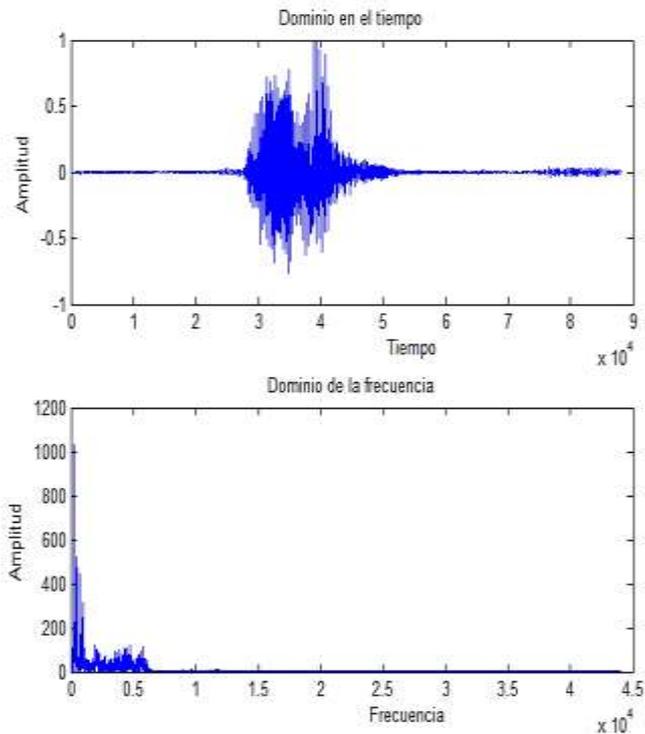


Fig. 3. Dominio del tiempo y frecuencia: a) Dominio del tiempo, b) Dominio de la frecuencia.

d) Comunicación: En esta última fase después de haber realizado el reconocimiento envía la palabra en forma de texto.

Diseño conceptual

El objetivo del sistema traductor es visualizar el avatar con las respectivas señales con cada uno de sus movimientos de acuerdo con la palabra ingresada ya sea por teclado o por voz mediante un dispositivo móvil para esto se establece las especificaciones técnicas requeridas por las necesidades del usuario, las cuales son:

- La función del sistema es visualizar el avatar con la señal de manos de acuerdo con cada palabra.
- Las dimensiones del avatar deben ser proporcionales al cuerpo humano.
- El avatar debe mover la parte superior del cuerpo cabeza, cuello, brazos y manos.
- El sistema es compatible en dispositivos Android.

Diseño del sistema del avatar

El diseño del avatar se lo realiza como la fisonomía de un humano teniendo en cuenta cada una de las partes que van a estar en movimiento para realizar cada una de las señas. El diseño del avatar se lo realiza en el Software de DAZ STUDIO.

El software cuenta con varias plantillas para la elección del avatar en este caso se eligió el Génesis 3 Male como se indica en la Fig. 4.a la cual se la ha editado hasta lograr la apariencia requerida como en la Fig. 4.b., donde también se ha propuesto la posición inicial de ejecución de los movimientos que va a realizar según la palabra.



a) Plantilla del génesis 3 Male
b) Avatar modificado
Fig. 4. Avatar

Implementación de la aplicación Android

La aplicación del sistema traductor debe ser sencilla de operar con el objetivo que pueda ser utilizada por niños, niñas y adultos por lo que se ha considerado el diseño de una interfaz amigable y con colores llamativos. Se integran todas las ventanas múltiples que se encuentran en la aplicación cada una de ellas realizando diferente acción (fig. 5).

En la pantalla principal de inicio del sistema se tiene dos botones de salir y del menú. La pantalla de menú tiene dos opciones que es el ingreso de la frase por teclado o por voz y cada opción consta su respectivo botones de ingreso, y el botón de salir es para cerrar la aplicación. Al accionar un botón de la ventana de menú se abre otra ventana que tiene la barra para ingresar las palabras en el caso que se acciono el ingreso por teclado en cambio en la de voz se tiene un botón para ingresar las palabras, también consta de tres botones: el botón traducir al ser accionado se visualiza el avatar con su respectiva señal de la palabra ingresada, el botón de borrar limpia la pantalla y el de salir regresa a la pantalla del menú.

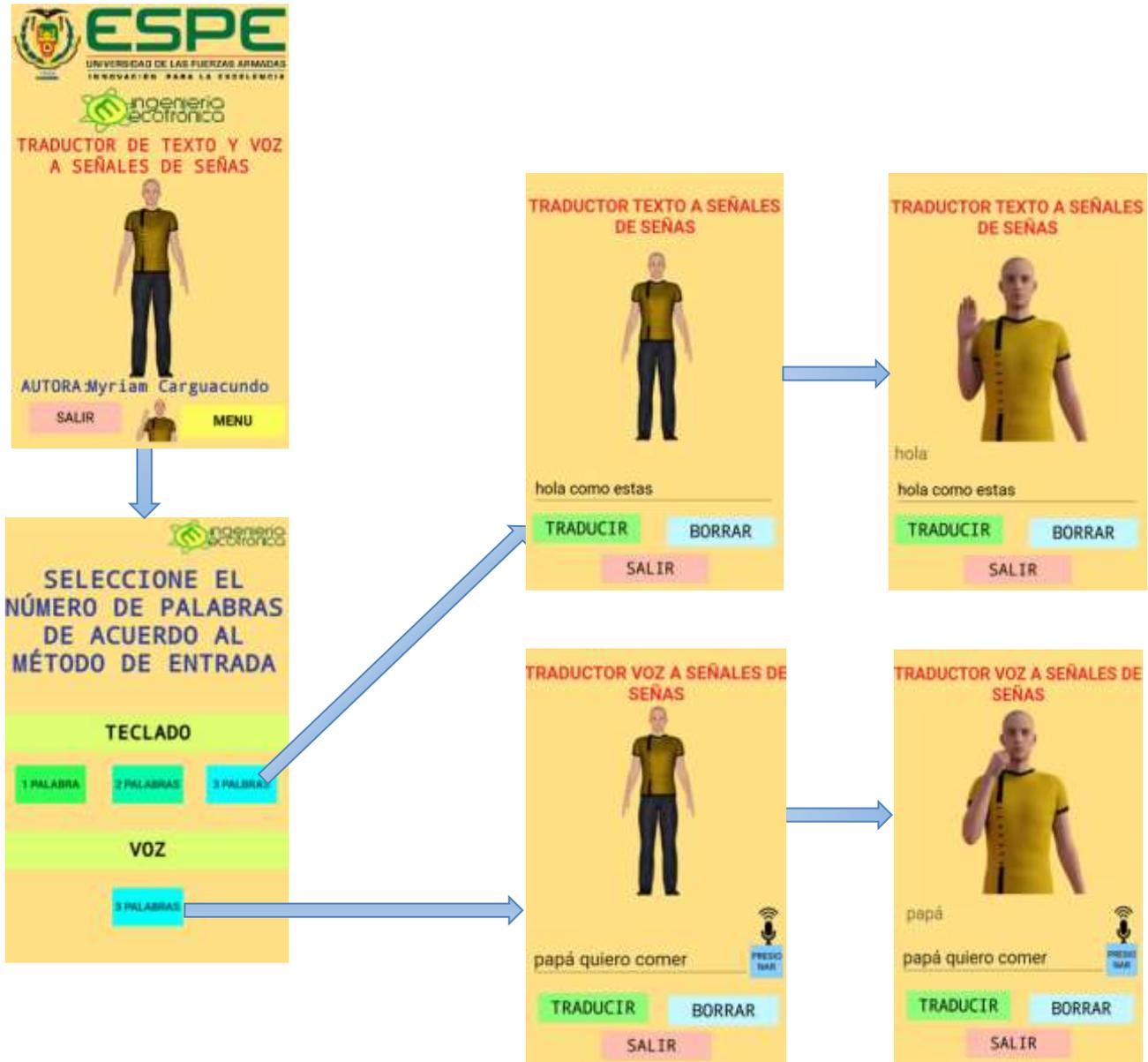


Fig. 5. Implementación de la aplicación.

El sistema consta de 120 palabras cada una con su respectiva señal, se ha considerado importante incluir el abecedario, los días de la semana, meses del año, número del uno al doce, saludos entre otros.

III. RESULTADOS

Se probó la aplicación del traductor con 10 personas, a cada persona se le asignó 17 palabras seleccionadas de la base de datos de manera aleatoria. Dentro de los usuarios se encontraban adultos, jóvenes y niños (Fig. 7).

Pruebas de reconocimiento de voz

La realización de esta prueba tiene como objetivo es comprobar que la palabra obtenida del motor de reconocimiento de voz se compare con una letra de la base de datos y se obtenga la señal respectiva. Para esta

prueba se obtuvo un total de 170 palabras.



Fig. 7. Traductor usado por un niño.

Los resultados de la prueba de voz entregaron un reconocimiento de 162 palabras mientras que 8 palabras no fueron reconocidas, con estos datos se calcula el error que es el 5% y un 95% de eficiencia.

Pruebas base de datos

Para comprobar el funcionamiento de la base de

LUNES



MARTES



5



datos, se realiza varias pruebas que se ingresa una, dos o tres palabras, las mismas que se comparan con las

palabras de la base de datos, si existe se visualiza el avatar con la señal respectiva caso contrario se presenta al usuario el mensaje de “NO EXISTE” y continúa con las demás palabras si fuera el caso. El tiempo de visualización del avatar depende de la duración de cada señal. Los datos generales que se obtuvo de las pruebas se observan en la tabla 1.

Tabla 1.
Datos obtenidos en las pruebas de base de datos.

Número de palabras	Se visualizó	No se visualizó
Una	164	6
Dos	163	7
Tres	162	8
TOTAL	163	7

El resultado de la prueba de la base de datos es de 163 palabras que se visualizaron correctamente en el avatar y 7 palabras no se visualizaron, dentro de las razones se encuentra que el usuario no logró escribir correctamente o no colocó tilde en las palabras; se obtuvo un error del 4% y una confiabilidad del 96%.

IV. CONCLUSIONES

El sistema traductor de texto y voz a lenguaje de señas a través de una interfaz gráfica con un avatar mediante dispositivos Android tiene un porcentaje de confiabilidad del 95% en la visualización del avatar con su respectiva señal.

Se ha logrado implementar un traductor de lengua de señas ecuatoriano el cuál pretende contribuir al aprendizaje del lenguaje no solo para personas sordomudas sino como herramienta de aprendizaje de familiares, amigos y personas interesadas en aprender LSEC y de esta manera lograr una verdadera inclusión de las personas con discapacidad.

La programación del sistema traductor se desarrolló en el Software libre de Android Studio para versiones

de 4.5 en adelante de Android.

El diseño de la interfaz gráfica se realizó con un entorno amigable, que sea de fácil uso para el usuario de diferentes edades.

Las señas del avatar se realizaron en DAZ Studio que es un software libre y de fácil manipulación que permite realizar cada una de las señas de las diferentes palabras que se encuentra en el sistema traductor.

La base de datos cuenta con 120 palabras entre las cuales están el abecedario, días de la semana, meses, números, entre otras las cuales han sido previamente realizadas en el avatar, se pretende en un futuro incrementar esta base de datos.

Se ha considerado implementar reconocimiento de voz, para llamar la atención de personas sin discapacidad auditiva y motivar el uso del mismo para aprender lenguaje básico de señas ecuatorianas.

REFERENCIAS

- [1] Jarque M. (2012) Las lenguas de signos: su estudio científico y reconocimiento legal. Anuari de Filologia. Estudis de Lingüística, num. 2, p. 33-48. [En línea]. Disponible en: <http://hdl.handle.net/2445/96020>
- [2] Pérez, Maria. (2012). Influencia de la aprobación previa de la lengua de señas en la comprensión lectora de los niños sordos. [En línea]. Disponible en: <https://dspace.ups.edu.ec/bitstream/123456789/3443/1/UPS-OT03016.pdf>
- [3] Ministerio de Educación. (2011). [En línea]. Disponible en: <https://educacion.gob.ec/instituciones-de-educacion-especial/>
- [4] MinConsejo Nacional para la igualdad de discapacidades registradas - CONADIS. (2018). [En línea]. Disponible en: <https://www.consejodiscapacidades.gob.ec/wp-content/uploads/downloads/2018/03/index.html>
- [5] Wilcox, S., & Occhino, C. (2016, 11). Historical Change in Signed Languages. Oxford Handbooks. [En línea]. doi:10.1093/oxfordhb/9780199935345.013.24
- [6] Diccionario de Lengua de Señas Ecuatoriana. (n.d.). [En línea]. <http://plataformaconadis.gob.ec/diccionario/>
- [7] Grupo BBVA. (2015). 5 ejemplos de tecnología para personas ciegas: más allá del Braille. [En línea]. Disponible en: <https://www.bbvaopenmind.com/tecnologia-para-invidentes-mas-alla-del-braille/>
- [8] Borja, Cynthia. (2012, Diciembre). ¿Qué es la tecnología asistiva? [En línea]. Disponible en: https://www.usfq.edu.ec/publicaciones/para_el_aula/Documents/para_el_aula_04/0006_para_el_aula_04.pdf

- [9] Torrejón M. (2016). Traductor de Lenguaje de Señas con Guante Electrónico y Aplicación Android. Rev Inv Tec [En línea]. vol.4, n.1 pp. 99-104 . Disponible en: http://www.revistasbolivianas.org.bo/scielo.php?script=sci_arttext&pid=S2306-05222016000100014&lng=es&nrm=iso. ISSN 2306-0522.
- [10] Pozo H., & Espinosa P. (2013). Traductor de lenguaje de señas por un guante electrónico. [En línea]. Disponible en: <https://dspace.ups.edu.ec/bitstream/123456789/4211/1/UPS-CT002598.pdf>
- [11] Almeida L., & Viteri P. (2016). Intérprete del lenguaje de señas elementales por un par de guantes. [En línea]. Disponible en: <http://repositorio.espe.edu.ec/xmlui/bitstream/handle/21000/12451/T-ESPEL-MEC-0078.pdf?sequence=1&isAllowed=y>
- [12] Veintimilla, M. (2014). [En línea]. Disponible en: <https://repositorio.espe.edu.ec/bitstream/21000/8873/1/T-ESPE-048054.pdf>
- [13] Edwin Chiguano, Nathaly Moreno, & Luis Corrales. traductor de lenguaje de señas a texto mediante visión artificial en un ambiente controlado. (2012). [En línea]. Disponible en: http://bibdigital.epn.edu.ec/bitstream/15000/4924/1/PAPER_CHIGUANO_MORENO.pdf
- [14] Cárdenas, J. (2012). Sistema de codificación de voz a texto usando Matlab. [En línea]. Disponible en: http://repositorio.uta.edu.ec/bitstream/123456789/2907/1/Tesis_t771e.c.pdf
- [15] Bernal S. (2014). Sistema Visc Sistema Inteligente De Reconocimiento De Voz Para La Traducción Del Lenguaje Verbal A La Lengua De Señas Colombiana. Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación. ISBN: 978-84-7666-210-6 – Artículo 349. [En línea]. Disponible en: <https://www.oei.es/historico/congreso2014/memoriactei/349.pdf>
- [16] Riotordo, A. (2012). ¿Cultura Sorda? - A.P.A.S.U. (23 de enero de 2018). [En línea]. Disponible en: <http://www.apasu.org.uy/b-sordos/d-cultura-sorda>
- [17] INEC. (2016). [En línea]. Disponible en: <http://www.ecuadorencifras.gob.ec/en-cinco-anos-se-quintuplicaron-los-usuarios-de-telefonos-inteligentes/>.
- [18] Biesa, M. (2015). MIQUE. [En línea]. Disponible en: <http://www.mique.es/la-importancia-del-color-en-el-diseno-grafico/>
- [19] UDLAP. (2016). Reconocimiento de voz. [En línea]. Disponible en: <http://ict.udlap.mx/people/ingrid/Clases/IS412/>
- [20] Nación. (2015). La inteligencia Artificial. [En línea]. Disponible en: <http://www.lanacion.com.ar/1810095-la-inteligencia-artificial-el-recurso-de-google-para-detener-el-spam-en-gmail>