

*investigación*



Actualidad TIC  
de encuestas telefónicas  
automatizadas

# ODEC ADET



y el encuestado se limita a contestar una de las posibles opciones. La respuesta emitida no tiene por qué seguir exactamente la sintaxis de las posibles contestaciones, sino que existe un margen para la omisión o inserción de palabras. El sistema es capaz de identificar cuál es la palabra relevante y, en consecuencia, transcribir la opción escogida. Éste sería el comportamiento lógico si la encuesta fuese realizada por una máquina, ya que se elimina la posibilidad de que el encuestado incurra en los comportamientos problemáticos.

## La tecnología desarrollada por el grupo se está aplicando con el fin de crear prototipos y productos que puedan ser útiles a las empresas

### b. Segunda Fase. Desarrollo del prototipo y evaluación.

En una segunda fase se procedió al desarrollo del prototipo. El objetivo en esta segunda fase fue el de construir un prototipo preliminar y evaluar cuál era su rendimiento. Para ello, se seleccionó una encuesta real con la que llevar a cabo una evaluación. La encuesta escogida tenía como tema "El debate político sobre el estado de la nación". Como muestras de voz para la evaluación se utilizaron dos corpus:

- **Corpus ODEC:** De los 75 ficheros de audio proporcionados por ODEC, cada uno de los cuales contenía la grabación de toda una encuesta, se seleccionaron 6 para evaluación. El criterio para seleccionarlas fue comprobar que en ellos se seguía, más o menos, el paradigma pregunta/respuesta. Con las encuestas seleccionadas se procedió a la extracción manual de los segmentos de voz correspondientes a las contestaciones de las preguntas de evaluación. Para cada muestra de audio extraída fue necesario transcribir manualmente cuál era la respuesta escogida.

- **Corpus ITI:** Utilizando el prototipo se procedió a que personal del ITI realizase la encuesta. De todas las encuestas adquiridas se seleccionaron al azar 6 correspondientes a personas distintas. Para cada muestra de audio fue necesario también transcribir cuál era la respuesta escogida.

FIGURA 2



En la evaluación se experimentaron con distintos modelos del reconocedor de habla. Todos estos modelos fueron generados siempre de forma automática a partir de las posibles respuestas de cada pregunta. La diferencia entre unos modelos u otros radica en la metodología utilizada para su generación. Esta metodología influye en la flexibilidad que tienen el modelo de admitir mayor inserción u omisión de palabras a la hora de emitir las respuestas.

En las siguientes tablas se muestra el acierto del prototipo (en tanto por cien (%) y número absoluto (Abs)) en asignar la respuesta correcta en función de lo que dice el encuestado. Como ya se ha indicado no siempre el encuestado contesta exactamente siguiendo la sintaxis de la respuesta. Lógicamente esto es más frecuente en el corpus ODEC.

En la siguiente tabla se muestran los mejores resultados. En la primera fila, se muestra los resultados obtenidos utilizando modelos generados de forma totalmente automática. Éste sería el caso de un sistema totalmente automático exento de supervisión. Estos resultados pueden ser mejorados añadiendo una mínima supervisión. Para algún tipo de pregunta los modelos que son más flexibles obtienen mejores resultados, mientras que para otros tipos ocurre al contrario. En la segunda fila de la tabla se muestran los mejores resultados utilizando en cada pregunta el modelo que se comporta mejor.

	Corpus ITI			Corpus ODEC		
	Nº muestras	Acierto	Abs.	Nº muestras	Acierto	Abs.
Sin supervisión	396	83.8%	331	227	69.2%	157
Con supervisión	396	87.4%	346	227	78.0%	177

Como se observa, una mínima evaluación de los modelos generados automáticamente permitiría escoger los modelos apropiados para cada pregunta, repercutiendo en una mejora de las prestaciones del sistema.

### c. Tercera fase. Desarrollo del prototipo final.

Una vez evaluado el prototipo preliminar de forma satisfactoria se procedió, en la tercera fase, a desarrollar el prototipo final incorporándole algunas mejoras.

Los objetivos en esta tercera fase del proyecto pueden resumirse en dos aspectos importantes:

1. Mejorar el rendimiento del prototipo desarrollado en la segunda fase. Para ello se propusieron dos líneas de actuación:

- Mejora de los modelos del reconocedor.
- Incluir la estimación de medidas de confianza en el prototipo.

Este aspecto permite al prototipo reaccionar ante sus propios errores, lo que repercute en una mejora de sus prestaciones. Por ejemplo, en el caso de que no se disponga de una alta fiabilidad de que la respuesta reconocida es la que ha elegido el encuestado, podría repetirse la pregunta.

2. Desarrollar un demostrador del prototipo que sea utilizable a través de Internet.

El prototipo obtenido como resultado de esta fase está accesible en la página web <http://prhltedemos.iti.upv.es/~adet>. En la siguiente sección se describen sus características más importantes.



FIGURA 3

### Prototipo Desarrollado

El sistema prototipo ha sido desarrollado para poder ser accedido directamente a través de Internet. En función de esto, el prototipo se compone de dos partes: la del cliente y la del servidor. [Ver figura 2].

**La interfaz ha sido implementada en JAVA, lo que la hace compatible con la mayoría de las plataformas existentes.**

El cliente es una interfaz con la que el usuario puede interactuar en el desarrollo de la encuesta. En la misma se visualizan las preguntas y respuestas reconocidas, así como la respuesta seleccionada por el sistema. La interfaz ha sido implementada completamente en JAVA, lo que la hace compatible con la mayoría de las plataformas existentes. [Ver figura 3]

La parte del servidor incorpora, en líneas generales, tres módulos diferentes:

1. El primer módulo se encarga del reconocimiento de las respuestas de voz telefónica para cada una de las preguntas. Este módulo obtiene, de cada pregunta, la respuesta emitida por el encuestado, y en función de ella, almacena la opción escogida como respuesta. Este módulo está basado en un motor de

reconocimiento del habla continuo, denominado ATROS (Automatically Trainable Recognizer Of Speech) desarrollado por el PRHLT.

2. El segundo módulo se especializa en realizar las preguntas de la encuesta. En este caso, reproducir ficheros de audio previamente grabados con cada una de las posibles preguntas.

3. El tercer módulo incorpora lo que se conoce como "medidas de confianza", cuyo objetivo es dotar al motor de reconocimiento de la capacidad de asignar a cada unidad transcrita (palabra) una "medida de confianza". La "medida de confianza" es un indicador de la certeza que se tiene sobre la corrección de la unidad transcrita de forma automática. Esta información es relevante, ya que permite al sistema reaccionar ante posibles errores del reconocimiento.

### Conclusiones

En este artículo se ha presentado el proyecto ADET. El prototipo desarrollado en este proyecto es capaz de realizar encuestas telefónicas generando los modelos necesarios de forma totalmente autónoma y automática. Este aspecto dota al sistema de una gran flexibilidad y permite su aplicación a empresas que no tengan conocimientos en este área de investigación. Podría ser idóneo en aplicaciones ya incipientes, como son encuestas telefónicas automáticas que se realizan, por ejemplo, para obtener datos sobre la satisfacción de clientes ante un servicio dado. ■