

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/236123076>

Satisfacción de usuarios del ámbito de la traducción en el uso de sistemas de búsqueda multilingüe de respuestas como recurso...

Conference Paper · October 2012

CITATIONS

0

READS

40

2 authors:



[María Dolores Olvera-Lobo](#)

University of Granada

93 PUBLICATIONS 245 CITATIONS

[SEE PROFILE](#)



[Juncal Gutiérrez-Artacho](#)

University of Granada

43 PUBLICATIONS 70 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Translation Training: Professional Approach and Information Technologies (Aula.Int) [View project](#)



Bibliometrics, webometrics and audiovisual translation. [View project](#)

All content following this page was uploaded by [Juncal Gutiérrez-Artacho](#) on 20 May 2014.

SATISFACCIÓN DE LOS USUARIOS EN LA BÚSQUEDA MULTILINGÜE DE RESPUESTAS COMO RECURSOS DE INFORMACIÓN TERMINOLÓGICA

María Dolores Olvera Lobo
Universidad de Granada
molvera@ugr.es

Juncal Gutiérrez Artacho
Universidad de Granada
juncalgutierrez@ugr.es

RESUMEN

Con el rápido crecimiento de Internet y el desarrollo de las nuevas tecnologías en los últimos años, los sistemas de búsquedas de respuestas (SBR) se han convertido en una alternativa a los tradicionales sistemas de Recuperación de Información. Aunque existe una prolífera producción sobre estos sistemas, pocos son los trabajos que se han desarrollado en la evaluación centrada en el usuario. El presente trabajo se centra exclusivamente en la evaluación de los SBR multilingüe ya que permite al usuario acceder a información terminológica no disponible en su lengua, y en la evaluación centrada en el usuario para entender las necesidades del usuario e identificar las dimensiones y factores en el desarrollo de un sistema de información con el fin de mejorar su aceptación.

El objetivo es conocer el grado de satisfacción del usuario para el SBR multilingüe HONqa (inglés, francés e italiano) utilizando la herramienta de los investigadores ONG et al. (2009) basado en un examen de modelos y teorías. Para ello, se ha analizado el grado de satisfacción de los profesionales de la traducción como recurso de información terminológica.

Un total de 122 especialistas en inglés, francés o italiano ha utilizado y evaluado el recurso terminológico desde la perspectiva de la traducción.

Tras el análisis de los resultados obtenidos respecto al SBR de HONqa los cuales han permitido evaluar la satisfacción de los alumnos de la Facultad de Traducción e interpretación de Granada se ha comprobado que el sistema les resulta a los usuarios-alumnos fácil y útil para la recuperación de información terminológica en todos los idiomas.

Palabras clave: Información multilingüe, Sistemas QA multilingües, Sistemas QA de dominio especializado, HONqa, Evaluación de la satisfacción de los usuarios, Información terminológica.

1. Introducción

La aparición y el consiguiente desarrollo de la Web han concedido al público general la capacidad de acceder a grandes volúmenes de información, lo que supone indudables beneficios. No obstante, esto también conlleva inconvenientes como la sobrecarga de información –que en este entorno es aún más acusada–, o el hecho de que buena parte de dicha información sea incorrecta, incompleta o imprecisa –ya sea de forma intencionada o no–. En consecuencia, se hace imprescindible avanzar en el desarrollo de herramientas y procedimientos que permitan al usuario obtener información fiable que sea relevante para su consulta en particular. Este es el reto al que se enfrenta la Recuperación de Información (en adelante, RI).

En los últimos años, algunos de los esfuerzos en la mejora de la RI en el entorno de la Web se han centrado en el diseño y desarrollo de los llamados sistemas de búsqueda de respuestas (en adelante SBR). El desarrollo de los SBR toma un importante impulso en el seno de la conferencia sobre recuperación de información TREC (Text REtrieval Conference¹) –principalmente a partir de TREC-8 (Vorhees, 1999)– la cual, desde 1992, constituye un importante foro internacional para aunar e incentivar la investigación en diferentes ámbitos de la recuperación de información.

En este trabajo se ha evaluado la satisfacción de los usuarios como información terminológica en el ámbito de los SBR de dominio restringido. Para ello, se ha utilizado la herramienta desarrollada por Ong et al. (2009) para la evaluación del grado de satisfacción de los usuarios del SBR multilingüe HONqa. A continuación se describen brevemente las principales características de los SBR, se detalla la metodología utilizada y se presentan los principales resultados obtenidos.

¹ TREC (Text REtrieval Conference), <http://trec.nist.gov/>

2. Sistemas de búsqueda de respuestas: Más allá de la Recuperación de Información

En el entorno de la Web la sobrecarga de información se deja sentir aún más que en otros contextos. De esta forma, en demasiadas ocasiones, al plantear una determinada consulta en las herramientas de búsqueda de información Web (buscadores, directorios o metabuscadores) el número de páginas Web recuperadas resulta excesivo y no todas ellas son relevantes ni útiles para los objetivos del usuario. Por ello, los profesionales de diversos ámbitos comienzan a reconocer la utilidad de otros tipos de sistemas, como los SBR, como método para la obtención de información especializada de forma rápida y efectiva (Crouch et al., 2005; Lee et al., 2006).

La importancia de la aparición de los SBR, en el ámbito de la información digitalizada, es actualmente innegable. Se presentan como una alternativa a los tradicionales sistemas de RI tratando de ofrecer respuestas precisas y comprensibles a preguntas factuales, en lugar de presentar al usuario una lista de documentos relacionados con la búsqueda (Jackson y Schilder, 2005), de modo que el usuario no ha de leer documentos completos para obtener la información requerida. La tarea de los SBR se puede definir como la acción de encontrar respuestas concretas a preguntas precisas y arbitrarias formuladas por los usuarios en lenguaje natural dentro de grandes volúmenes de información (como por ejemplo la World Wide Web). A pesar del avance que supone el poder contar con herramientas de búsqueda de información de este tipo, los SBR presentan algunas restricciones.

El funcionamiento de los SBR se basa en los modelos de respuestas cortas (Blair-Goldensohn, 2004), ya que la pregunta es dividida en palabras claves a las que se le asigna una etiqueta que permite distinguir y recuperar la información deseada. El sistema reemplaza esa etiqueta por las palabras adecuadas para poner a disposición de los usuarios una selección de textos que responden correctamente a la consulta (Pérez-Coutiño et al., 2004). La ventaja principal es que el usuario no ha de leer documentos completos para obtener la información requerida puesto que el sistema ofrece la respuesta correcta en forma de un número, un sustantivo, una frase corta o un fragmento breve de texto. Planteada la pregunta en el motor de búsqueda del sistema, se procede a analizar la pregunta separando la palabra o palabras claves, luego se localiza y extrae una respuesta a partir de diferentes fuentes dependiendo de la especialización del sistema se utilizarán unas u otras fuentes (Olvera-Lobo y Gutiérrez-Artacho, 2010)–, y finalmente, se evalúa y elimina aquella información redundante o que no responde correctamente a la pregunta planteada para, posteriormente, elaborar y presentar una o varias respuestas concretas que supuestamente satisfacen la necesidad del usuario (Cui et al. 2004; Tsur 2003).

En cuanto a los tipos de la lengua hay un gran número de clasificaciones propuestas (Adiwibowo y Adriana, 2007; Izquierdo et al., 2007; Roger et al., 2007; Solorio et al., 2005) que permite la distinción entre sistemas monolingües y multilingües, aunque en algunos casos un tercer tipo podrían ser incorporados relacionados con los SBR en varios idiomas que utilizan el inglés como lengua fundamental (García-Cumbreras et al., 2006). Por otro lado, la cobertura temática de los documentos en la base de datos es también un aspecto a tener en cuenta (Harabagiu et al, 2000; Magnini et al, 2001; Moldovan et al, 2003.; Roger et al., 2007). Así, mientras que unos pocos son de dominio abierto, es decir, los SBR permiten a los usuarios realizar preguntas sobre temas de carácter general o

multidisciplinarios, sin restricciones, otros tienen una colección especializada de un determinado campo temático.

En un entorno multilingüe como el de la Web la mayoría de los sistemas de RI tienen la limitación de encontrar documentos sólo en el idioma en el que se escribe la consulta o bien incorporan sistemas de traducción automática que únicamente resultan útiles cuando los documentos ya han sido localizados y tampoco siempre, pero no facilitan un medio efectivo para salvar la barrera del idioma en el proceso de búsqueda. Dentro de los SBR, hay que mencionar los vinculados a la RI multilingüe, los cuáles implican, al menos, dos lenguas, y permiten plantear las consultas en varios idiomas y recuperar información en todas las lenguas aceptadas por el sistema (Diekema, 2003).

Estos sistemas son capaces de operar en una colección de documentos multilingües dada una consulta determinada, y recuperar aquella información relevante que responda a la misma, independientemente del idioma utilizado al plantear la consulta (Grefenstette, 1998). Dentro de la RI multilingüe, se debe destacar el caso especial del objeto del estudio, los sistemas multilingües de SBR, que son aquellos sistemas donde el idioma en el que se plantea la pregunta, puede ser diferente a la lengua en la que está escrito el documento recuperado, pero se diferencian del resto de sistemas de la RI multilingües en que éstos no recuperan documentos completos sino que responden con una respuesta corta a la consulta planteada. Normalmente el funcionamiento de estos sistemas es muy similar a los de SBR monolingües, solamente se incorpora el módulo de la traducción y/o la herramienta o recurso lingüístico que llevará a cabo la recuperación translingüe.

3. Interactividad de los Sistemas de búsqueda de respuestas

La capacidad de interacción con el usuario es un aspecto muy importante de los SBR. Aunque es notable el interés creciente por los SBR capaces de mantener un diálogo continuado con los usuarios, la mayor parte de los sistemas aún no alcanzan un grado de comunicación conversacional. La ventaja de los sistemas plenamente interactivos es que utilizan cada nuevo input para conseguir cierta retroalimentación y afinar en la búsqueda de la respuesta más adecuada. Por ejemplo, están más preparados para resolver con éxito las preguntas enlazadas, que incluyen referencias a preguntas o a respuestas anteriores (¿quién descubrió la penicilina?, ¿dónde nació?). Los sistemas interactivos de búsqueda de respuestas encajan perfectamente con el enfoque de los tradicionales servicios de referencia. Las aplicaciones en este ámbito (Pomerantz, 2005) parten de la premisa de que la consulta primera del usuario suele ser errónea, ya que éste puede tener cierta dificultad para expresar su necesidad de información. Se cuestiona la validez de la pregunta inicial y se busca una reformulación más exacta ofreciendo mayor orientación durante la consulta. Por el contrario, los sistemas de búsqueda de respuestas que no cuentan con la perspectiva del servicio de referencia, es decir, la inmensa mayoría, consideran que el usuario formula correctamente su pregunta.

A la par de los SBR, se han desarrollado medidas para evaluar el desempeño de estos sistemas. Estas medidas se han propuesto desde diferentes perspectivas, tales como: la utilización de colecciones de prueba (Voorhees & Tice, 2000), el uso de pruebas de lectura y comprensión de textos (Charniak et al., 2000) y la aplicación de sistemas automáticos que

evalúan la validez de las respuestas dadas por los sistemas, esto es mediante su comparación con las respuestas generadas por los humanos a las mismas preguntas (Breck et al., 2000). La perspectiva de mayor éxito ha sido la utilización de colecciones de prueba, la cual consiste de un conjunto de documentos, un conjunto de preguntas y respuestas y una medida del rendimiento del sistema (Vicedo et al., 2003).

3.1 Evaluación centrada en el usuario

Partimos de un modelo de evaluación de SBR a partir de modelos de sistemas de satisfacción del usuario y aceptación de la tecnología. El concepto fundamental fue inspirado por la Teoría de la Acción Razonada (TRA) (Fishbein & Ajzen, 1980), elaborado de la psicología social, que se ha utilizado para predecir una amplia gama de los comportamientos ya que se plantea la hipótesis de que las actitudes están determinadas por las características que los observadores asocian con un objeto se basan en las características del contexto y de la muestra. También se ha basado en la teoría del comportamiento planificado (TPB) (Ajzen, 1991) que ayuda a entender cómo podemos cambiar el comportamiento de la gente.

Se adoptan tres dimensiones de la calidad (DeLone & McLean, 2003) de los modelos de éxito de los sistemas de información: calidad de la información, calidad de los sistemas y calidad del servicio. La satisfacción también tiene tres dimensiones: satisfacción de la información, satisfacción de los sistemas y satisfacción del servicio.

Existen creencias acerca de que la calidad afecta a la satisfacción. En el que la satisfacción de la información y la satisfacción del sistema se cree que influye en la utilidad y la facilidad de uso percibida. Muchos investigadores consideran que la intención de un individuo para utilizar un sistema se encuentra correlacionada significativamente con su uso real que a su vez es un indicador de la aceptación de un sistema de información. En resumen la satisfacción del usuario puede ser considerada como una función de la facilidad y utilidad de uso percibida y es probable que conduzca a una aceptación y éxito. El éxito de la evaluación de un SI y la satisfacción del usuario son temas importantes en el campo de la gestión de la información, especialmente en el caso de los sistemas de gestión on line. Básicamente, se utilizan modelos de evaluación para comprender las necesidades de los usuarios e identificar importantes dimensiones y factores que intervienen en el desarrollo de sistemas con el fin de ampliar su aceptación.

Con el rápido crecimiento en los últimos años, los SBR han surgido importantes aplicaciones. Por ello, han recibido la atención de los investigadores, en particular la de los de recuperación de información y procesamiento del lenguaje natural. Hay SBR que están enfocados a la interacción con el usuario, permitiendo que el proceso de búsqueda vaya refinándose con la intervención del usuario. Este tipo de sistemas tienen en cuenta a los usuarios y sus interacciones con las preguntas y los documentos donde se realiza la búsqueda (Hersh, 2006). El principal objetivo de la recuperación de la información de los SBR es satisfacer las necesidades de sus usuarios. El primer intento de definir “la satisfacción del usuario” como concepto, Tessier et al. (1977) afirmó que la satisfacción es “En última instancia una vivencia en el interior de la cabeza del usuario” y por lo tanto fue una respuesta “que puede ser a su vez emocional e intelectual”. La evaluación centrada en

el usuario se utiliza para entender las necesidades del usuario e identificar importantes dimensiones y factores en el desarrollo de un sistema de información con el fin de mejorar su aceptación.

4. Metodología

En este trabajo se ha llevado a cabo un estudio para evaluar la opinión de los usuarios acerca del uso de una herramienta que puede ser útil como fuente de información de traducción terminológica en un sistema biomédico. Con este fin, se ha analizado y evaluado el funcionamiento del SBR, de dominio especializado en el dominio biomédico HONqa. HONqa está desarrollado por la *Health On the Net Foundation (Fundación de la salud en la red)*, una organización suiza sin ánimo de lucro. El usuario accede directamente a la información requerida y no tiene que pasar por todos los documentos propuestos por el motor de búsqueda, una de las características principales de este sistema es que permite al usuario evaluar positiva o negativamente la respuesta recuperada, ofreciendo esa información a los desarrolladores (véase figura 1). Además, para cada pregunta planteada al sistema por el usuario se indica de qué tipo es (definición, factual o de lista), el número total de respuestas recuperadas, la lengua en la que están, así como sobre qué aspecto médico versa (tratamiento, síntomas, diagnóstico...). Los objetivos del sistema son los de ofrecer respuestas pertinentes a una pregunta en lenguaje natural, promover la creación de información médica de calidad y veraz y facilitar el acceso a los datos médicos más recientes e importantes a través del uso de Internet.

Answers:	Language:	Expected question type:	Expected medical type:	Rate answers:
26	English	definition	disease	Is it appropriate or not ?
1: Answer from familydoctor.org				
Attention-deficit hyperactivity disorder (ADHD) is the name of a group of behaviors found in many children and adults.				✓ ✗
2: Answer from www.adhd.org.nz				
attention deficit/hyperactivity disorder: combined type attention deficit/hyperactivity disorder: predominantly inattentive attention deficit/hyperactivity disorder: predominantly hyperactive or impulsive				✓ ✗
3: Answer from www.webmd.com				
Attention-deficit/hyperactivity disorder (ADHD) is also known as hyperactivity or attention deficit disorder (ADD).				✓ ✗
4: Answer from www.nimh.nih.gov				
Attention deficit/hyperactivity disorder across the lifespan.				✓ ✗

Figura 1: Página de resultados de sistema de búsqueda de respuestas HONqa

El modelo de evaluación propuesto proporciona un marco de trabajo para el diseño de SBR desde la perspectiva del usuario y que podría ayudar a incrementar la aceptación del usuario de los SBR. En nuestro estudio partimos de la herramienta propuesta por los investigadores Ong et al. (2009), en su trabajo *“The measurement of user satisfaction with question answering systems”* que permite medir la satisfacción de los usuarios. La herramienta propuesta ha sido probada de una forma rigurosa proporcionando un alto grado

de confianza en la fiabilidad y la validez de las escalas, los resultados confirmaron la existencia de cuatro factores con valores propios superiores a 1, que en conjunto representaron el 78,5% de la varianza total, confirmando la carga significativa de todos los elementos en un solo factor indicado de forma unidimensional, esto junto al hecho de que no haya carga cruzada de dichos ítems, apoyan la validez discriminante del instrumento. La fiabilidad se determinó mediante el uso de alfa de Cronbach para evaluar la consistencia interna de los elementos que representan a cada factor. El instrumento de 18 ítems (véase tabla 1) tenían una alta fiabilidad (0,92) superando el estándar mínimo sugerido para la investigación básica. La escala utilizada para medir el grado de satisfacción es una escala Likert de siete puntos que va desde “Totalmente en desacuerdo” a “Muy de acuerdo”. Esta medida supone que los usuarios no estarán satisfechos con un sistema si este no proporciona información de forma satisfactoria. La calidad técnica de un sistema es irrelevante, ya que un sistema no se consideraría exitoso si no satisface las necesidades de los usuarios. Esto explica por qué la calidad del sistema no juega un papel en la medición de la satisfacción del usuario.

Para ello se ha utilizado un cuestionario que se estructura en tres partes, una primera donde se presenta los datos demográficos (edad, sexo y curso), una segunda parte que es la autopercepción de los usuarios respecto al grado de experiencia con herramientas de recuperación y al grado de conocimientos médicos, que se evalúa con una escala de tres puntos (alto=3, medio=2, bajo=1). Y por último, una tercera parte que consta de 17 preguntas en donde se recoge la herramienta diseñada para medir la satisfacción de los usuarios en los SBR (Ong et al., 2010). Estas preguntas han sido evaluadas por una escala de Likert de 7 puntos (0= Totalmente en desacuerdo, 1= Muy en desacuerdo, 2 = En desacuerdo en ciertos aspectos, 3 = Indeciso, 4 = De acuerdo en ciertos aspectos, 5 = Muy de acuerdo, 6= Totalmente de acuerdo).

CODIGO	ORIGINAL	DESCRIPCIÓN
E1	Q27	<i>La interacción con el sistema de búsqueda de respuestas es clara y comprensible</i>
E2	Q22	<i>Resulta sencillo aprender a utilizar el sistema</i>
E3	Q25	<i>Es fácil llegar a estar completamente familiarizado con el uso del sistema</i>
E4	Q24	<i>Encuentro fácil utilizar el sistema para que realice lo que yo quiero que haga</i>
E5	Q26	<i>Encuentro fácil utilizar el sistema</i>
U1	Q20	<i>Utilizar el sistema podría mejorar mi eficacia en el desarrollo de mi trabajo</i>
U2	Q18	<i>El sistema podría ser útil en el desarrollo de mi trabajo</i>
U3	Q16	<i>El uso del sistema podría mejorar el rendimiento en mi trabajo</i>
U4	Q19	<i>El uso del sistema en mi trabajo podría incrementar mi productividad</i>
U5	Q21	<i>El uso del sistema podría hacer más fácil mi trabajo</i>
S1	Q12	<i>Este sistema de búsqueda de respuestas es fiable</i>
S2	Q13	<i>Los técnicos del sistema de búsqueda de respuestas ofrecen un servicio rápido a los usuarios</i>

S3	Q11	<i>La tecnología utilizada por el sistema y la información que ofrece están actualizados</i>
S4	Q14	<i>Los técnicos de los sistemas de búsquedas de respuestas están bien formados para hacer bien su trabajo</i>
I1	Q2	<i>La información facilitada por el sistema es fácil de comprender</i>
I2	Q4	<i>La información facilitada por el sistema es relevante</i>
I3	Q1	<i>La información facilitada por el sistema es completa</i>
I4	Q3	<i>La información facilitada por este sistema está personalizada</i>

E =Fácil de usar **U** =Utilidad **S** =Calidad del Servicio **I** =Calidad de la Información

Tabla 2: *Herramienta de trabajo para medir la satisfacción de los usuarios*

Estas 17 preguntas se agrupan en cuatro factores: un primer factor, la facilidad de uso del sistema (la cinco primeras preguntas), un segundo factor, la utilidad del sistema (de la sexta a la décima), un tercero, la calidad del servicio (de la undécima a la décima tercera) y un cuarto, la calidad de la información (de la décimo cuarta a la décimo séptima) (véase tabla 2). Los cuestionarios se han realizado en la Facultad de Traducción e Interpretación de la Universidad de Granada con alumnos de 1^a, 2^o, 3^o y 4^o curso de la especialidades de Inglés, Francés e Italiano, la elección de estas tres lenguas es porque coinciden con las del sistema multilingüe HONqa.

La interacción con el sistema de búsqueda de respuestas es clara y comprensible: **Facilidad de uso**
 Resulta sencillo aprender a utilizar el sistema: **Facilidad de uso**
 Es fácil llegar a estar completamente familiarizado con el uso del sistema: **Facilidad de uso**
 Encuentro fácil utilizar el sistema para que realice lo que yo quiero que haga: **Facilidad de uso**
 Encuentro fácil utilizar el sistema **Facilidad de uso**
 Utilizar el sistema podría mejorar mi eficacia en el desarrollo de las tareas académicas: **Utilidad**
 El sistema podría ser útil en el desarrollo de mis tareas: **Utilidad**
 El uso del sistema podría mejorar el rendimiento en mi trabajo: **Utilidad**
 El uso del sistema en mis tareas podría incrementar mi productividad: **Utilidad**
 El uso del sistema podría hacer más fácil mi trabajo: **Utilidad**
 Este sistema de búsqueda de respuestas es fiable: **Calidad del servicio**
 La ayuda ofrecida por el sistema es suficiente: **Calidad del servicio**
 La tecnología utilizada por el sistema y la información que ofrece están actualizados: **Calidad del servicio**
 La información facilitada por el sistema es fácil de comprender: **Calidad de la información**
 La información facilitada por el sistema es relevante: **Calidad de la información**
 La información facilitada por el sistema es completa: **Calidad de la información**
 La información facilitada por este sistema está personalizada: **Calidad de la información**

Tabla 2: *Tabla de ítems del cuestionario agrupada en cuatro factores*

Los datos se recogieron de un total de 122 alumnos (58 alumnos de inglés, 32 alumnos de francés y 32 alumnos de italiano). De las 122 encuestados 27 eran hombres, 95 mujeres. Todos los encuestados no tenían experiencia previa en el uso de los SBR. Antes de realizar los cuestionarios se les dio una breve información sobre los SBR y de forma más detallada sobre el funcionamiento del SBR HONqa, además se les dejó unos días para que interactuaran con el sistema realizándoles preguntas facilitadas (en inglés, francés e italiano) y se les propuso que realizarán su propias preguntas. Cada grupo interactuó con las

preguntas en un único idioma, según su especialidad. Todas las preguntas pertenecen al grupo de definición, con el objetivo de poder valorar la utilidad de sistema como fuente de información terminológica para este tipo de usuarios, y con la particularidad de obtener respuestas específicas.

5. Resultados y discusión

Los resultados se presentan en los siguientes cuatro apartados: datos demográficos, autopercepción de los usuarios respecto al grado de experiencia con herramientas de recuperación y al grado de conocimientos médicos, satisfacción de los usuarios respecto al sistema según la herramienta de evaluación aplicada, y correlaciones entre el grado de satisfacción en relación a la facilidad de uso, utilidad, calidad del servicio y calidad de la información.

5.1 Datos demográficos

Los datos recogidos de un total de 122 alumnos se dividen en las siguientes lenguas: 58 alumnos de inglés, 32 alumnos de francés y 32 alumnos de italiano. En los 122 encuestados las edades oscilan entre los 17 y los 39 años obteniendo una media de edad de 21,48. Aunque existe cierta variación en la edad de los alumnos según la lengua, la media de edad en los tres idiomas es prácticamente igual (20,50) en francés, (21,83) en inglés y (21,84) en italiano (véase tabla 3).

Estadísticos

Edad		
N	Válidos	122
	Perdidos	0
Media		21,48
Mediana		21,00
Moda		20
Mínimo		17
Máximo		39

Tabla 3. Estadísticas de la edad

De los 122 alumnos 27 son hombres y 95 son mujeres. Esta diferencia entre el número mujeres y hombres se ve reflejada en el número de alumnos y alumnas del Grado y Licenciatura en Traducción e Interpretación, siendo casi el triple el número de alumnas que el de alumnos. La mayoría de los estudiantes son de tercero (89 alumnos). Le siguen en número los alumnos de cuarto (20) y primero (12) (véase tabla 4). La mayoría de los alumnos son de la especialidad de inglés (58) le sigue a continuación los alumnos de francés e italiano ambos con igual número de alumnos (32).

Curso

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Primero	12	9,8	9,8	9,8
	Segundo	1	,8	,8	10,7
	Tercero	89	73,0	73,0	83,6
	Cuarto	20	16,4	16,4	100,0
	Total	122	100,0	100,0	

Tabla 4. Estadísticas de los alumnos según curso

5.2 La autopercepción de los usuarios respecto al grado de experiencia con herramientas de recuperación y al grado de conocimientos médicos.

El grado de experiencia global (1,75) corresponde según la escala de 3 puntos utilizada (bajo, medio, alto) a un valor bajo. El grupo de italiano (2,07) supera esta media obteniendo un nivel medio en cuanto al grado de experiencia en el uso de herramientas de recuperación, le sigue el grupo de francés (1,73) y en último lugar el de inglés (1,60). La media global (1,22) obtenida en el grado de conocimientos médicos no es significativa, los resultados son muy similares en los tres grupos obteniéndose un nivel de conocimientos médicos bajo (Figura 2).

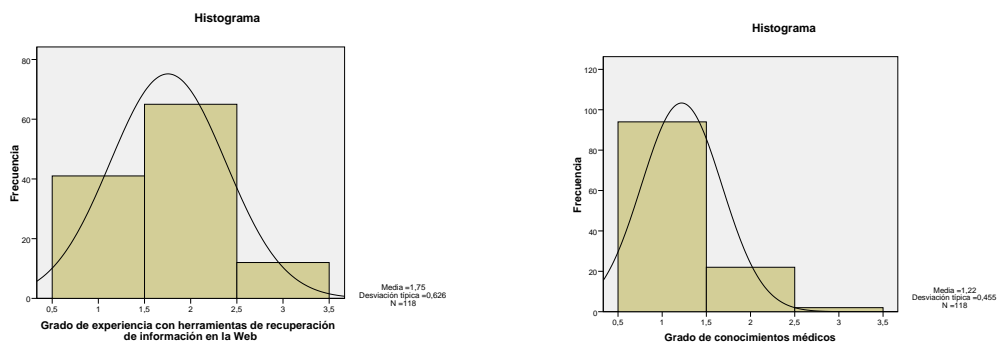


Figura 2. Histograma del grado de conocimiento de los estudiantes

5.3 La satisfacción de los usuarios respecto al sistema, según la herramienta de evaluación aplicada.

El grado de satisfacción de los alumnos de la Facultad de Traducción e interpretación que tienen acerca del SBR HONqa se evaluó en el cuestionario en el cual se encontraban dos diferentes escalas una escala de tres puntos y una escala de Likert de 7 puntos los resultados se explicaran a continuación a través de la descripción de las diferentes relaciones resultantes del análisis entre los tres grupos de las 3 especialidades de idioma (inglés, francés e italiano) realizado con la ayuda del programa SPSS (*Statistical Program For Social Sciences*) A la pregunta realizada, “la interacción con el sistema de búsqueda de respuestas es clara y comprensible” la media global obtenida es de 4,04 mostrándose de acuerdo en ciertos aspectos. El grupo de inglés (4,59) al igual que el de francés (4,38) muestran el mismo grado de conformidad en ciertos aspectos y el grupo de italiano se muestran indecisos (2,70). Los resultados obtenidos a la pregunta, “resulta sencillo aprender a utilizar el sistema” la media global obtenida (4,88) de los tres grupos reflejan que se encuentran de acuerdo en ciertos aspectos. A los alumnos del grupo de inglés le resulta sencillo el aprendizaje (5,34), en cambio al grupo de francés (4,75) y al de italiano no les resulta sencillo en su totalidad (4,16).

El resultado obtenido a la pregunta es “fácil llegar a estar completamente familiarizado con el uso del sistema”, la media global (4,49) obtenida por los tres grupos indican acuerdo en ciertos aspectos. Los del grupo de inglés dan una puntuación mayor (4,95) les siguen los del grupo de francés (4,23) y por último los del grupo de italiano (3,90) que se muestran indecisos.

A la pregunta “encuentro fácil utilizar el sistema para que realice lo que yo quiero que haga” la media global obtenida en (3,47) los tres grupos muestra una clara indecisión. El grupo de inglés (3,98) da una puntuación superior al de francés y al grupo de italiano (2,69) que se muestra en desacuerdo en ciertos aspectos. A la pregunta “encuentro fácil utilizar el sistema, la media global del grupo (4,61) nos indica que están casi de acuerdo en este aspecto. Por el contrario el grupo de inglés (5,00) está muy de acuerdo con la facilidad del sistema, mientras que el grupo de francés (4,69) opina al igual que la media grupal y el grupo de italiano (3,84) se encuentra indeciso. A la pregunta “utilizar el sistema podría mejorar mi eficacia en el desarrollo de las tareas académicas” la media global grupal (3,70) indica indecisión. El grupo de inglés (4,14) se muestra más de acuerdo en ciertos aspectos, mientras que el grupo de francés (3,43) y el de italiano (3,13) se encuentran indecisos. A la pregunta “el sistema podría ser útil en el desarrollo de mis tareas” la media global (3,91) grupal muestran indecisión. El promedio entre las tres lenguas no es significativo, no hay diferencias entre los tres grupos. A la pregunta “el uso del sistema podría mejorar el rendimiento en mi trabajo” la media global (3,63) grupal indica vacilación. El grupo de inglés (4,12) muestra algo más de seguridad, mientras que el grupo de francés y el de italiano vuelven a mostrar esa indecisión. A la pregunta “el uso del sistema en mis tareas podría incrementar mi productividad” la media grupal (3,58) sigue mostrando indecisión. Los tres grupos por separado muestran el mismo grado de satisfacción puntuando mejor el grupo de inglés (3,98). A la pregunta “el uso del sistema podría hacer más fácil mi trabajo” la media global (3,77) continua con esa indecisión ante las preguntas relacionadas con el sistema. El grupo de inglés (4,29) sigue estando entre de acuerdo e indeciso. El grupo de

francés (3,56) y el de italiano (3,06) comparten el mismo grado de indecisión. A la pregunta “este sistema de búsqueda de respuestas es fiable” la media global (3,68) grupal indica un desconocimiento sobre la fiabilidad del sistema. El grupo de inglés (4,25) muestra un mayor grado de confianza en la fiabilidad mientras que esta confianza va decayendo en el de francés (3,55) e italiano (2,81).

A la pregunta “la ayuda ofrecida por el sistema es suficiente” se obtienen los segundos peores resultados en el cuestionario con una media global (2,99) muy baja mostrando su desacuerdo la ayuda ofrecida por el programa. El promedio entre los tres grupos no es significativo no hay diferencias entre las tres lenguas. A la pregunta “la tecnología utilizada por el sistema y la información que ofrece están actualizados” la media global (3,54) indican desconocimiento sobre este aspecto. El promedio de los tres grupos no es significativo, expresan el mismo grado de satisfacción de indecisión. A la pregunta “La información facilitada por el sistema es fácil de comprender” la media grupal (4,04) indica mejor valoración relativa, se muestran de acuerdo en ciertos aspectos. El grupo de inglés (4,58) tiene el mismo grado de satisfacción que la media, pero el grupo de francés (3,90) y el de italiano (3,22) se muestran indecisos. A la pregunta “la información facilitada por el sistema es relevante” la media grupal (3,66) indica desconocimiento de la materia. El grupo de inglés (4,29) muestra mayor conocimiento que los del el grupo de francés (3,34) y mucho más que los del grupo de italiano (2,88) que están en desacuerdo parcialmente. A la pregunta “la información facilitada por el sistema es completa” la media obtenida (3,22) grupal considera que no es del todo completa ni incompleta. El grupo de inglés (3,76) y el de francés (3,10) valoran con un grado de satisfacción igual a la media y el grupo de italianos (2,38) muestra su desacuerdo parcial. A la pregunta “la información facilitada por este sistema está personalizada” se obtiene la puntuación más baja estando la media global grupal (2,77) en desacuerdo. El grupo de inglés (3,05) se muestra indeciso, el grupo de francés (2,71) y el de italiano (2,34) muestran su desacuerdo. Los resultados obtenidos de los 17 ítems agrupados en 4 factores muestran valoraciones dispares, dando puntuaciones muy altas a la facilidad y utilidad del sistema y más bajas a la calidad del servicio y de la información prestada por el mismo. Estos resultados son similares tanto a nivel global como por lengua.

5.4 Correlaciones entre el grado de satisfacción en relación a la facilidad de uso, utilidad, calidad del servicio y calidad de la información.

Antes de comentar el grado de correlación entre los cuatro factores realizaré un breve comentario sobre la correlación de la variable edad y sexo en relación a las respuestas de los alumnos. No hay correlación significativa entre la variable edad y la respuesta de los alumnos. En cuanto a la variable sexo no hay grandes diferencias en la respuestas a excepción de las preguntas: “Utilizar el sistema podría mejorar mi eficacia en el desarrollo de las tareas académicas” ($P= 0,047$) y “Este sistema de búsqueda de respuestas es fiable” ($P= 0,038$) en el que la mujeres dan mejor puntuación. Los resultados utilizados para interpretación se encuentra contenidos en las tablas, de estos se tomaron en cuentan aquéllos cuya relación tuviera un valor superior o igual a 0,6 debido a que se considera que este valor implica una relación fuerte. A nivel global los cuatro factores nos dan como

resultado una correlación más alta entre la calidad de servicio y la calidad de la información (668) cuanto mayor sea el grado en que un alumno muestra su comprensión y está de acuerdo con la situación, capacidad de respuesta y tiempo de reacción del servicio, más satisfecho se mostrará con la calidad de los datos proporcionados por el sistema. Esta correlación también se refleja a nivel grupal de francés (763) e italianos (754).

6. Conclusiones

Tras el análisis de los resultados obtenidos respecto al SBR de HONqa los cuales han permitido evaluar la satisfacción de los alumnos de la Facultad de Traducción e interpretación de Granada se ha comprobado que el sistema les resulta a los usuarios-alumnos fácil y útil para la recuperación de información, pero se muestran indecisos al evaluar la calidad del servicio y la calidad de la información proporcionada. Esto es debido a que el interfaz de sistema es simple y fácil de manejar y su utilidad se aprecia en seguida. Pero por otro lado no se ven capaces de evaluar la calidad de la información proporcionada puesto que ellos muestran un nivel bajo de conocimientos médicos. Otro aspecto interesante es que si comparamos los resultados obtenidos en los tres grupos de lengua, los del grupo de inglés en general se muestran más satisfechos con el sistema ya que han dado a los 17 ítems los valores más altos. A continuación le sigue los del grupo francés y por último los de italiano. En definitiva desde el punto de vista del usuario, la satisfacción de sus necesidades de información, es la medida, de eficacia del sistema por excelencia. La opinión de los usuarios sobre la satisfacción o insatisfacción, deriva directamente de la calidad de un servicio, de sus características y de la información que le proporciona. Constituyendo una información muy importante acerca del éxito o fracaso del sistema para cubrir sus expectativas. No es suficiente con que el sistema funcione, es necesario que el servicio que se presta sea satisfactorio para el usuario y que éste así lo perciba.

Referencias bibliográficas

- Adiwibowo, S.; Adriani, M (2007). Finding answers using resources in the internet. En: Working notes for the CLEF 2007 Workshop .Disponible en: http://www.clef-campaign.org/2007/working_notes/adiwibowoCLEF2007.pdf
- Ajzen, I. (1991). The theory of planned behavior. En: *Organizational and Human Decision Processes*, Vol. 50, 179-211.
- Blair-Goldensohn, S., McKeown, K., & Schlaikjer, A. H. (2004). Answering Definitional Questions: A Hybrid Approach, *New Directions in Question Answering*, 4, 47-58.
- Breck E. et al. (2000). How to Evaluate Your Question Answering System Every Day ... and Still Get Real Work Done, *Proceedings of Second International Conference on Language Resources and Evaluation*. Atenas: Grecia.
- Cui, Hang; Kan, Min-Yen; Chua, Tat-Seng y Jing Xiao (2004). A Comparative Study on Sentence Retrieval for Definitional Question Answering. *SIGIR Workshop on Information retrieval for Question Answering (IR4QA)*, Sheffield.
- Crouch, D., Saurí, R. y Fowler, A. (2005). *AQUAINT Pilot Knowledge-Based Evaluation: Annotation Guidelines*. Palo Alto Research Center. Disponible en:

- http://www2.parc.com/isl/groups/nltp/papers/aquaint_kb_pilot_evaluation_guide.pdf[consulta 19/11/2011]
- Charniak E. et al. (2000, mayo). Reading Comprehension Programs in a Statistical-Language-Processing Class. ANLP/NAACL Workshop on Reading Comprehension Tests as Evaluation for Computer-Based Language Understanding Systems. Washington: Seattle.
- DeLone, W.H.; E.R. McLean (2003). DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, (19:4), 9-30
- Diekema, Anne R. (2003). Translation Events in Cross-Language Information Retrieval: Lexical ambiguity, lexical holes, vocabulary mismatch, and correct translations. Tesis Doctoral. University of Syracuse.
- García-Cumbreras, M.A., Ureña, L.A., & Martínez-Santiago, F. (2006, april). BRUJA: Question Classification for Spanish. Using Machine Translation and an English Classifier. Workshop on Multilingual Question Answering.
- Grefenstette, Gregory (1998). Cross-Language Information Retrieval. Kluwe academic publishers 1.
- Harabagiu, S. M.; Pascal, M. A.; Maiorano, S. J (2000). Experiments with open domain textual question answering. En: Proceedings of the Coling. Association for Computational Linguistics . Harman, D. K. (1998). Text retrieval conferences (TREC): providing a test-bed for information retrieval systems. *Bulletin of the American Society for Information Science*, vol. 24 (4), 11-13.
- I Izquierdo, R., Fernández, O., Fernández, S., Tomás, D., Vicedo, J.L., Martínez, P., & Suarez, A. (2007). Question Answering Learning technologies in a multilingual and multimodel Environment, 38, 43-47
- Jackson, Peter y Frank Schilder (2005). Natural Language Processing: Overview. En K. Brown :. *Encyclopedia of Language & Linguistics* 2, 503-518. Amsterdam: Elsevier Press.
- Jones, K. Sparck, & Jackson, David M. (1970). The use of automatically-obtained keyword classifications for information retrieval. *Information Storage and retrieval*, 5(4), 175-201.
- Magnini, B.; Negri, M.; Prevete, R. (2001) "Open domain question/answering on the Web". En: Congress of the Italian Association for Artificial Intelligence, n. 2175 (pp. 273-284).
- Moldovan, D. et al. (2003). Performance issues and error analysis in an open-domain question answering system. *ACM transactions on information systems*. Esta indicado: v. 21, n. 2, 133 -154. 67
- Olvera-Lobo, M. D. & Gutiérrez-Artacho, J. (2010). Question-answering systems as efficient sources of terminological information: Evaluation. *Health Information and Library Journal*, 27 (4): 268-274
- Ong Cc, Day M, Hsu M. (2009). The measurement of user satisfaction with question answering systems. *Information & Management*, 46(7): 397-403
- Pérez-Coutiño, M.; Solorio, T.; Montes-Gomez, M.; López López, A.; Villaseñor-Pineda, L. "The use of lexical context in question answering for Spanish". En: Workshop of the cross language evaluation forum (CLEF2004). Bath, UK. September 2004.
- http://www.clefcampaign.org/2004/working_notes/WorkingNotes2004/46.pdf Pomerantz, J. (2005). "A linguistic analysis of question taxonomies". *Journal of the American Society for Information Science and Technology*, v. 56, n. 7, 715-728.
- R Roger, S., Fernández, A., Peral, J., Fernández, S., & P.López. (2007, march). An Inference Mechanism for Question Answering. In: *Journal of Computer Science and Technology*. Argentina: San Luis.

- Solorio, T.; Pérez-Coutiño, M.; Montes-Gómez, M.; Villaseñor-Pineda, L.; López-López, A. (2004). A language independent method for question classification. En: The 20th Intl. conf. on computer linguistics , 1374-1380
- Vicedo J. et al (2003). Los sistemas de Búsqueda de Respuestas desde una perspectiva actual. Revista de la Sociedad Española para el Procesamiento del Lenguaje Natural, n.31, 2003.
- Voorhees, E. M. (2002). Overview of the TREC 2002 Question Answering Track. Proceedings of the Eleventh Text REtrieval Conference.
http://comminfo.rutgers.edu/~muresan/IR/TREC/Proceedings/t11_proceedings/t11_proceedings.html [5 noviembre 2011] 69
- Voorhees, E. M. y Tice, D. (1999). The TREC-8 question answering track evaluation. En Greece: Athens. .