
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 tecnologías de la información y la comunicación 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. El presente trabajo se centra en los SBR multilingües —que permiten al usuario acceder a información terminológica no disponible en su lengua— y, concretamente, en su evaluación desde la perspectiva del usuario —la cual hace posible aprehender sus necesidades e identificar las dimensiones y factores relevantes en el desarrollo de los sistemas con el fin de mejorar su aceptación—.

El objetivo del estudio es conocer el grado de satisfacción de los usuarios especializados en el ámbito de la traducción respecto a los SBR de dominio restringido como fuente de información terminológica. Para ello se ha llevado a cabo una evaluación centrada en el usuario del SBR multilingüe (inglés, francés e italiano) HONqa como recurso de información terminológica. Como herramienta de evaluación se ha aplicado un riguroso cuestionario, ya testado y validado (Ong et al., 2009), que surge tras una exhaustiva revisión de modelos y teorías relacionados con la aceptación y uso de la tecnología.

El análisis de los resultados pone de manifiesto que el sistema les resulta a los usuarios fácil y útil para la recuperación de información terminológica en todos los idiomas evaluados.

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 respecto a los SBR de dominio restringido como fuente de información terminológica. Para ello, se ha utilizado la herramienta desarrollada por Ong et al. (2009), a modo de cuestionario, 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.

2. SISTEMAS DE BÚSQUEDA DE RESPUESTAS: MÁS ALLÁ DE LA RECUPERACIÓN DE INFORMACIÓN

En demasiadas ocasiones, al plantear una determinada consulta en las herramientas de búsqueda de información web (buscadores, directorios o metabuscadores) el volumen de páginas web recuperadas resulta excesivo, y no todas ellas son relevantes ni útiles para los objetivos del usuario. 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).

El funcionamiento de los SBR se basa en los modelos de respuestas cortas (Blair-Goldensohn, 2004), ya que la pregunta es dividida en palabras clave a las que se les asigna una etiqueta que permite distinguir y recuperar la información deseada. 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 que extrae de muy diversas fuentes (Olvera-Lobo y Gutiérrez-Artacho, 2010). Finalmente, 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 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, en ocasiones, incorporan sistemas de traducción automática que únicamente resultan útiles cuando los documentos ya han sido localizados, pero no facilitan un medio efectivo para salvar la barrera del idioma en el proceso de búsqueda.

Los SBR translingües —inexistentes por el momento en el entorno de la Web— 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). Normalmente su funcionamiento es muy similar al de los de SBR monolingües pero se les incorpora un módulo de traducción y/o una herramienta o recurso lingüístico que llevará a cabo la recuperación translingüe. Por su parte, los SBR multilingües, tipo al que pertenece el sistema evaluado en la presente investigación, —también muy escasos en la Web— cuentan con varias bases de datos independientes en distintas lenguas. Admiten preguntas en diferentes idiomas y presentan la respuesta en la misma lengua en la que se formuló la consulta.

3. EVALUACIÓN DE LOS SISTEMAS DE BÚSQUEDA DE RESPUESTAS CENTRADA EN EL USUARIO

Para analizar el desempeño de los sistemas de búsqueda de respuestas se han desarrollado diferentes métodos. Estos encaran la evaluación desde diversas 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), o 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 aproximación de mayor éxito ha sido la utilización de colecciones de evaluación, la cual consta de un conjunto de documentos, un conjunto de preguntas y respuestas, y medidas del rendimiento del sistema (Vicedo et al., 2003).

Sin embargo, el principal objetivo de los sistemas de RI en general y de los SBR en particular, es satisfacer las necesidades de información de sus usuarios. En uno de los primeros intentos de definir la «satisfacción del usuario» se afirma que es «en última instancia una vivencia en el interior de la cabeza del usuario» y, por lo tanto, es una respuesta «que puede ser a su vez emocional e intelectual» (Tessier et al., 1977). La evaluación centrada en el usuario se utiliza para entender sus necesidades e identificar las dimensiones y factores relevantes en el desarrollo de los sistemas con el fin de mejorar su aceptación.

El modelo de evaluación centrado en el usuario de los SBR propuesto por Ong et al. (2009) surge a partir de modelos de satisfacción de usuario y de aceptación de la tecnología. El concepto fundamental fue inspirado por la Teoría de la Acción Razonada (TRA) (Fishbein & Ajzen, 1980), elaborado por la psicología social —que se ha utilizado para predecir una amplia gama de comportamientos— y por la Teoría del Comportamiento Planificado (TPB) (Ajzen, 1991) que ayuda a entender cómo se puede alterar el comportamiento humano.

Para articular el modelo se adoptan tres dimensiones de la calidad de los sistemas de información (DeLone & McLean, 2003), a saber, calidad de la información, calidad de los sistemas y calidad del servicio. La satisfacción también tiene tres dimensiones, satisfacción respecto a la información, satisfacción respecto a los sistemas y satisfacción respecto al servicio.

Por tanto, con gran frecuencia, se vincula la calidad con la satisfacción del usuario. Según esta perspectiva, la satisfacción respecto a la información y la satisfacción respecto al sistema se cree que influyen en la utilidad y la facilidad de uso de ese sistema percibidas por el usuario. Muchos investigadores consideran que la intención de un individuo al utilizar un sistema se encuentra correlacionado significativamente con su uso real —el cuál, a su vez, es un indicador de la aceptación del sistema por parte de los usuarios—. 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 de esa tecnología.

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 terminológica. Con este fin, se ha analizado y evaluado el funcionamiento del SBR de dominio especializado ámbito biomédico HONqa. HONqa está desarrollado por la organización suiza *Health On the Net Foundation* con el objetivo de ofrecer respuestas pertinentes a preguntas planteadas 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 de Internet.

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...).

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.				<input checked="" type="checkbox"/> <input type="checkbox"/>
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				<input checked="" type="checkbox"/> <input type="checkbox"/>
3: Answer from www.webmd.com Attention-deficit/hyperactivity disorder (ADHD) is also known as hyperactivity or attention deficit disorder (ADD).				<input checked="" type="checkbox"/> <input type="checkbox"/>
4: Answer from www.nimh.nih.gov Attention deficit/hyperactivity disorder across the lifespan.				<input checked="" type="checkbox"/> <input type="checkbox"/>

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

El cuestionario que se ha aplicado en este estudio se estructura en tres partes: una primera, donde se presentan los datos demográficos (edad, sexo y curso); una segunda parte, que refleja la autopercepción de los usuarios respecto al grado de experiencia con herramientas de recuperación y en relación al grado de conocimientos médicos —y que se evalúa en una escala de tres niveles (alto=3, medio=2, bajo=1)—; y, por último, una tercera parte que consta de 17 preguntas donde se pretende medir la satisfacción de los usuarios respecto a los SBR. En esta última parte del cuestionario se utilizó una escala Likert de siete niveles que va desde «totalmente en desacuerdo» (=1) a «muy de acuerdo» (=6).

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 =Facilidad de uso, U =Utilidad, S =Calidad del Servicio, I =Calidad de la Información
Qn= número de pregunta en nuestro cuestionario

Tabla 1. *Herramienta de evaluación para medir la satisfacción de los usuarios*

Estas 17 preguntas se agrupan en cuatro factores: un primer factor, la facilidad de uso del sistema (que recoge las cinco primeras preguntas), un segundo factor, la utilidad del sistema (de la sexta a la décima), el tercero, centrado en la calidad del servicio (de la

undécima a la décima tercera) y el cuarto factor, cuyas preguntas se refieren a la calidad de la información (de la décimo cuarta a la décimo séptima). Los cuestionarios se aplicaron a estudiantes de 1.º, 2.º, 3.º y 4.º curso de las especialidades de inglés, francés e italiano —lenguas en las que trabaja el sistema multilingüe HONqa— en la Facultad de Traducción e Interpretación de la Universidad de Granada.

Los datos se recogieron de un total de 122 estudiantes y se analizaron con la ayuda del programa SPSS (*Statistical Program For Social Sciences*). Antes de realizar los cuestionarios se les ofreció a los estudiantes una breve información sobre los SBR y, de forma más detallada, sobre el funcionamiento del SBR HONqa. Durante unos días interactuaron libremente con el sistema planteándoles preguntas que se les había facilitado (en inglés, francés o italiano, dependiendo de su lengua de trabajo), y otras de su elección. Cada grupo interactuó con las preguntas en un único idioma, según su especialidad. Las preguntas que se les sugirió fueron definicionales («¿Qué es...?») con el objetivo de poder valorar la utilidad de sistema como fuente de información terminológica para este tipo de usuarios, con la particularidad de poder obtener respuestas específicas.

5. RESULTADOS Y DISCUSIÓN

A continuación se muestran los principales resultados obtenidos.

5.1. Datos demográficos

De los 122 estudiantes (58 de inglés, 32 alumnos de francés y 32 de italiano), 27 eran hombres y 95 mujeres —en nuestra universidad el número de alumnas triplica al de alumnos en este grado— con edades comprendidas entre los 17 y los 39 años —si bien el 66,4% contaba con entre 20 y 22 años—. Ningún encuestado tenía experiencia previa en el uso de los SBR. La mayoría de los estudiantes eran de tercer curso (89 alumnos, 73%) o de cuarto (20, 16,4%) frente a los de primer (2, 9,8%) o segundo curso (1, 0,8%).

5.2. Autopercepción sobre experiencia y conocimientos

En la autovaloración de los usuarios respecto a su grado de experiencia con herramientas de recuperación de información en la Web y al grado de conocimientos médicos —dominio temático en el que se centra el sistema HONqa— los resultados muestran lo siguiente.

Como se observa en la figura 2 los estudiantes manifiestan un grado de experiencia respecto al uso de sistemas de recuperación no muy alto (1,75) según la escala de 3 puntos utilizada (bajo, medio, alto) en esta sección del cuestionario. El hecho de que este tipo de sistemas no esté aún muy extendido y que los estudiantes no los utilicen con frecuencia explica esta autopercepción. Asimismo, en cuanto a la autovaloración del grado de conocimientos médicos, como era de prever, y al no tratarse de especialistas en ese dominio temático, los usuarios declaran tener escasos conocimientos en este ámbito temático (1,22).

SATISFACCIÓN DE LOS USUARIOS EN LA BÚSQUEDA MULTILINGÜE

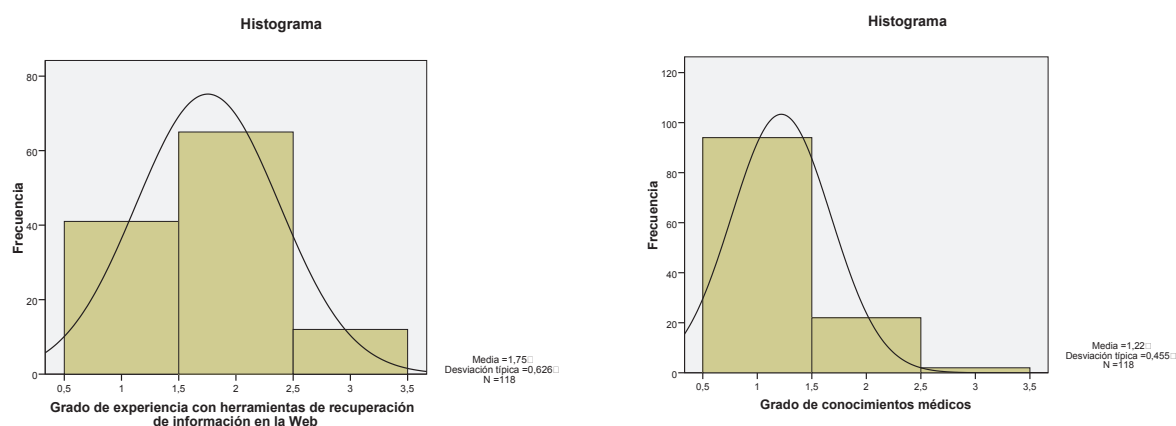


Figura 2. Histograma del grado de conocimiento de los estudiantes

5.3. Satisfacción de los usuarios respecto al sistema

Las tablas 2, 3 y 4 muestran los resultados agrupados en los cuatro factores en el que se organizan las preguntas, para los usuarios de inglés, de francés y de italiano, respectivamente.

	N	Media	Desv. típ.
Facilidad de uso	58	4,7678	,84276
Utilidad	58	4,1621	1,13413
Calidad del servicio	58	3,7615	1,05738
Calidad de la información	58	3,9488	1,00886

Tabla 2. Resultados agrupados por factores para usuarios de inglés

Estadísticos descriptivos

	N	Media	Desv. típ.
Facilidad de uso	32	4,2953	1,27882
Utilidad	32	3,4927	1,27936
Calidad del servicio	32	3,4032	1,48684
Calidad de la información	32	3,3091	1,41474

Tabla 3. Resultados agrupados por factores para usuarios de francés

Estadísticos descriptivos

	N	Media	Desv. típ.
Facilidad de uso	32	3,4406	1,15775
Utilidad	32	3,2078	1,43696
Calidad del servicio	32	2,8073	1,19670
Calidad de la información	32	2,7031	1,07283
N válido (según lista)	32		

Tabla 4. Resultados agrupados por factores para usuarios de italiano

Los resultados obtenidos de los 17 ítems agrupados en 4 factores muestran valoraciones dispares dependiendo del grupo de usuarios y del factor del que se trate. Mientras que la valoración de los usuarios es bastante aceptable respecto a la facilidad de uso (con valores entre 4,77 y 3,44) y a la utilidad del sistema (con valor máximo para el grupo de inglés en 4,16 y 3,2 para italiano), suelen ser más críticos al evaluar la calidad del servicio (3,76 y 2,81) y de la información ofrecida (3,95 y 2,7). Asimismo, para todos los factores, se observa que los usuarios del sistema en lengua inglesa muestran mejor opinión acerca del SBR que los de lengua francesa y, sobre todo, italiana.

Así ante a la pregunta, «resulta sencillo aprender a utilizar el sistema», que presenta un valor promedio de 4,88 puntos para todos los encuestados, los usuarios de inglés consideran muy sencillo su aprendizaje (5,34) mientras que al grupo de francés (4,75) y al de italiano (4,16) no les resulta tanto. Otro ejemplo, a la pregunta «la interacción con el sistema de búsqueda de respuestas es clara y comprensible» la media global obtenida es de 4,04 pero, mientras el grupo de inglés (4,59) y el de francés (4,38) muestran grado similar de acuerdo el grupo de italiano valora mucho peor este aspecto (2,70).

Un último ejemplo. A la pregunta «La información facilitada por el sistema es fácil de comprender» (con un promedio total de 4,04) el grupo de inglés lo valora con una media de 4,58, el de francés con 3,90, y el de italiano desciende a 3,22.)

Es decir, el sistema es claramente mejor valorado para todas las preguntas del cuestionario en el caso de hacer uso de él en inglés, que en francés o en italiano (siendo estos últimos usuarios los más descontentos).

Estos resultados, basados en la percepción y opiniones de los usuarios —perspectiva de evaluación centrada en el usuario— apoyan los resultados de un estudio anterior (Olvera-Lobo & Gutiérrez-Artacho, 2011a) en el que se aplicaron ocho medidas específicas para la evaluación de la calidad de los resultados ofrecidos por los sistemas de búsqueda de respuestas —desde la perspectiva de la evaluación centrada en el sistema—, lo que pone de manifiesto que, al obtener respuestas menos relevantes —como ocurre para las consultas en italiano— la valoración del sistema desciende en todos sus aspectos.

Si se analizan las correlaciones entre los cuatro factores contemplados, se observa que, 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). Esto constata el hecho de que, cuanto mayor sea el grado en que un usuario está satisfecho respecto a los resultados obtenidos mejor valora la calidad del servicio, y viceversa .

Correlaciones

			Facilidad de uso	Utilidad	Calidad del servicio	Calidad de la información
Rho de Spearman	Facilidad de uso	Coefficiente de correlación	1,000	,388(**)	,510(**)	,549(**)
		Sig. (bilateral)	.	,000	,000	,000
		N	122	122	121	120
	Utilidad	Coefficiente de correlación	,388(**)	1,000	,452(**)	,532(**)
		Sig. (bilateral)	,000	.	,000	,000
		N	122	122	121	120
	Calidad del servicio	Coefficiente de correlación	,510(**)	,452(**)	1,000	,668(**)
		Sig. (bilateral)	,000	,000	.	,000
		N	121	121	121	119
	Calidad de la información	Coefficiente de correlación	,549(**)	,532(**)	,668(**)	1,000
		Sig. (bilateral)	,000	,000	,000	.
		N	120	120	119	120

** La correlación es significativa al nivel 0,01 (bilateral).

Tabla 5. Correlaciones entre los cuatro factores evaluados

6. CONCLUSIONES

El estudio realizado ha permitido evaluar la satisfacción de los usuarios del ámbito de la traducción respecto al uso de los SBR multilingües como fuente de información terminológica. El análisis de los resultados de la evaluación mediante HONqa pone de manifiesto que el sistema les resulta a los usuarios fácil y útil para la recuperación de información. Esto es debido a que la interfaz de sistema es simple y fácil de manejar y su utilidad se aprecia enseguida.

No obstante, se muestran menos generosos al evaluar la calidad del servicio y la calidad de la información proporcionada. Esto podría explicarse por dos razones, por un lado, quizá se sienten algo inseguros al evaluar la calidad de la información ofrecida por el sistema, puesto que declaran tener un nivel bajo de conocimientos médicos. Pero, por otro lado, como puso también de manifiesto un estudio anterior que realizamos aplicando otro enfoque metodológico y una perspectiva diferentes de evaluación —en ese caso evaluación centrada en el sistema— (Olvera-Lobo & Gutiérrez-Artacho, 2011b) la calidad de las respuestas ofrecidas para consultas en lengua francesa e inglesa en HONqa es ciertamente mejorable, lo que probablemente hace descender el grado de satisfacción de los usuarios y el valor promedio para este factor. En efecto, 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 otorgado a los 17 ítems los valores más altos. A continuación le sigue los usuarios 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 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

- AJZEN, I. 1991. «The theory of planned behavior». *Organizational and Human Decision Processes* 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. J., BURGER, J. D., FERRO, L., HIRSCHMAN, L., HOUSE, D., LIGHT, M. 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, H., KAN, M.Y.; CHUA, T.S., XIAO, J. 2004. «A Comparative Study on Sentence Retrieval for Definitional Question Answering». *SIGIR Workshop on Information retrieval for Question Answering (IR4QA)*, Sheffield.
- CHARNIAK, E., ALTUN, Y., DE SALVO BRAZ, R., GARRETT, B., KOSMALA, M., MOSCOVICH, T., PANG, L., PYO, C., SUN, Y., WY, W., YANG, Z., ZEILER, S., ZORN, L. 2000. «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.; MCLEAN, E.R. 2003. «Model of Information Systems Success: A Ten-Year Update». *Journal of Management Information Systems*, 19:4, 9-30
- DIEKEMA, A. R. 2003. «Translation Events in Cross-Language Information Retrieval: Lexical ambiguity, lexical holes, vocabulary mismatch, and correct translations». Tesis Doctoral. University of Syracuse.
- FISHBEIN, M., AJZEN, I. 1980. *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- GREFENSTETTE, G. 1998. *Cross-Language Information Retrieval*. Kluwe academic publishers 1.
- JACKSON, P., SCHILDER, F. 2005. «Natural Language Processing: Overview». En *Brown: Encyclopedia of Language & Linguistics* 2, 503-518. Amsterdam: Elsevier Press.
- 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
- OLVERA-LOBO, M. D., GUTIÉRREZ-ARTACHO, J. 2011a. «Evaluation of Open- vs. Restricted- Domain Question Answering Systems in the Biomedical Field». *Journal of Information Science*, 37 (2), 152-162.
- OLVERA-LOBO, M. D., GUTIÉRREZ-ARTACHO, J. 2011b. «Multilingual Question-Answering System in Biomedical Domain on the Web: An evaluation». *Lecture Notes in Computer Science (LNCS)* 6941: 83:88
- ONG, C. C., DAY M., HSU M. 2009. «The measurement of user satisfaction with question answering systems». *Information & Management*, 46 (7): 397-403
- TESSIER, J. A., CROUCH, W. W., ATHERTON, P. 1977. «New Measures of User Satisfaction with ComputerBased Literature Searches,» *Special Libraries*, 68: 383-388
- TSUR, O. 2003. *Definitional Question-Answering Using Trainable Text Classifiers*. Tesis doctoral. Amsterdam: University of Amsterdam.
- VICEDO GONZÁLEZ, J. L., RODRÍGUEZ HONTORIA, H., PEÑAS PADILLA, A., MASSOT BAYÉS, M. 2003. «Los sistemas de Búsqueda de Respuestas desde una perspectiva actual». *Revista de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 31.
- VOORHEES, E. M. 2002. «Overview of the TREC 2002 Question Answering Track». *Proceedings of the Eleventh Text REtrieval Conference*. 69
- VOORHEES, E. M., TICE, D. 1999. *The TREC-8 question answering track evaluation*. Greece: Athens